Solvable Lattice Models: Algebraic and Combinatorial Theory

Daniel Bump and Anne Schilling

Preface

If you find mistakes or typos, please let us know so we can fix them. And if you do send feedback, please include the following version information, so we will know which draft you are using:

Git version: 8b5aeab Date: Wed Feb 18 09:11:01 AM PST 2026

Contents

CHAPTER 1

Origin in Statistical Mechanics

Lattice models were introduced in statistical mechanics in order to study realistic systems. Statistical mechanics will *not* be a focus of this book. Indeed, it turns out that solvable lattice models have important connections with representation theory, for example of quantum groups regarding the underlying mechanism, and other areas such as representation theory of p-adic groups, algebraic combinatorics, algebraic geometry, and conformal field theory. We will review the statistical mechanical origins of the theory, referring to Baxter 1982 for much more information, before turning away to other subjects. Section 4 will, however, set up some notation that will be used throughout the book.

1. Thermodynamics

The purpose of this section is to give a quick account of the origins of our subject in statistical mechanics. Since we will soon migrate away from this subject, readers can skip to Section 4 without loss of continuity.

Statistical mechanics is a development from thermodynamics. Thermodynamics was an empirical discovery which started with the theory of gases, motivated by considerations related to engines and refrigeration.

Thermodynamics can be *axiomatized* in the form of several laws, most importantly the *second law of thermodynamics* which contains a subtle and important concept, entropy. The laws of Thermodynamics are sometimes stated thus:

(1) Energy is conserved in a closed system.
(2) Entropy is increasing.
(3) If the temperature is decreased to zero, entropy approaches a fixed value, called the *residual entropy*.

The concept of entropy is of great importance, and universal in its surprising applicability to different areas, such as information theory and black holes. It has important philosophical implications, since it gives a direction to the arrow of time. This is paradoxical since the laws of physics are invariant under time reversal (CPT symmetry).

We take for granted the concept of energy, and its conservation. In thermodynamics and statistical mechanics, it is important to take into account both closed systems, that do not interact with their environment, and systems that do interact. Thus we imagine that energy can be put into a system, or extracted from it. *Work* can be described as energy that is extracted from a system, for example by operating a piston or generating electricity.

Heat is a form of energy that we now understand to be due to the kinetic energy of molecules in a substance. Carnot, whose investigations of the steam engine led to the concepts of thermodynamics, thought of heat as a fluid like water, that can flow from higher levels to lower, and in the process can be made to do work. The first law of thermodynamics

can be expressed in the formula

$$dU = dQ + dW,$$

where U is a variable expressing the total amount of energy in the system, Q is the amount of heat, and W is a variable expressing work, energy that is put into a system, or extracted from it.

As Carnot realized, certain processes are reversible. We may imagine a perfectly efficient engine, with frictionless parts, where energy is put in, in the form of fire or electricity, and mechanical work is extracted. But other processes, such as friction, are irreversible. In friction, work is transformed into heat, and this is energy that can never be extracted from the system. A processes involving friction is irreversible.

Again, if a system consists of two bodies of different temperatures, energy can be extracted as work by a mechanism such as a dipping bird. But if heat flows from one body to the other, until they reach the same temperature, the energy still exists, but can no longer be extracted as work. Thus the cooling of a hot object is an irreversible process.

The second law of thermodynamics regulates such irreversible processes. The second law postulates that there is a quantity S, called *entropy*, that can only increase. Irreversible processes are precisely those that increase the entropy. Conversely, a process is reversible if it does not increase entropy. If a system is at *maximal entropy*, the entropy can no longer increase. An example would be a system in thermal equilibrium, where all parts are at the same temperature.

Also related to the second law is the notion of *free energy*. This is the amount of energy that can be extracted from a system as useful work. Thus the entropy of the system is maximal if the free energy is zero.

2. Statistical mechanics and the partition function

The physical basis for thermodynamics is *statistical mechanics*. Thus heat is understood as being the kinetic energy of atoms and molecules, and the laws of thermodynamics can be derived from statistical considerations.

We will consider a system with many possible states, which is not strictly subject to the first law, in that not all states have the same energy. The source of this uncertainty is usually interaction with the environment. For example, one considers a system that is in contact with a heat bath at a constant temperature. The system itself is assigned a temperature that may be constant, or could vary within the medium. The system may also depend on other parameters, such as pressure or the strength of an applied electromagnetic field.

An important question that is investigated in statistical mechanics is the behavior of a system at a phase transition point. We may consider the melting or boiling of a substance as an example. In an idealized form, we may imagine the process as follow. In a "frozen" state, there are correlations between the local structure of the system at locations that are separated in distance, but in the "melted" form, there are no such correlations. The phase transition point or critical temperature is the value where the structure changes from frozen to melted.

A statistical mechanical system $\mathfrak{S}$ is an ensemble of *states*. Each state $\mathfrak{s}$ has an energy $e(\mathfrak{s})$, and there is a probability measure on $\mathfrak{S}$, with high energy states being less probable. The system may depend on some external parameters, notably the *temperature* of the system. The probability of the state $\mathfrak{s}$ with energy $E(\mathfrak{s})$ is proportional to $\beta(\mathfrak{s}) = e^{-E(\mathfrak{s})/kT}$, where k is *Boltzmann's constant*. Since the sum of the probabilities must be 1, the actual probability

is

$$\frac{1}{Z}\beta(\mathfrak{s}), \qquad Z = Z(\mathfrak{S}, T) := \sum_{\mathfrak{s}} \beta(\mathfrak{s}).$$

The quantity $\beta(\mathfrak{s})$ is called the *Boltzmann weight* of the state, and the quantity Z is called the *partition function*. Note that as the temperature increases, energetic states become more probable.

The partition function is very important in statistical mechanics, since it controls characteristics of the system such as energy and entropy, and how they depend on temperature and other parameters. For example the mean energy is

$$\langle E \rangle := \frac{1}{Z} \sum_{\mathfrak{s}} \beta(\mathfrak{s})\, E(\mathfrak{s}) = kT^2 \frac{\partial}{\partial T} \log(Z).$$

The free energy, which we recall is the amount of energy that can be extracted from the system as work, equals

$$F = -kT\, \log(Z),$$

and the entropy is

$$S = k\, \log(Z) + \frac{1}{kT}\langle E \rangle.$$

If the partition function depends on other parameters such as a magnetic field strength, differentiating with respect to those will yield other values of significance.

The partition function also occurs in other areas of physics, such as quantum field theory. For us, the partition function will be a main object of study, even though we will soon leave its origins in statistical mechanics behind.

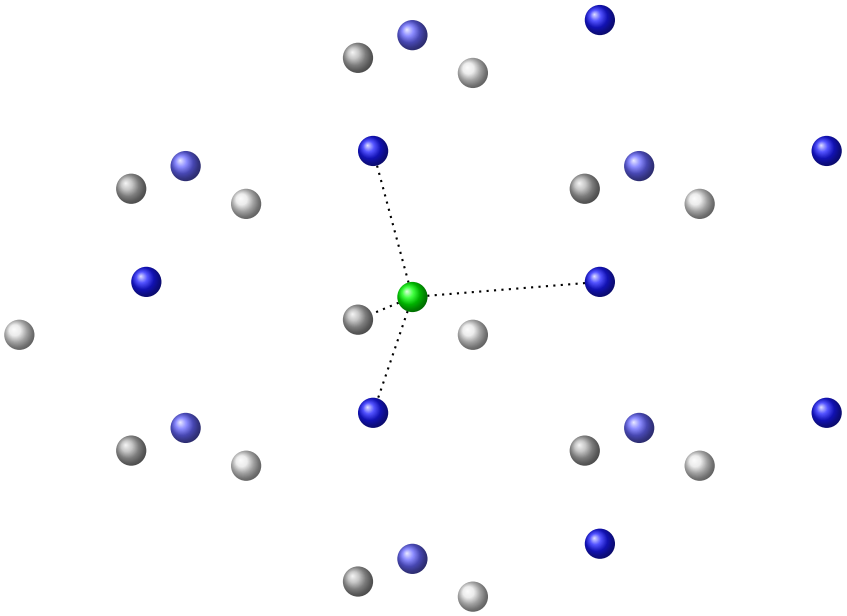
3. Ice

We may consider ice (frozen H_2O), where the larger oxygen atoms have fixed locations at the vertices of a grid. In its usual form (called Ice I_h) these oxygen atoms are arranged in a three-dimensional hexagonal lattice. We can envision the oxygen atoms as lying on the vertices of a three-dimensional hexagonal crystal lattice. Each oxygen atom will have four neighbors, lying at the vertices of a tetrahedron. We may consider the 4-regular graph Γ whose vertices are the oxygen atoms and whose edges are the segments joining them to the four nearby atoms.

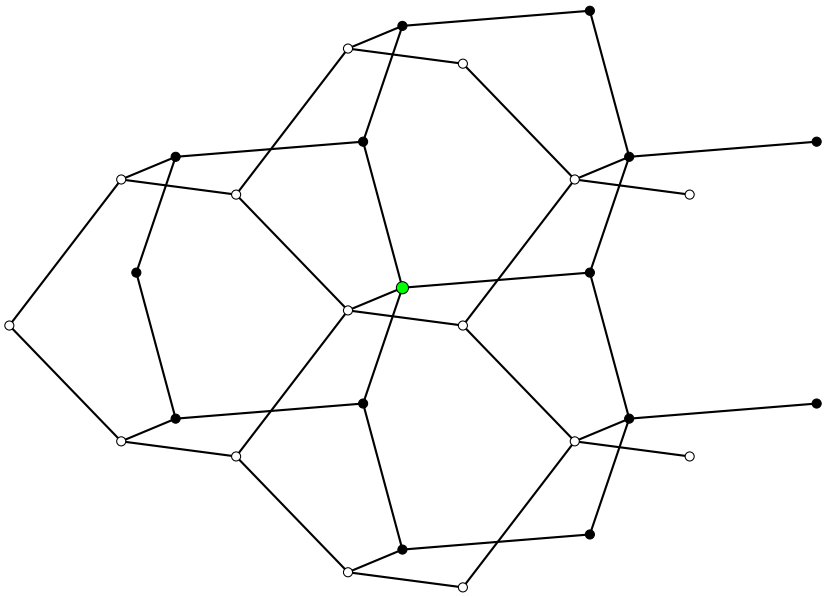
Linus Pauling Pauling 1935 computed the entropy and free energy of ice by means of a three-dimensional lattice model. Let us describe a grid whose vertices are the oxygen atoms in a crystal. We consider two oxygen atoms *adjacent* if they share a hydrogen bond. They then form a graph Γ that is nearly 4-regular in that each oxygen atom, except those at the boundary of the crystal, have 4 neighbors. (Here we are ignoring a detail about boundary edges, and we will give a proper discussion of Γ below in Section 4.)

Ice has many possible crystalline structures. Under normal conditions, Ice I_h is the usual one. This crystal occurs in sheets or layers. The graph is bivalent. Each layer is tesselated by hexagons, with oxygen atoms at their vertices. Furthermore, each atom has a bond with one in either the layer above or below, depending on its valence.

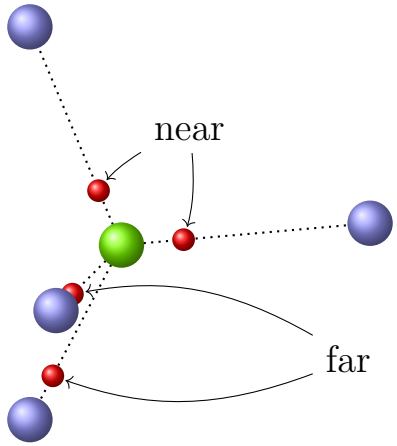
Here is the hexagonal Ice I_h lattice, showing the segments joining a sample oxygen atom (green) to its four neighbors.



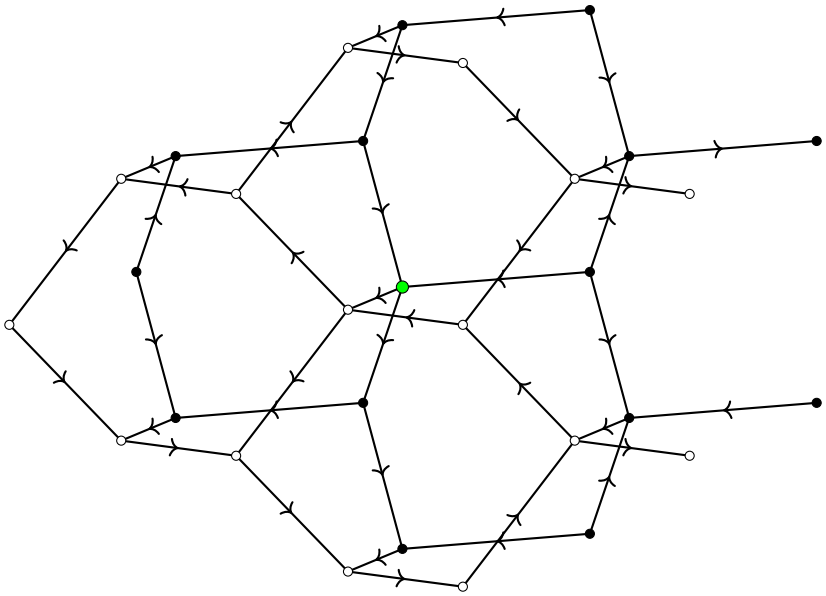
Here is the graph Γ showing two adjacent layers.



While the location of the oxygen atoms is fixed, and forced into a crystalline pattern, the location of the hydrogen nuclei (protons) is another matter. Due to its position in the periodic table, oxygen is allowed two covalent bonds. The oxygen atom will therefore borrow electrons from two hydrogen atoms. This causes the protons to lie on the segments between two adjacent oxygen atoms, but each proton will be closer to one or the other of the two oxygen atoms. There are many possible configurations, which are subject to quantum superposition.



We may represent this graphically by making the graph Γ into a *directed* graph. We decorate the edges with arrows, each pointing towards the hydrogen atom on the edge.



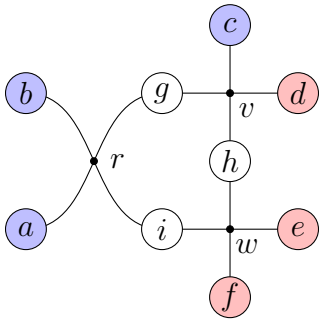
Then we obtain the following model: we have a 4-regular graph, based on the three-dimensional hexagonal lattice. A state of the system is a refinement of the graph to a directed graph, with every vertex having two incoming and two outgoing arrows.

4. A class of lattice models

4.1. Graphs. We have formalized the ice crystal into a system based on a *graph*, which is almost but not the same as a graph in the usual combinatorial definition. Let us define a graph to be a set of *vertices* and a set of *edges* with an *incidence relation*. This means that some edges are *adjacent to* or *through* certain vertices. We will assume that every edge is through either exactly two vertices, or a single vertex. The edges that are through a single vertex will be called *boundary edges*. The edges that connect two vertices are *interior edges*.

As an example, let us consider this graph:

(1.1)



Here we have a graph with three vertices, labeled v, w and r. There are nine edges, labeled a, b, c, d, e, f and g, h, i. The edges a, b, c, d, e, f are boundary edges. For later reference we have colored the boundary edges. (See Section 6.)

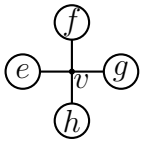
The graph is *planar* if it can be embedded in the plane. We will consider mainly planar graphs.

4.2. Spins and edge types. In the class of models we will consider, every edge e will be assigned a set Σ_e called its *spinset*. Elements of Σ_e will be called *spins*. We will also denote by V_e the *spinspace* which is by definition the free vector space on Σ_e. A *state* of the model is a function that assigns to every edge e an element of Σ_e.

We will require that the spins of the boundary edges are fixed, and are part of the data describing the system. On the other hand, the spins of the interior edges are variable. Two edges e, f may be classified as the same *type* if the spin sets Σ_e and Σ_f are equal.

For example, in the Ice model, every edge has two possible configurations, and we can take the spinsets $\Sigma_e = \{+, -\}$. In this example, all edges have the same type.

Two edges on opposite sides of a vertex will (usually) have the same spinset. Thus in a configuration



we require $\Sigma_e = \Sigma_g$ and $\Sigma_f = \Sigma_h$. Thus e and g have the same edge type, and f, h have the same edge type. Consequently all edges on a single line through the configuration have the same edge type.

4.3. States. A *state* of the model is an assignment of an element of its spinset to every edge of the model. We will assume that the boundary spins have fixed assignments. Indeed, this will be part of the data describing the model.

Almost always there will be *local constraints* at each vertex on the possible configurations of spins adjacent to a particular vertex. We will call a state in which these constraints are satisfied at every vertex *admissible*.

For example in the Ice I_h model that we have described, the spins are directions or orientations of the edges, which we can represent by arrows, and the constraint is that there there are two "in" arrows and two "out" arrows. This means that there are $\binom{4}{2} = 6$ possible configurations of local spins at the vertex.

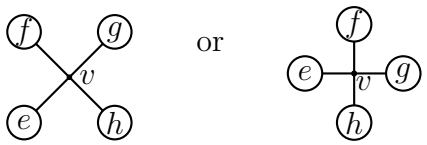
4.4. Local Boltzmann weights. Every vertex v is adjacent to four edges, which we label e, f, g, h. We assume that there is a given map

$$\beta_v \colon \Sigma_e \times \Sigma_f \times \Sigma_g \times \Sigma_h \longrightarrow \mathbb{C}$$

that assigns to every configuration of spins on these adjacent vertices a value, called the *Boltzmann weight*. If $(a, b, c, d) \in \Sigma_e \times \Sigma_f \times \Sigma_g \times \Sigma_h$, we may denote the value of $\beta_v(a, b, c, d)$ as

$$\beta_v \begin{pmatrix} b & c \\ a & d \end{pmatrix} \qquad \text{or} \qquad \beta_v \begin{pmatrix} & b & \\ a & & c \\ & d & \end{pmatrix}$$

depending on the orientation of the vertex:



In applications to statistical mechanics or probability this would be a nonnegative real number, but we will not assume this. The local constraints on the spins at the vertex may be formulated as the assumption that $\beta_v(a, b, c, d) \neq 0$.

4.5. Vertex types. Two vertices v, v' in the same or different systems are *equivalent* if there is a bijection between the edges adjacent to v and the edges adjacent to v' that transports the local Boltzmann weights at v to the local Boltzmann weights at v'. A *vertex type* is an equivalence class of vertices.

4.6. Partition function. Every admissible configuration $\mathfrak{s}$ is therefore a state of the system $\mathfrak{S}$, which is the ensemble of all states. In such a state, if v is a vertex and e, f, g, h are the adjacent edges, the state assigns spins in $\Sigma_e, \Sigma_f, \Sigma_g, \Sigma_h$, and by abuse of notation we will denote by $\beta_v(\mathfrak{s})$ the Boltzmann weight at v determined by these four spins. Now define

$$\beta(\mathfrak{s}) = \prod_v \beta_v(\mathfrak{s}).$$

The state is admissible if $\beta(\mathfrak{s})$ is nonzero. The partition function is

$$Z(\mathfrak{S}) = \sum_{\mathfrak{s}} \beta(\mathfrak{s}).$$

We may sum over all states, or over admissible states.

Remark 1.1. In our pictures, we will often decorate the edges of the graph with a "bubble" that can contain a symbol. But for narrative flexibility the contents of the bubble will vary with context. It may be the name of the edge. It may be a symbol denoting an element of the spinset, for example to specify the boundary conditions. It may also be a vector space, typically the free vector space on the spinset, or a module of a quantum group. This should cause no confusion since the meaning of the symbol in the bubble is defined somewhere in the text.

5. The six-vertex model

Certain lattice models are called *solvable* since algebraic methods based on the Yang–Baxter equation, which will be a major focus of this book, allow the partition function to be computed exactly. Historically the first example was Onsager's 1944 study of the 2-dimensional Ising model. However we will start with an even simpler model, the *six-vertex model*, which is also related to ice.

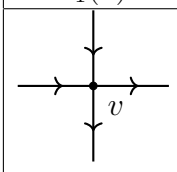
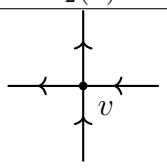
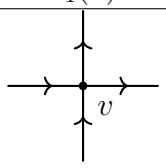
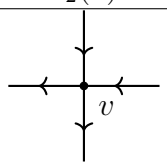
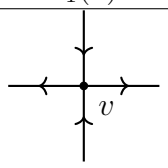
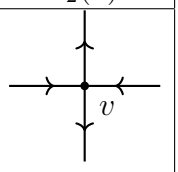
Solvable lattice models are almost exclusively 2-dimensional. This means that the underlying graph is planar. The Ice I_h model that we considered is not solvable as far as we know, and its graph is not planar.

While Pauling had considered the realistic problem of 3-dimensional Ice and heuristically computed the number of states, one can also consider 2-dimensional Ice, in which the oxygen atoms are restricted to a plane, and form a crystal with the oxygen atoms at the vertices of a square lattice. This was investigated by NagleNagle 1966, after which LiebLieb 1967a,b,c and Sutherland Sutherland 1967 found exact solutions for the entropy problem. Baxter introduced the *Yang–Baxter equation* and applied it to Ice-type models, as well as the more difficult eight-vertex model. (See Baxter 1982, Chapter 9.)

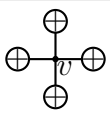
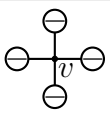
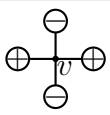
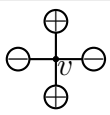
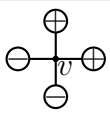
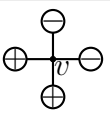
For 2-dimensional Ice, Lieb found that the residual entropy was $kN\log(W)$ with $W = (4/3)^{3/2}$.

The mathematical model of 2-dimensional ice is the famous six-vertex model, which is the archetype of a large class of important solvable lattice models. It is realistic enough to have a phase transition, which was of great interest to the early investigators. We will therefore discuss it at length.

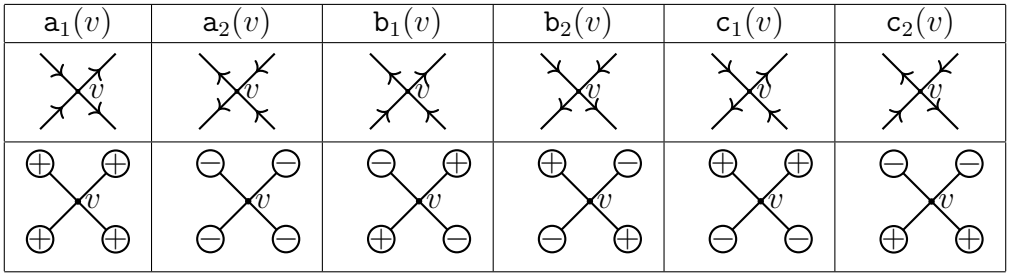
The six-vertex model is nearly identical to the I_h models we have discussed, except that the underlying crystal is 2-dimensional, based on a square lattice. We will give two versions of the Boltzmann weights. Recall that the spinset of an edge is a set of possible states. For the six-vertex model, the spinset has cardinality two. In one version of the six-vertex model, the spinset of an edge is an orientation. The Boltzmann weights depend on six parameters, $\mathtt{a}_1$, $\mathtt{a}_2$, $\mathtt{b}_1$, $\mathtt{b}_2$, $\mathtt{c}_1$ and $\mathtt{c}_2$, which may depend on the vertex v, so we may write $\mathtt{a}_1(v)$, etc. We label the possible states as follows:

$\mathtt{a}_1(v)$	$\mathtt{a}_2(v)$	$\mathtt{b}_1(v)$	$\mathtt{b}_2(v)$	$\mathtt{c}_1(v)$	$\mathtt{c}_2(v)$
					

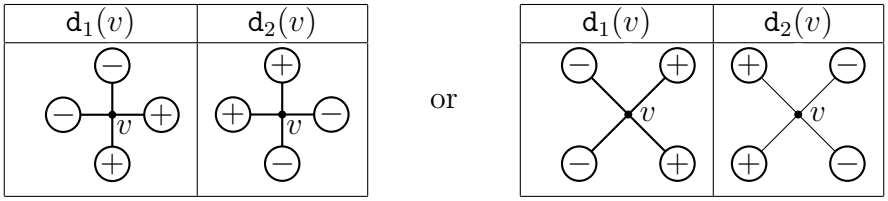
On the other hand, it will also be convenient to dispense with the orientation and take the spinset to be the 2-element set $\{\oplus, \ominus\}$. Then the labeling of the states is as follows:

$\mathtt{a}_1(v)$	$\mathtt{a}_2(v)$	$\mathtt{b}_1(v)$	$\mathtt{b}_2(v)$	$\mathtt{c}_1(v)$	$\mathtt{c}_2(v)$
					

Although the lattice model will be based on a rectangular grid, we will also encounter vertices that are in a rotated orientation, and we will use the following labels for these.



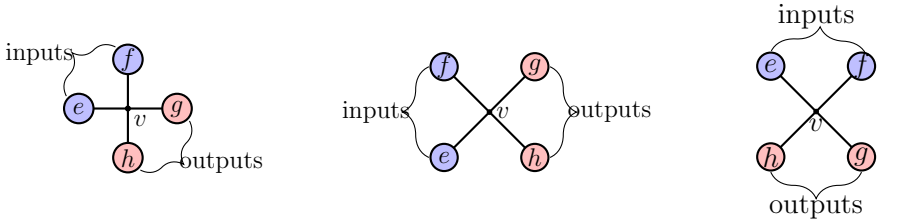
Similar to the six-vertex model, but more difficult, is the *eight-vertex model*. This adds two more admissible configurations:



The eight vertex model is more difficult than the six-vertex model, but nevertheless Baxter was able solve it, simultaneously solving a problem in quantum mechanics, the analysis of the XYZ Hamiltonian. See Baxter 1982, Chapter 10.

6. Boltzmann weights as linear transformations

It will be advantageous later to regard the Boltzmann weights as linear transformations. We will usually draw the vertices in one of the following orientations:



Note that we have classified some of the edges as inputs and some as outputs. We've colored the input edges blue and the output edges red. We shall adopt the following rule: when two vertices are connected by an edge, input edges are allowed to connect to output edges, but inputs cannot be connected to inputs, and outputs cannot be connected to outputs. Thus when an edge connects two vertices v and w, it is an output edge for v if and only if it is an input edge for w. We will put the label for a vertex between the two output edges.

Remark 1.2. The division of edges into inputs and outputs extends to boundary edges of larger graphs. For example in (1.1) we have chosen to consider the boundary edges a, b and c as inputs, and correspondingly we have colored them blue. The boundary edges d, e and f are outputs and we have colored them red. On the other hand the interior edges g, h and i have ambiguous classification. Indeed g is an output edge for the vertex r and an input edge for s, so we do not consider it either input or output for the overall graph.

As explained in Section 4.2, the edges e and g have the same spinsets, i.e. $\Sigma_e = \Sigma_f$, and similarly $\Sigma_f = \Sigma_h$. Let V be the free vector space on the spinset $\Sigma_e = \Sigma_g$, and let W be the free vector space on $\Sigma_f = \Sigma_h$. We think of the spins attached to e and f as "inputs"

and those attached to g and h as outputs. Then we may define a linear transformation $R_v : V \otimes W \longrightarrow W \otimes V$ such that $\beta_v(x, y, z, w)$ is the coefficient of $w \otimes z$ in $R_v(x \otimes y)$, when $x \in \Sigma_e$, $y \in \Sigma_f$, $z \in \Sigma_h$ and $w \in \Sigma_g$. We may sometimes abuse notation and denote R_v as just v when it simplifies the notation.

We will sometimes use *Dirac notation*, which we now explain. The vector space V comes with a distinguished basis $\Sigma_e = \Sigma_g$, and so the dual V^* has the dual basis Σ_e^*. If $x \in \Sigma_e$, let $x^* \in V^*$ be the corresponding linear functional. Alternatively, we may denote x as $|x\rangle$ when we want to emphasize that it is in the spinset of an input edge, say Σ_e, and we will then denote x^* as $\langle x|$, which we regard as an element of the output edge spinset Σ_g. If additionally $y \in \Sigma_f$ then we write $|x \otimes y\rangle$ instead of $|x\rangle \otimes |y\rangle$. Then with this notation

$$\beta(x, y, z, w) = \langle w \otimes z | R_v | x \otimes y \rangle, \qquad x, w \in V,\ y, z \in W.$$

Now the output edges g and h are (unless boundary edges) also be input edges for other vertices in the grid, so the notation $|w \otimes v\rangle$ also has meaning. We will also write:

$$R_v|x \otimes y\rangle = \sum_{\substack{w \in \Sigma_h \\ z \in \Sigma_g}} \beta(x, y, z, w)|w \otimes z\rangle.$$

Let us consider six-vertex model. In this case all four edges have the same spinset $\{\oplus, \ominus\}$. The free vector space V on the spinset $\{\oplus, \ominus\}$ may be identified with $\mathbb{C}^2$. Let v_+ and v_- be the standard basis v_+ and v_- be the standard basis. We may then use the basis

$$\{v_+ \otimes v_+, v_- \otimes v_+, v_+ \otimes v_-, v_- \otimes v_-\}$$

for $V \otimes V$, and we find that the six-vertex model vertex v corresponds to the linear transformation

$$R_v = \begin{pmatrix} \mathsf{a}_1(v) & & & \\ & \mathsf{c}_1(v) & \mathsf{b}_1(v) & \\ & \mathsf{b}_2(v) & \mathsf{c}_2(v) & \\ & & & \mathsf{a}_2(v) \end{pmatrix}. \tag{1.2}$$

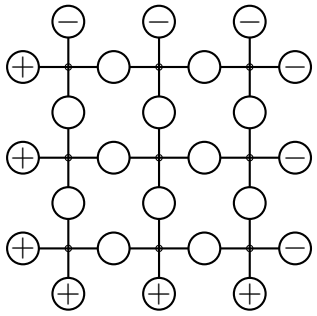
In Dirac notation we have:

$$\begin{aligned} a_1(v) &= \langle v_+ \otimes v_+ | R_v | v_+ \otimes v_+ \rangle, & a_2(v) &= \langle v_- \otimes v_- | R_v | v_- \otimes v_- \rangle, \\ b_1(v) &= \langle v_- \otimes v_+ | R_v | v_+ \otimes v_- \rangle, & b_2(v) &= \langle v_+ \otimes v_- | R_v | v_- \otimes v_+ \rangle, \\ c_1(v) &= \langle v_- \otimes v_+ | R_v | v_- \otimes v_+ \rangle, & b_2(v) &= \langle v_+ \otimes v_- | R_v | v_+ \otimes v_- \rangle. \end{aligned}$$

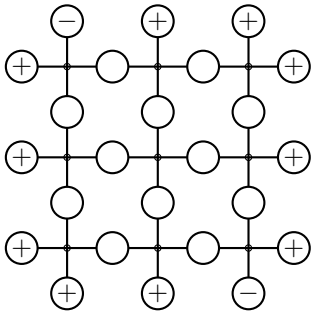
Exercises

Exercise 1.1. Check that (1.2) is the right matrix for the linear transformation of the six-vertex model v. What is the matrix for the eight-vertex model?

Exercise 1.2. Here are two models based on a 3×3 grid.



Model 1



Model 2

(a) For Model 1, prove that this model has 7 states for the six-vertex model, and 16 states for the eight-vertex model.

(b) For Model 2, prove that this model has 6 states for the six-vertex model, and 16 states for the eight-vertex model.

(c) Consider a model based on this 3×3 grid with an arbitrary distribution of $\oplus$ and $\ominus$ spins on the 12 boundary edges. Prove that the model has 16 states in the eight-vertex model if the number of boundary $\ominus$ spins is even, and no states if the number of $\ominus$ spins is odd.

CHAPTER 2

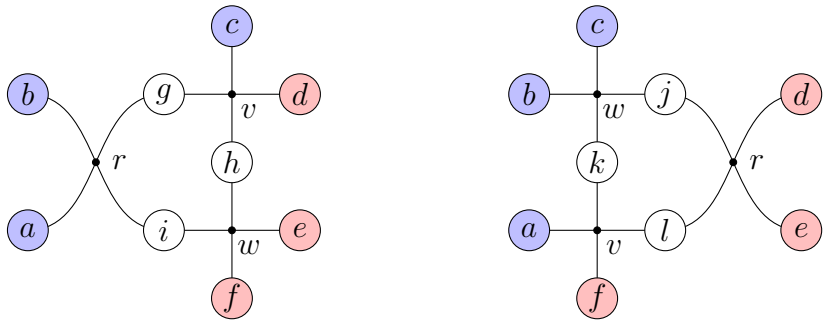
The Yang–Baxter Equation

1. Solvability

Baxter introduced an important method of studying certain vertex models, and he used it to solve not just the six-vertex model, but the more difficult eight-vertex model, and with it the XYZ Heisenberg spin chain, a related quantum mechanical problem. This method is based on the *Yang–Baxter equation*, so named by Faddeev. The study of the Yang–Baxter equation leads to interesting mathematics, namely braided categories and quantum groups. The same technology explains knot invariants such as the Jones polynomial. The six-vertex model was a key example leading to this mathematics.

We resume the discussion from Section 5 of Chapter 1. We will say that a class of models is *solvable* if, when v and w are vertex types that can occur adjacent to each other in the class, there is another vertex type that we will denote r such that the two following systems are equivalent.

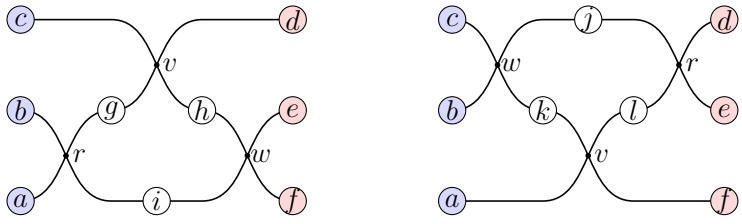
(2.1)



This means that for every possible assignment of spins to the six boundary edges a, b, c, d, e, f, the partition functions of the two systems are equivalent. Thus we sum over all possible assignments of spins to the interior edges, g, h, i on the left-hand side, or j, k, l on the right-hand side. If this is so, we say the *Yang–Baxter equation* is satisfied for these vertices v, w, r.

In (2.1), one vertex r is "rotated." We can rotate the other two vertices, and consider instead the equality of the two following systems:

(2.2)

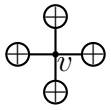
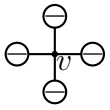
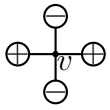
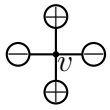
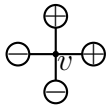
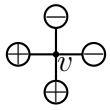


We will refer to a vertex in a rotated orientation as an *R-vertex*, and a vertex (such as v and w in (2.1)) aligned parallel to the coordinate axes as a *T-vertex*. So the Yang-Baxter equation in (2.1) contains one R-vertex and two T-vertices. Therefore we refer to this as an *RTT*

equation. In (2.2), there are three R-vertices, so we will refer to this as an *RRR equation*. Mathematically, the RTT and RRR equations are equivalent. But there is a difference in how the R-vertices and T-vertices are used: typically the T-vertices are assembled into a grid, and the R-vertices are attached to the grid to obtain information about the partition function.

One goal of this Chapter is to introduce this subject by showing how, following Baxter Baxter 1982, Yang–Baxter equations can be used to prove global properties of the partition function. An archetypal example that we will discuss is the commutativity of the row-transfer matrices for a model. We begin by introducing one family of solutions to the Yang–Baxter equation within the six-vertex model.

1.1. Field-free six-vertex model. The vertex v with Boltzmann weights $\mathsf{a}_i(v)$, $\mathsf{b}_i(v)$ and $\mathsf{c}_i(v)$ is called *field-free* if, in the notation of Chapter 1, Section 5 we have $\mathsf{a}_1(v) = \mathsf{a}_2(v)$, $\mathsf{b}_1(v) = \mathsf{b}_2(v)$ and $\mathsf{c}_1(v) = \mathsf{c}_2(v)$. We will suppress the subscript in the field free case and write just $\mathsf{a}(v) = \mathsf{a}_1(v) = \mathsf{a}_2(v)$, etc.

$\mathsf{a}(v)$	$\mathsf{a}(v)$	$\mathsf{b}(v)$	$\mathsf{b}(v)$	$\mathsf{c}(v)$	$\mathsf{c}(v)$
					

Let v be a field-free vertex. If $\mathsf{a}(v)$ and $\mathsf{b}(v)$ are nonzero, let

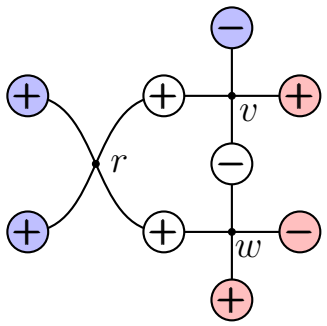
$$\Delta(v) = \frac{\mathsf{a}^2(v) + \mathsf{b}^2(v) - \mathsf{c}^2(v)}{2\mathsf{a}(v)\,\mathsf{b}(v)}. \tag{2.3}$$

We say that $\Delta(v)$ is *defined* if $\mathsf{a}(v)$ and $\mathsf{b}(v)$ are nonzero.

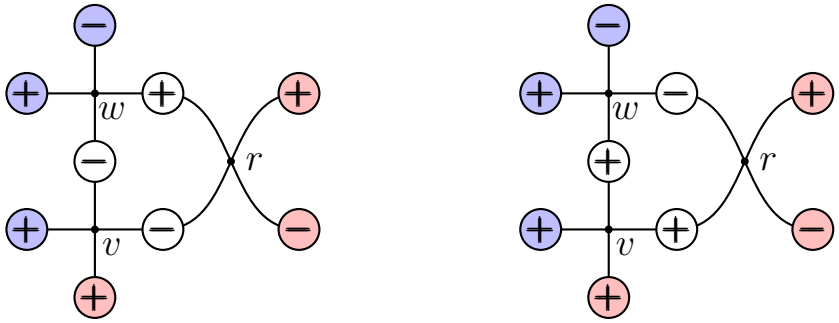
The case $\Delta = 0$ is called *free-fermionic*, and we will discuss it later (without the field-free condition). The cases $\Delta = \pm 1$ are also special since they are phase transition points (Baxter 1982, Section 8.10). For an interesting and important generalization, see Baxter's analysis of the field-free eight-vertex model in Baxter 1982, Section 10.4.

Theorem 2.1 (Baxter)**.** *Let v and w be field-free vertices such that $\Delta(v)$ and $\Delta(w)$ are both defined. Suppose that $\Delta(v) = \Delta(w)$. Then there exists a nonzero field-free vertex r such that the Yang–Baxter equation (2.1). If $\Delta(r)$ is defined, then $\Delta(r) = \Delta(v) = \Delta(w)$.*

Proof. There are three equations that must be satisfied for the Yang–Baxter equation to be satisfied. First take $(a, b, c, d, e, f) = (+, +, -, +, -, +)$ in (2.1). The left-hand side of the Yang–Baxter equation has one admissible state:



This has Boltzmann weight $\mathtt{b}(v)\mathtt{c}(w)\mathtt{a}(r)$. On the other hand, there are two admissible states on the right-hand side:



These have weights $\mathtt{c}(v)\mathtt{b}(w)\mathtt{c}(r)$ and $\mathtt{a}(v)\mathtt{c}(w)\mathtt{b}(r)$. So we obtain the equation

$$\mathtt{b}(v)\mathtt{c}(w)\mathtt{a}(r) = \mathtt{c}(v)\mathtt{b}(w)\mathtt{c}(r) + \mathtt{a}(v)\mathtt{c}(w)\mathtt{b}(r). \tag{2.4}$$

Taking $(a,b,c,d,e,f) = (+,+,-,-,+,+)$ gives

$$\mathtt{c}(v)\mathtt{a}(w)\mathtt{a}(r) = \mathtt{c}(v)\mathtt{b}(w)\mathtt{b}(r) + \mathtt{a}(v)\mathtt{c}(w)\mathtt{c}(r), \tag{2.5}$$

and taking $(a,b,c,d,e,f) = (+,-,+,-,+,+)$ gives

$$\mathtt{b}(v)\mathtt{a}(w)\mathtt{c}(r) = \mathtt{c}(v)\mathtt{c}(w)\mathtt{b}(r) + \mathtt{a}(v)\mathtt{b}(w)\mathtt{c}(r). \tag{2.6}$$

Taking other combinations of a,b,c,d,e,f give a total of 12 equations altogether, but they turn out to be these same three equations, repeated. So we need to show that we can construct the vertex r to satisfy (2.4–2.6).

Since we are assuming that $\Delta(v) = \Delta(w)$ we have

$$(\mathtt{a}(v)^2 + \mathtt{b}(v)^2 - \mathtt{c}(v)^2)\mathtt{a}(w)\mathtt{b}(w) = (\mathtt{a}(w)^2 + \mathtt{b}(w)^2 - \mathtt{c}(w)^2)\mathtt{a}(v)\mathtt{b}(v).$$

Since $\Delta(v)$ and $\Delta(w)$ are defined, both $\mathtt{a}(w), \mathtt{b}(v) \neq 0$ implying

$$\frac{\mathtt{b}(v)\mathtt{a}(w)\mathtt{b}(w) - \mathtt{a}(v)\mathtt{b}(w)^2 + \mathtt{a}(v)\mathtt{c}(w)^2}{\mathtt{a}(w)} = \frac{\mathtt{a}(v)\mathtt{b}(v)\mathtt{a}(w) - \mathtt{a}(v)^2\mathtt{b}(w) + \mathtt{c}(v)^2\mathtt{b}(w)}{\mathtt{b}(v)},$$

and we define this to be $\mathtt{a}(r)$. Then we define

$$\mathtt{b}(r) = \mathtt{b}(v)\mathtt{a}(w) - \mathtt{a}(v)\mathtt{b}(w), \qquad \mathtt{c}(r) = \mathtt{c}(v)\mathtt{c}(w).$$

Now it may be checked that the identities (2.4), (2.5), (2.6) are satisfied. For example, to prove (2.4), the right-hand side equals

$$\mathtt{c}(w)(\mathtt{a}(v)\mathtt{b}(v)\mathtt{a}(w) - \mathtt{a}(v)^2\mathtt{b}(w) + \mathtt{c}(v)^2\mathtt{b}(w))$$

and using the second expression for $\mathtt{a}(r)$ this equals $\mathtt{a}(r)\mathtt{c}(w)\mathtt{b}(v)$. We leave the other two cases to the reader. We leave the fact that $\Delta(r) = \Delta(v) = \Delta(w)$ to the reader (Exercise 2.1). □

We will explain later how this Yang–Baxter equation can be applied to study the partition functions for the field-free two-dimensional ice models, and what some of the applications are.

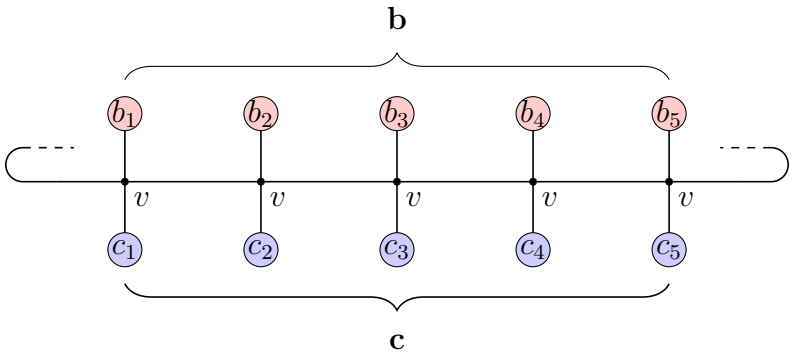
2. Commuting transfer matrices

We will consider systems $\mathfrak{S}$ built up from graphs Γ as in Chapter 1. Recall that a *graph* for us consists of vertices and edges, with an incidence relation between them. Every edge is adjacent to one or two vertices. An edge that is adjacent to two vertices is called *interior*, and an edge that is adjacent to only one vertex is called a *boundary* edge. Every edge $\mathcal{E}$ is assigned a *spinset* $\Sigma_{\mathcal{E}}$ of possible states called *spins*. The spins of boundary edges are fixed, and are part of the data defining the system. A state of the system consists of an assignment of spins to the interior edges.

Also required for the specification of the system $\mathfrak{S}$ is, for every vertex $v \in \Gamma$ a rule β that assigns to a state $\mathfrak{s}$ and a vertex v a weight $\beta_v(\mathfrak{s})$. This should only depend on the spins of the edges adjacent to v. The Boltzmann weight $\beta(\mathfrak{s})$ is the product of the $\beta_v(\mathfrak{s})$ over all vertices, and the partition function is

$$Z(\mathfrak{S}) = \sum_{\text{states } \mathfrak{s}} \beta(\mathfrak{s}).$$

We wish to discuss the concatenation of two systems. To have an example in mind, consider a system consisting of a single row of vertices:



Every vertex has the same Boltzmann weight β_v, so we give every vertex the same label. We imagine that one edge wraps around the back, so the system is periodic. We will refer to this as *cyclindric boundary conditions*.

For simplicity, assume that every vertical edge here has the same spinset Σ. Let V be the spinspace, which we recall is the free vector space on Σ. Let $b_i, c_i \in \Sigma$ be the boundary spins. We may collect the data $\mathbf{b} = (b_1, \ldots, b_n) \in \Sigma^n$ and $\mathbf{c} = (c_1, \ldots, c_n) \in \Sigma^n$ into data representing the boundary spins on the top and bottom edges. The free vector space on Σ^n is naturally isomorphic to $\otimes^n V$, so we may also denote $\mathbf{b} = b_1 \otimes \cdots \otimes b_n$ by abuse of notation.

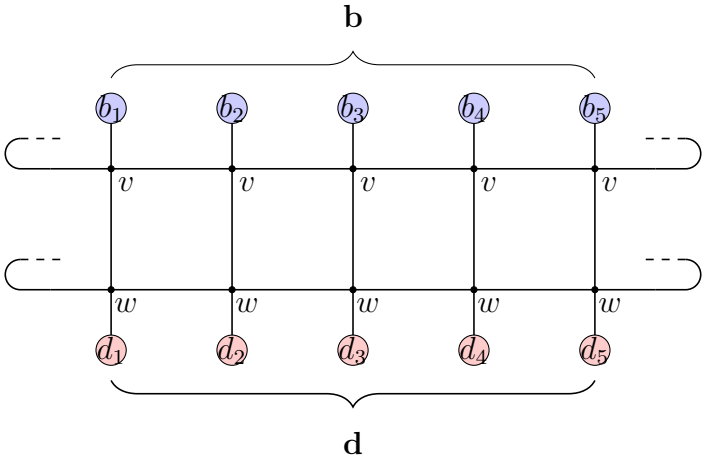
We have already introduced Dirac notation in Section 6 of Chapter 1. Let us see how it is applicable here. To quickly review, let W be a *frame*, by which we mean a vector space with a distinguished basis Φ. The dual space W^* then contains the dual basis Φ^*. If $\beta \in \Phi$, we sometimes denote β as $|\beta\rangle$ and the corresponding element of Φ^* as $\langle\beta|$. Then if $T : W \longrightarrow W$ is a linear transformation, and if $\beta, \gamma \in \Phi$, we may apply T to $|\beta\rangle$ to obtain $T|\beta\rangle$. Then we may apply the linear functional $\langle\gamma|$ to this vector to obtain $\langle\gamma|T|\beta\rangle$. With this notation, the adjoint operator $T^* : W^* \to W^*$ is the map that sends $\langle\gamma|$ to $\langle\gamma|T$. Dirac invented this notation for Hermitian or self-adjoint operators that are prevalent in quantum mechanics, where it is natural to use the same symbol T to denote an operator and its adjoint.

Applying this with $W = \otimes^n V$, we may consider the partition function of this one-rowed system to be the matrix of an operator $T_v : W \to W$ and write

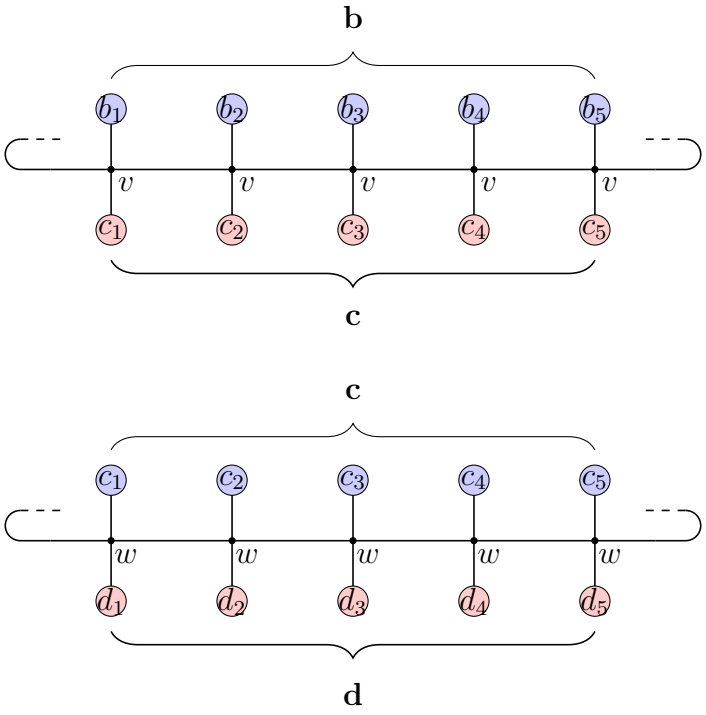
$$Z(\mathfrak{S}) = \langle \mathbf{c} | T_v | \mathbf{b} \rangle.$$

We call T_v the *row transfer matrix*. Note that thinking of the partition function as an operator this way treats the top boundary edges as inputs, and the bottom boundary edges as outputs.

One of Baxter's great insights was the use of the Yang–Baxter equation to prove that under certain conditions, row transfer matrices commute. To see how this works, let w be another vertex type. Consider a system with two layers:



We can express this in terms of the product $\langle \mathbf{d} | T_w T_v | \mathbf{b} \rangle$ of two row transfer matrices. We may concatenate the two smaller systems:



Now the common edges, labeled $\mathbf{c}$ in both cases have become interior edges, so by our rules, we have to sum over the possible states, to obtain:

$$\sum_{\mathbf{c}} \langle \mathbf{d}|T_w|\mathbf{c}\rangle\langle \mathbf{c}|T_v|\mathbf{b}\rangle = \langle \mathbf{d}|T_wT_v|\mathbf{b}\rangle,$$

consistent with the usual rule for matrix multiplication. This explains the term "row transfer matrix."

In preparation for applying the Yang–Baxter equation, let $a, b, c \in \mathbb{C}$, and let $v = v(a, b, c)$ denote the corresponding vertex type for the field-free six-vertex model. Thus the Boltzmann weights of this vertex type are such that $\mathtt{a}_1(v) = \mathtt{a}_2(v) = a$, $\mathtt{b}_1(v) = \mathtt{b}_2(v) = b$ and $\mathtt{c}_1(v) = \mathtt{c}_2(v) = c$.

Theorem 2.2 (Baxter)**.** *Let $\Delta \in \mathbb{C}^\times$, and let a, b, c, a', b', c' be such that*

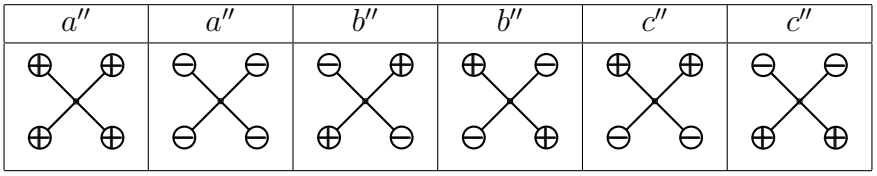
$$\frac{a^2 + b^2 - c^2}{2ab} = \frac{(a')^2 + (b')^2 - (c')^2}{2a'b'} = \Delta.$$

Let $v = v(a, b, c)$ and $w = v(a', b', c')$ be the two corresponding vertex types. Then the corresponding row transfer matrices commute:

$$T_wT_v = T_vT_w.$$

We should think of this in the context of "diagonalizing" the matrix T_v, for it is often easier to diagonalize a large family of commuting operators than a single operator.

Proof. To prove this, we will make use of the Yang–Baxter equation, with the R-vertex r from the last section. This is the vertex $v(a'', b'', c'')$ which we draw in a rotated orientation, thus:



where

$$a'' = \frac{ba'b' - a(b')^2 + a(c')^2}{a'} = \frac{aba' - a^2b' + c^2b'}{b},$$

$$b'' = ba' - ab', \qquad c'' = cc'.$$

We recall that also

$$\frac{(a'')^2 + (b'')^2 - (c'')^2}{2a''b''} = \Delta.$$

Then r is invertible in the following sense. As in Section 6 in Chapter 1, we may encode the Boltzmann weights of the vertex r by the linear transformation with matrix

$$r = \begin{pmatrix} a'' & & & \\ & c'' & b'' & \\ & b'' & c'' & \\ & & & a'' \end{pmatrix}$$

with inverse (as usual matrices):

$$\begin{pmatrix} a''' & & & \\ & c''' & b''' & \\ & b''' & c''' & \\ & & & a''' \end{pmatrix} = \begin{pmatrix} a'' & & & \\ & c'' & b'' & \\ & b'' & c'' & \\ & & & a'' \end{pmatrix}^{-1}.$$

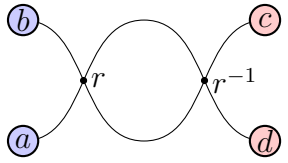
It may be computed that

$$a''' = \frac{1}{a''}, \qquad b''' = \frac{-b''}{(c'')^2 - (b'')^2}, \qquad c''' = \frac{c''}{(c'')^2 - (b'')^2}.$$

Then we compute that also

$$\frac{(a''')^2 + (b''')^2 - (c''')^2}{2a'''b'''} = \Delta.$$

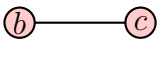
Now we may concatenate the matrices r and r^{-1}, and this is done by matrix multiplication. In other words, if we compute the partition function of the following system:

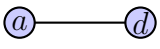


we get 1 if $a = d$ and $b = c$ but 0 otherwise. This is because summing over the middle column (four possibilities) really amounts to just multiplying matrices:

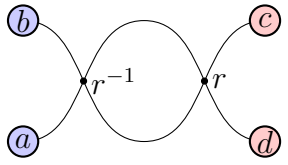
$$\begin{pmatrix} a'' & & & \\ & c'' & b'' & \\ & b'' & c'' & \\ & & & a'' \end{pmatrix} \begin{pmatrix} a''' & & & \\ & c''' & b''' & \\ & b''' & c''' & \\ & & & a''' \end{pmatrix}.$$

So this concatenation of r and r^{-1} is equivalent to:

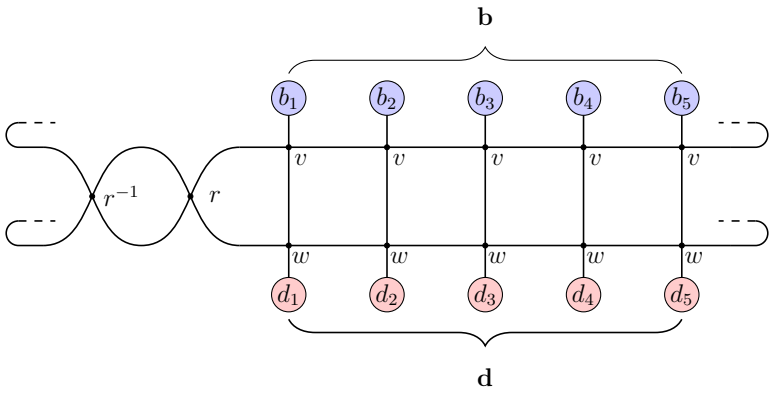




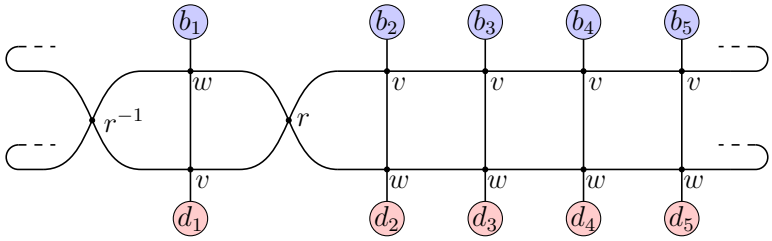
This is also equivalent to:



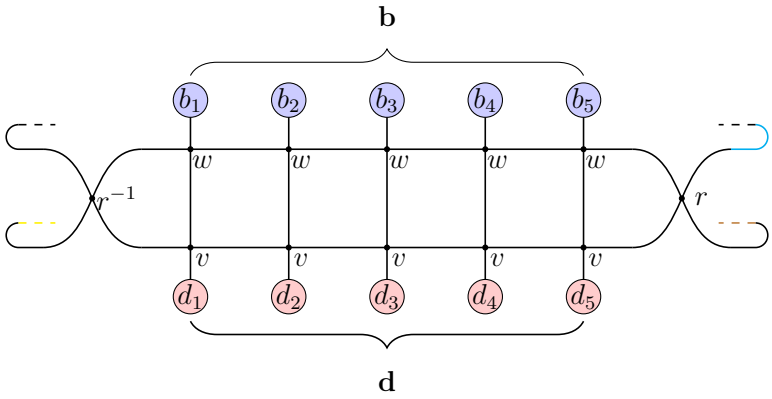
We may insert r and r^{-1} into our system representing $\langle \mathbf{d}|T_w T_v|\mathbf{b}\rangle$ to obtain:



Now we use the Yang–Baxter equation to see that this system is equivalent to:



We may repeat this process several times to obtain this system:



Due to the cylindric boundary conditions, the r and r^{-1} are again adjacent and may be discarded. But now the system represents $\langle \mathbf{d}|T_v T_w|\mathbf{b}\rangle$. We have proven that the two row transfer matrices commute. □

The argument in the last proof, where the R-vertex r moves past two rows of the grid to interchange two rows, is known as the *train argument*.

3. Vector Yang–Baxter equation

We will give another notion of the Yang–Baxter equation. Soon we will connect it with the familiar one that we used in Chapter 1 and earlier in Chapter 2.

Let U, V and W be vector spaces. Suppose that we are given three linear transformations:

$$r : U \otimes V \longrightarrow V \otimes U, \qquad s : U \otimes W \longrightarrow W \otimes U, \qquad t : V \otimes W \longrightarrow W \otimes V.$$

We will consider two homomorphisms $U\otimes V\otimes W \longrightarrow W\otimes V\otimes U$. The first is the composition

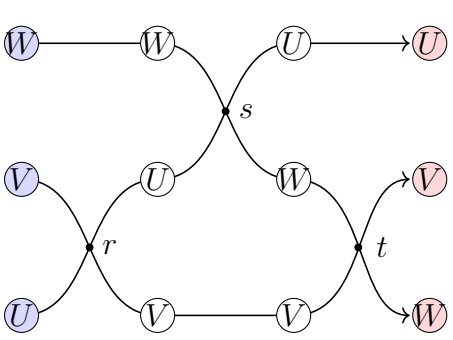
$$U\otimes V\otimes W \xrightarrow{R\otimes I_W} V\otimes U\otimes W \xrightarrow{I_V\otimes S} V\otimes W\otimes U \xrightarrow{T\otimes I_U} W\otimes V\otimes U\,.$$

We may more compactly denote $r_{12}=r\otimes I_W$, $s_{23}=I_V\otimes s$ and $t_{12}=t\otimes I_U$. In this notation, r_{ij} means r applied to the i and j components of a tensor. (The subscript notation is popular in the Hopf algebra and quantum group literature.) The other homomorphism is

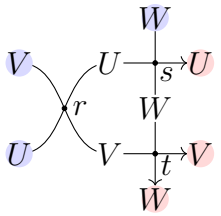
$$U\otimes V\otimes W \xrightarrow{I_U\otimes T} U\otimes W\otimes V \xrightarrow{S\otimes I_V} W\otimes U\otimes V \xrightarrow{R\otimes I_V} W\otimes V\otimes U\,.$$

We can diagram the homomorphisms graphically as follows.

$$U\otimes V\otimes W\xrightarrow{\rho_{12}}V\otimes U\otimes W\xrightarrow{s_{23}}V\otimes W\otimes U\xrightarrow{t_{12}}W\otimes V\otimes U$$

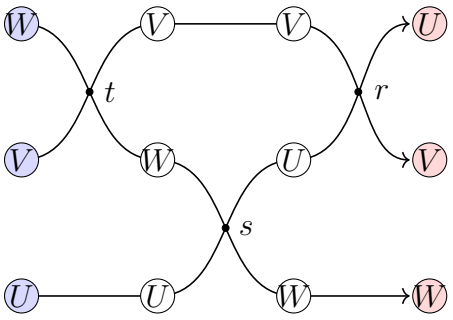


Alternative orientation:

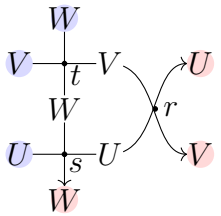


and

$$U\otimes V\otimes W\xrightarrow{t_{23}}U\otimes W\otimes V\xrightarrow{s_{12}}W\otimes U\otimes V\xrightarrow{r_{23}}W\otimes V\otimes U$$



Alternative orientation:



The equality of these two homomorphisms $U\otimes V\otimes W\to W\otimes V\otimes U$ is the *vector Yang–Baxter equation*.

Another useful notation for writing the Yang–Baxter equation involves the *Yang–Baxter commutator*

$$[\![r,s,t]\!] = (t\otimes I_V)(I_V\otimes s)(r\otimes I_V) - (I_V\otimes r)(s\otimes I_V)(I_V\otimes t)$$

in $\mathrm{End}(V\otimes V\otimes V)$. The vector Yang–Baxter equation in this notation is $[\![r,s,t]\!]=0$.

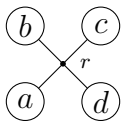
We will now explain how this vector Yang–Baxter equation is related to the Yang–Baxter equations we have previously described in terms of Boltzmann weights.

By a *frame* we mean a vector space with a distinguished basis. Thus if Σ_e is the spinset of the edge e, and if V_e is the spinspace, then (V_e,Σ_e) is a frame. Alternatively, if $a\in\Sigma_e$ we

will denote by v_a the corresponding element of V_a. If $V = V_e$ we may denote Σ_e as Σ_V to avoid needing to mention the edge e.

Example 2.3. In the six-vertex model $a \in \Sigma_e = \{\oplus, \ominus\}$. If U, V, W are six-vertex spinspaces, each has a basis u_+, u_- for U, v_+, v_- for V or w_+, w_- for W.

Let us start with a vertex r with chosen Boltzmann weights $\mathsf{a}_1(r), \mathsf{a}_2(r)$, etc. As in Section 6 of Chapter 1, we may encode these weights in a linear transformation $r : U \otimes V \longrightarrow V \otimes U$ by the following rule. If $a, b, c, d \in \Sigma$ then the Boltzmann weight of the state



is to be the coefficient of $v_d \otimes u_c$ in $r(u_a \otimes v_b)$. We will write this coefficient in Dirac notation as $\langle v_d \otimes u_c | r | u_a \otimes v_b \rangle$, or if we are thinking of it as a Boltzmann weight as

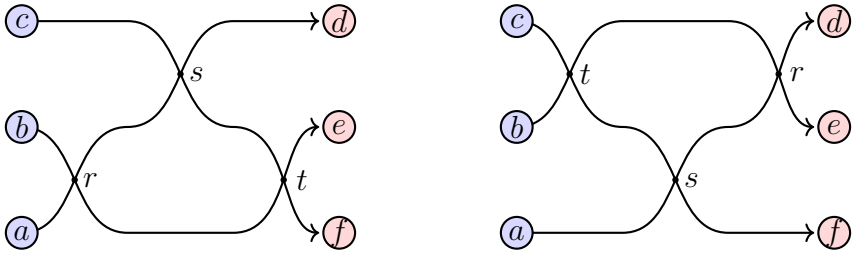
$$\beta_r \begin{pmatrix} b & c \\ a & d \end{pmatrix}.$$

So

$$r|u_a \otimes v_b\rangle := r(u_a \otimes v_b) = \sum_{c,d} \beta_r \begin{pmatrix} b & c \\ a & d \end{pmatrix} |v_d \otimes u_c\rangle.$$

Lemma 2.4. *The partition functions of the systems*

(2.7)



equal $\langle w_f \otimes v_e \otimes u_d | t_{12} s_{23} r_{12} | u_a \otimes v_b \otimes w_c \rangle$ *and* $\langle w_f \otimes v_e \otimes u_d | r_{13} s_{23} r_{12} | u_a \otimes v_b \otimes w_c \rangle$.

Proof. We've labeled the interior edges for reference in the following calculation.

$$t_{12} s_{23} r_{12} | u_a \otimes v_b \otimes w_c \rangle = \sum_{g,h} \beta_r \begin{pmatrix} b & g \\ a & h \end{pmatrix} t_{12} s_{23} | v_h \otimes u_g \otimes w_c \rangle$$

$$= \sum_{g,h} \sum_{i,d} \beta_r \begin{pmatrix} b & g \\ a & h \end{pmatrix} \beta_s \begin{pmatrix} c & d \\ g & i \end{pmatrix} t_{12} | v_h \otimes w_i \otimes u_d \rangle$$

$$= \sum_{g,h} \sum_{d,i} \sum_{e,f} \beta_r \begin{pmatrix} b & g \\ a & h \end{pmatrix} \beta_s \begin{pmatrix} c & d \\ g & i \end{pmatrix} \beta_t \begin{pmatrix} i & e \\ h & f \end{pmatrix} | w_f \otimes v_e \otimes u_d \rangle.$$

Therefore

$$\langle w_f \otimes v_e \otimes u_d | t_{12} s_{23} r_{12} | u_a \otimes v_b \otimes w_c \rangle = \sum_{g,h,i} \beta_r \left(\begin{smallmatrix} b & g \\ a & h \end{smallmatrix}\right) \beta_s \left(\begin{smallmatrix} c & d \\ g & i \end{smallmatrix}\right) \beta_t \left(\begin{smallmatrix} i & e \\ h & f \end{smallmatrix}\right).$$

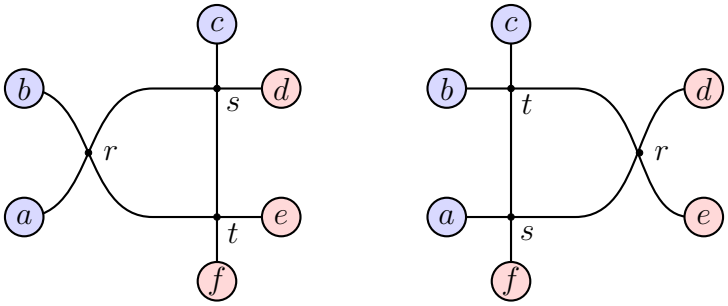
The right hand side is the partition function of the left-side of the Yang–Baxter equation system. As usual, the boundary spins a, b, c, d, e, f are fixed, and the spins of the interior edges g, h, i or j, k, l are summed over in the partition function. We leave the reader to check the other side. □

Therefore:

Theorem 2.5. *Let r, s, t be vertex types, and let U, V, W be as above, and define homomorphisms $r : U \otimes V \longrightarrow V \otimes U$ as above. A necessary condition that for all choices of boundary spins the partition functions of the systems (2.7) agree is that $[\![r, s, t]\!] = 0$ is satisfied.*

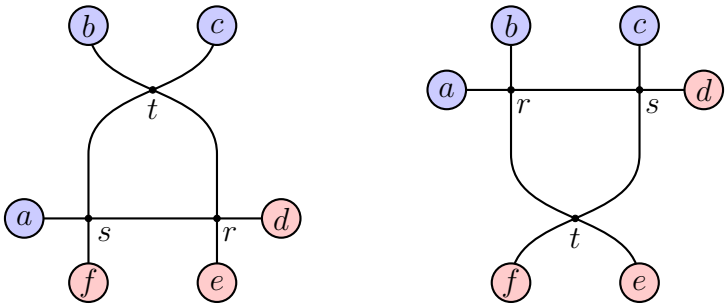
One may also reorient the edges and work instead with the systems:

(2.8)



or

(2.9)

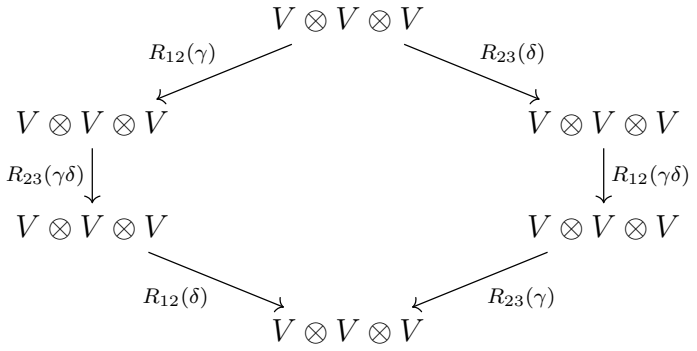


Lemma 2.6. *The two Yang-Baxter equations in (2.8) and (2.9) are equivalent, to each other and to the system in Theorem 2.5.*

Proof. Note that the left-hand system in (2.8) can be deformed into the left-hand system in (2.7) or to the right-hand system of (2.9). Similarly the right-hand system in (2.8) can be deformed into the right-hand system in (2.7) or the left-hand system of (2.9). Both systems are equivalent to $[\![r, s, t]\!] = 0$. □

4. Parametrized Yang–Baxter equations

Let Γ be a group, and let V be a vector space. Let $R\colon \Gamma \longrightarrow \mathrm{GL}(V \otimes V)$ be a map such that for every $\gamma, \delta \in \Gamma$, we have a vector Yang–Baxter equation:



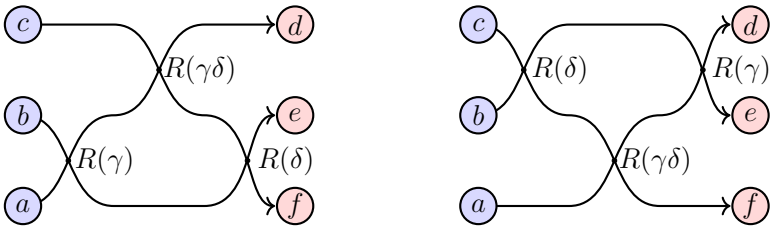
Then we say that we have a *parametrized Yang–Baxter equation* with parameter group Γ. In terms of the Yang–Baxter commutator

$$[\![R(\gamma), R(\gamma\delta), R(\delta)]\!] = 0.$$

We also require that $R(1_\gamma)$ is a scalar matrix, that is, a constant multiple of the identity matrix, and that furthermore for every $\gamma \in \Gamma$ that $R(\gamma)R(\gamma^{-1})$ is a scalar matrix.

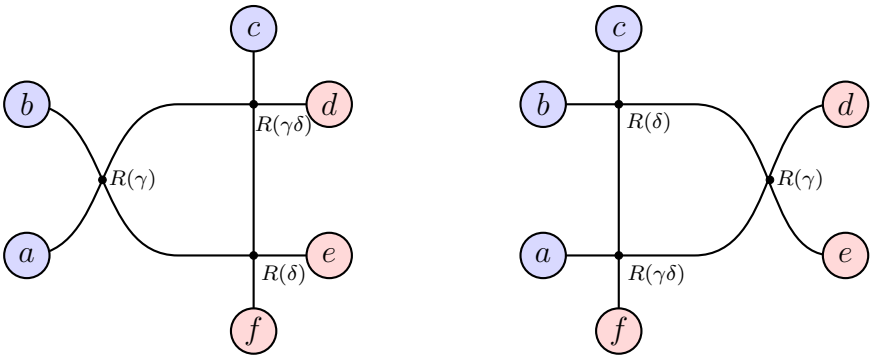
Usually the space V has a fixed basis Σ serving as the spinset of edges in the models where the Boltzmann weights come from the parametrized system. Then as in Section 3 we get vector Yang-Baxter equations in which $R(\gamma)$ encodes the Boltzmann weights for a vertex type. By Theorem 2.5, for all a, b, c, d, e, f the two following partition functions are equal:

(2.10)



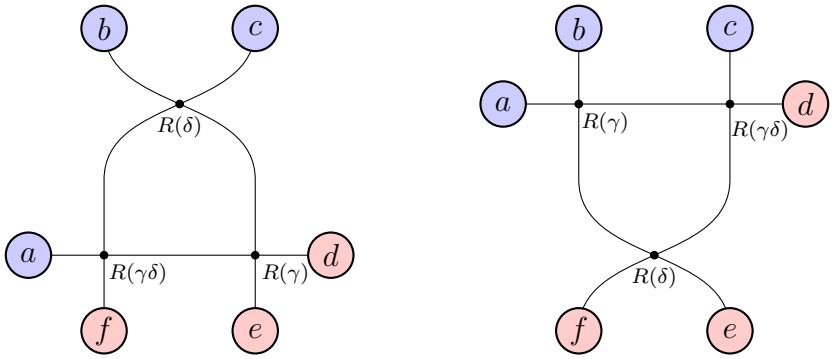
We could alternatively orient the edges as follows:

(2.11)



or equivalently (by Lemma 2.6):

(2.12)



In all cases, the procedure in Section 3 produces a vector Yang–Baxter equation, with V being the free vector space on the spinset Σ.

Proposition 2.7. *The three parametrized Yang-Baxter equations in (2.10), (2.11) and (2.9) are equivalent.*

Proof. This is a special case of Lemma 2.6. □

Example 2.8. We will show in the next Chapter that the field-free Yang–Baxter equation of Section 1.1 gives an example of a parametrized Yang–Baxter equation.

Example 2.9. Here is a parametrized Yang–Baxter equation in the six-vertex model with parameter group $\mathbb{C}^\times$. If $z \in \mathbb{C}^\times$, let $R(z)$ be the vertex with Boltzmann weights:

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
$1-q^2z$	$1-q^2z$	$q(1-z)$	$q(1-z)$	$1-q^2$	$z(1-q^2)$

Then (Exercise 2.3.)

$$[\![R(z), R(zw), R(w)]\!] = 0$$

Note that this parametrized Yang–Baxter equation is *almost* field-free since $\mathtt{a}_1 = \mathtt{a}_2$ and $\mathtt{b}_1 = \mathtt{b}_2$, although $\mathtt{c}_1$ and $\mathtt{c}_2$ differ. This example can be deduced from Example 2.8 by methods explained in Chapter 5.

Example 2.10. Here is another parametrized Yang–Baxter equation in the six-vertex model with parameter group $\mathbb{C}^\times$. If $z \in \mathbb{C}^\times$, let $R(z)$ be the vertex with Boltzmann weights

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
$z-q^2$	$1-q^2z$	$q(1-z)$	$q(1-z)$	$1-q^2$	$z(1-q^2)$

Then (Exercise 2.3.)

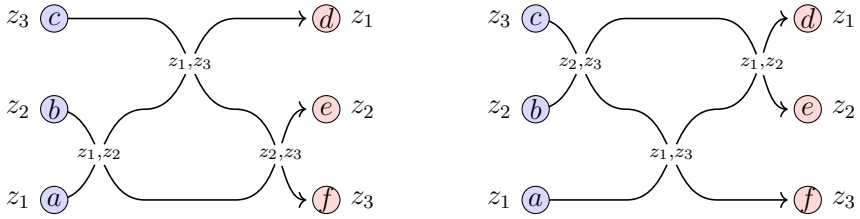
$$[\![R(z), R(zw), R(w)]\!] = 0.$$

This example is very similar to Example 2.9 since only the $\mathtt{a}_1$ weights differ. Yet the similarity is misleading, for unlike Example 2.9 this Yang–Baxter equation is *not* related to the field-free Yang–Baxter equation, or to Example 2.9 by any simple transformation. This example is *free-fermionic*, which means that the Boltzmann weights satisfy the identity $\mathtt{a}_1\mathtt{a}_2 + \mathtt{b}_1\mathtt{b}_2 - \mathtt{c}_1\mathtt{c}_2 = 0$.

4.1. Homogeneous parametrized Yang-Baxter Equations. As a variant of the parametrized Yang-Baxter equation, we will also consider systems in which the edge types are parametrized by a complex parameter z, and where edges with parameters z_1 and z_2 meet there is an R-matrix $R(z_1, z_2) \in \mathrm{End}(V \otimes V)$. We require that the R-matrix will be linear in both variables and given three parameters z_1, z_2 and z_3 there is a Yang-Baxter equation

$$[\![R(z_1, z_2), R(z_1, z_3), R(z_2, z_3)]\!] = 0. \tag{2.13}$$

This means the following two systems are equivalent:



Furthermore, we ask that $R(z, z)$ is a scalar linear transformation. This means that there exists a constant c such that $R(z, z) = czI_{V\otimes V}$. We also require that $R(z, w)R(w, z)$ is a scalar linear transformation. If these conditions are satisfied, we call this a *homogenous parametrized Yang-Baxter equation*.

Example 2.11. Take the Boltzmann weights from Example 2.9, and take $z = z_2/z_1$, $w = z_3/z_2$ so $zw = z_3/z_1$. We may adjust the R-matrix $R(z_1, z_2)$ by multiplying by z_1, which does not affect the validity of the Yang-Baxter equation, since then both sides are multiplied by the same value $z_1^2 z_2$. Thus we obtain (2.13) with the following R-matrix:

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
$z_1 - q^2 z_2$	$z_1 - q^2 z_2$	$q(z_1 - z_2)$	$q(z_1 - z_2)$	$z_1(1 - q^2)$	$z_2(1 - q^2)$

Example 2.12. Similarly Example 2.10 gives another homogeneous parametrized Yang-Baxter equation with the following weights:

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
$z_2 - q^2 z_1$	$z_1 - q^2 z_2$	$q(z_1 - z_2)$	$q(z_1 - z_2)$	$z_1(1 - q^2)$	$z_2(1 - q^2)$

5. Solvability for parametrized systems

In this section we define solvability and then specialize to the case where the Boltzmann weights come from a parametrized Yang–Baxter equation. In this special case, we will prove the equivalence of row and column solvability, and classify the solvable models with some very simple data.

We consider a grid with n rows and N columns. For every vertex, we assume that there is given a set of Boltzmann weights so that we have a lattice model.

Definition 2.13. The model is row-solvable if we can use the train argument to interchange rows. That is, let $R_1, \ldots, R_n$ be the rows. If $1 \leqslant a, b \leqslant n$ we may consider the two rowed grid consisting of R_a on top of R_b. Then there is assumed to be an R-matrix that we may attach and run the train argument, interchanging the rows.

Remark 2.14. Note that in this definition the rows R_a and R_b are not assumed to be adjacent. Since the train argument moves the rows around, this is needed if the train argument is to be used more than once.

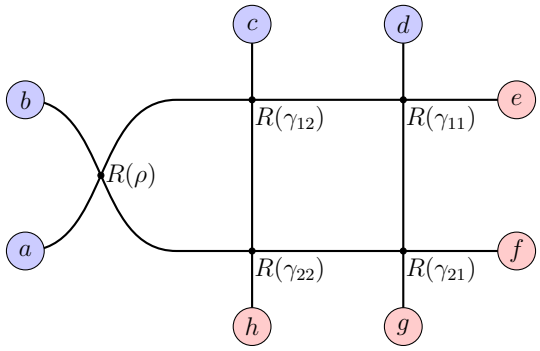
Column solvability is defined the same way with requiring R-matrices to interchange columns (not necessarily adjacent). If a system is both row and column solvable we will call it *solvable*.

We now specialize to the case where all the Boltzmann weights are assumed to come from a parametrized Yang–Baxter equation $R : G \longrightarrow \mathrm{End}(V \otimes V)$ for some group G. We will assume the map R is injective, so if $R(\gamma) = R(\delta)$ then $\gamma = \delta$.

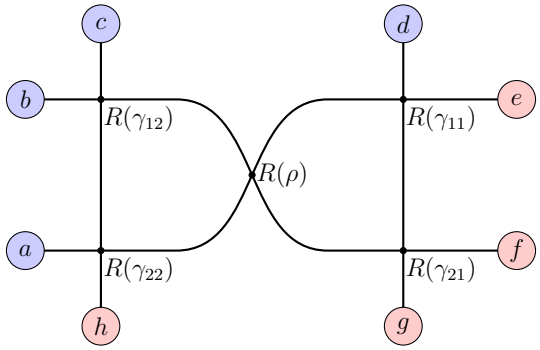
Let $\gamma_{ij} \in G$ be the element describing the Boltzmann weights in the i-th row and j-th column. As usual, we number the columns in decreasing order.

Theorem 2.15. *In this situation, row solvability is equivalent to column solvability.*

Proof. One may reduce to the case of a 2×2 grid, and we will explain that case. With only two rows, we attach the R-matrix, called $R(\rho)$.



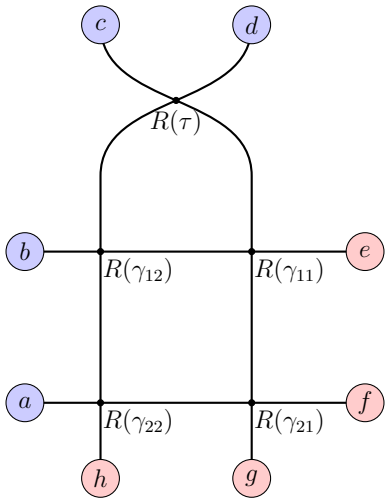
We remind the reader that we label columns in descending order, so γ_{11} is in the upper left. Remembering the parametrized Yang–Baxter equation, in G we must have $\gamma_{12} = \rho\gamma_{22}$, so $\rho = \gamma_{12}\gamma_{22}^{-1}$. Now using the Yang–Baxter equation, this equals the following partition function:



To use the Yang–Baxter equation again, we need $R(\rho)R(\gamma_{21}) = R(\gamma_{11})$, and so $\rho = \gamma_{11}\gamma_{21}^{-1}$. Combining this with our previous formula for ρ we obtain the condition for row-solvability which is that $\gamma_{12}\gamma_{22}^{-1} = \gamma_{11}\gamma_{21}^{-1}$. We prefer to take inverses here and write this condition

$$\gamma_{22}\gamma_{12}^{-1} = \gamma_{21}\gamma_{11}^{-1}. \tag{2.14}$$

Now let us similarly derive a condition for column solvability. We have to attach the vertical R-matrix $R(\tau)$ for some $\tau \in G$.



By Proposition 2.7 we may use (2.9) here and obtain $\gamma_{12} = \gamma_{11}\tau$, so $\tau = \gamma_{11}^{-1}\gamma_{12}$. But at the next stage of the train argument, we need $\tau = \gamma_{21}^{-1}\gamma_{22}$, and so we obtain

$$\gamma_{11}^{-1}\gamma_{12} = \gamma_{21}^{-1}\gamma_{22}. \tag{2.15}$$

This is the condition for column solvability. But (2.14) is equivalent to (2.15) and so row solvability is equivalent to column solvability. □

Since row solvability and column solvability are equivalent for the class of models we are now considering, we will simply call such models solvable.

Corollary 2.16. *A necessary and sufficient condition for solvability is that when* $1 \leqslant a, b \leqslant n$ *and* $1 \leqslant c, d \leqslant N$ *then*

$$\gamma_{bd}\gamma_{ad}^{-1} = \gamma_{bc}\gamma_{ac}^{-1}. \tag{2.16}$$

Proof. This generalizes (2.14). One may apply the argument to the two row grid consisting of the a and b rows for row solvability as in Definition 2.13. Column solvability is treated the same way. The fact that the row and column solvability criteria are equivalent is similar to the proof of (2.16). □

Now we can classify the solutions to 2.16 to construct solvable models with a parametrized Yang–Baxter equation. In fact, the following result is a complete classification. See Chapter 4 Section 6 for an illustration of this result.

Theorem 2.17. *Let* $\{\phi(a) \mid 1 \leqslant a \leqslant n\}$ *and* $\{\psi(b) \mid 1 \leqslant b \leqslant N\}$ *be arbitrary sequences in the parameter group* G*. Define*

$$\gamma_{ab} = \phi(a)\psi(b). \tag{2.17}$$

Then (2.16) is satisfied, so these Boltzmann weights define a solvable lattice model. Conversely, if (2.16) is satisfied, then functions ϕ *and* ψ *may be found such that the Boltzman weights are given by (2.17).*

Proof. Given ϕ and ψ with γ_{ab} as in (2.17) then

$$\gamma_{bd}\gamma_{ad}^{-1} = (\phi(b)\psi(d))(\phi(a)\psi(d))^{-1} = \phi(b)\phi(a)^{-1},$$

which is independent of d, hence equals $\gamma_{bc}\gamma_{ac}^{-1}$, proving (2.16).

To prove the converse, assume (2.16) and take $a = 1, c = 1$ to write

$$\gamma_{bd} = \gamma_{b1}\gamma_{11}^{-1}\gamma_{1d} = \phi(b)\psi(d)$$

where $\phi(b) = \gamma_{b1}$ and $\psi(d) = \gamma_{11}^{-1}\gamma_{1d}$. This proves (2.17). □

Exercises

The next two exercises are related to Theorem 2.1.

Exercise 2.1.
(i) In the proof of Theorem 2.1, we checked that the advertized r satisfies (2.4) but we did not check (2.5) and (2.6). Check these facts.
(ii) We also did not check that if $\Delta(r)$ is defined, then $\Delta(r) = \Delta(v) = \Delta(w)$. Prove this fact.
(iii) Assume that $\Delta(v) = \Delta(w)$ and that $c(v)$ and $c(w)$ are nonzero. Prove that r is unique up to constant multiple.

Exercise 2.2. Suppose that v and w are field-free vertices such that $c(v)$ and $c(w)$ are nonzero. If there exists a nonzero vertex r such that the Yang–Baxter equation (2.1) is satisfied, then $\Delta(v) = \Delta(w)$. **Hint**: use the three equations (2.4), (2.5) and (2.6).

Exercise 2.3.
(i) Prove the parametrized Yang–Baxter equation in Example 2.9.
(ii) Prove the parametrized Yang–Baxter equation in Example 2.10.

Exercise 2.4. Let $R(z)$ be the R-matrix in either Example 2.11 or Example 2.12. Show that $R(z)R(1/z)$ is the scalar $(q^2 - z)(q^2 - 1/z)$,

The goal of the next problem is to prove the relationship between two different forms of the Yang–Baxter equation.

Exercise 2.5. Let $r \in \mathrm{End}(V \otimes V)$. We have already introduced the notation $r_{ij} \in \mathrm{End}(V \otimes V \otimes V)$ where (i, j) is one of $(1, 2)$, $(2, 3)$ or $(1, 3)$. But to recapitulate, this means that r is applied to the i-th and j-th component of $V \otimes V \otimes V$, so $r_{12} = r \otimes I_V$ and $r_{23} = I_V \otimes r$; and if we expand r as a sum $\sum r_i' \otimes r_i''$ with r_i' and $r_i'' \in \mathrm{End}(V)$, then $r_{13} = \sum r_i' \otimes I_V \otimes r_i''$. The same notation applies also map (for example)

$$r \in \mathrm{Hom}(U \otimes V, V \otimes U) \qquad \text{to} \qquad r_{12} \in \mathrm{Hom}(U \otimes V \otimes W, V \otimes U \otimes W).$$

Let furthermore

$$s \in \mathrm{Hom}(U \otimes W, W \otimes U), \qquad t \in \mathrm{Hom}(V \otimes W, W \otimes V).$$

The first form of the Yang–Baxter equation is the identity $[\![r, s, t]\!] = 0$ in $\mathrm{Hom}(U \otimes V \otimes W, W \otimes V \otimes U)$. This can be written

$$t_{12}s_{23}r_{12} = r_{23}s_{12}t_{23}. \tag{2.18}$$

The other form of the Yang–Baxter equation is, for $R \in \mathrm{End}(U \otimes V)$, $S \in \mathrm{End}(U \otimes W)$, $T \in \mathrm{End}(V \otimes W)$ the identity

$$T_{12}S_{13}R_{12} = R_{12}S_{13}T_{12} \tag{2.19}$$

in $\mathrm{End}(U \otimes V \otimes W)$. Let $\tau : U \otimes V \to V \otimes U$ be the *flip map* $\tau(u \otimes v) = v \otimes u$. Now let $R \in \mathrm{End}(U \otimes V)$ and define $r \in \mathrm{Hom}(U \otimes V, V \otimes U)$ by $r = \tau \circ R$. Similarly, let $s = \tau \circ S$ and $t = \tau \circ T$. Then prove that the identity (2.18) is equivalent to (2.19).

CHAPTER 3

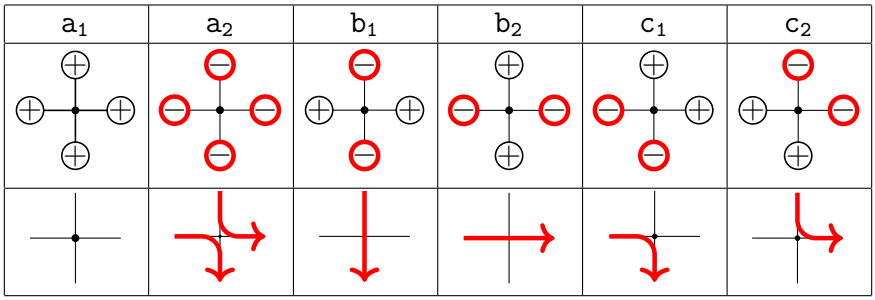
The Six-Vertex Model

We introduced the six-vertex model in Chapter 1, Section 5. In this chapter, we will study this model further. There are actually two representation theoretic aspects to our study. The first is a *global* one having to do with the theory in the large, such as information about the partition function, or about states of the system. In this chapter we show that the states of the six-vertex model can be parametrized by Gelfand–Tsetlin patterns, which have meaning in terms of the representation theory of $\mathrm{GL}(n, \mathbb{C})$. The *local* aspect consists of the properties of Boltzmann weights and the Yang–Baxter equation. We started our study of the Yang–Baxter equation for the six-vertex model in Chapter 2, and here we continue that study.

We remind the reader of the basic setup from Chapter 1, Sections 5 and 6. The spinset for every edge in the six-vertex model is the fixed set $\{\oplus, \ominus\}$. Let V be $\mathbb{C}^2$ with basis v_+, v_-, corresponding to $\oplus$ and $\ominus$. If v is a vertex with Boltzmann weights $\mathtt{a}_1(\mathtt{v})$, $\mathtt{a}_2(\mathtt{v})$, $\mathtt{b}_1(\mathtt{v})$, $\mathtt{b}_2(\mathtt{v})$, $\mathtt{c}_1(\mathtt{v})$, $\mathtt{c}_2(\mathtt{v})$, labeled as in Section 1.5, then as explained in Section 1.6, v may be associated with the linear transformation (1.1.2), which is an endomorphism of $V \otimes V$.

1. Paths

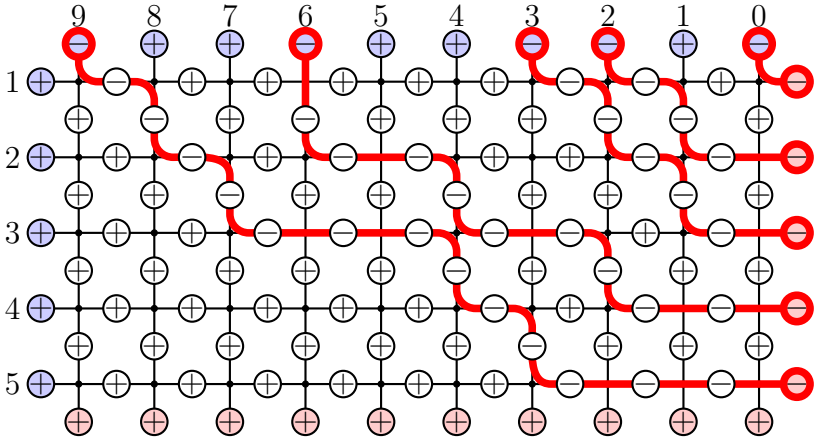
In many models we may visualize states in terms of paths (or lines) through the lattice. For the six-vertex model, we interpret a $\ominus$ (resp. $\oplus$) state as the presence (resp. absence) of a particle. We visualize the particles as moving from top to bottom, and from left to right.



We have drawn the particles in red, then visualized the paths they must take. In the case of $\mathtt{a}_2$ we have elected not to allow the paths to cross, though in other schemes they *might* cross.

The classification of the boundary edges is inputs or outputs (Remark 1.2 in Chapter 1) works well for the six-vertex model. If we call the top and left boundary edges inputs, and the right and bottom edges outputs, then every path connects an input to an output. Here is an example of a state of the six-vertex model with the paths drawn in red. In this example they start at the top boundary and finish on the right edge. Depending on the boundary

conditions some paths might start at the left boundary, or finish at the bottom.



In Chapter 2 we considered cylindric boundary conditions, wrapping the grid around into a cylinder. We might also consider *toroidal* boundary conditions, additionally wrapping the top to bottom so that there are no boundary edges.

Now, however, we want to do no wrapping, envisioning a rectangular grid with boundary edges on the left, right, top and bottom. Now let N be the number of columns and n be the number of rows of a rectangular grid. Let us assume that $N > n$.

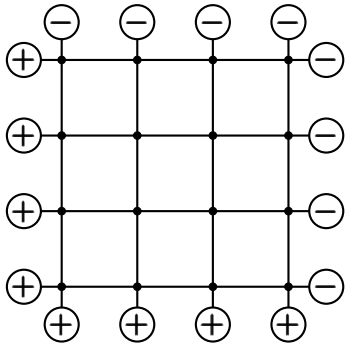
To specify the system, we must specify the boundary spins. These are constrained by the following fact.

Lemma 3.1. *The number of $\ominus$ on the top and left must equal the number of $\ominus$ on the right and bottom, or else the system has no admissible states.*

Proof. Every path must start at the top or left and finish on the right or bottom. This gives a bijection between the $\ominus$ spins on the top or left and those on the right or bottom. □

If the grid is square, we may put $\oplus$ on the left and bottom boundary edges, and $\ominus$ on the top and right boundary edges.

(3.1)

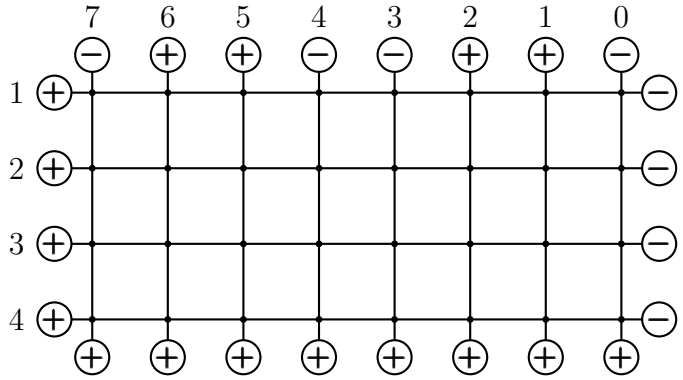


We refer to this specification as *domain-wall boundary conditions*. The assumption that the grid is square is needed here, since otherwise Lemma 3.1 shows that the system has no states.

In this text we will often have more columns than rows. Thus assuming $N > n$, we may (as with the domain-wall boundary conditions) put $\oplus$ on the left and bottom boundary, and $\ominus$ on the right boundary. But Lemma 3.1 shows that we must then have exactly n spins of $\ominus$ on the top, and the rest must be $\oplus$. We will call this arrangement *extended wall boundary*

conditions. It looks like this:

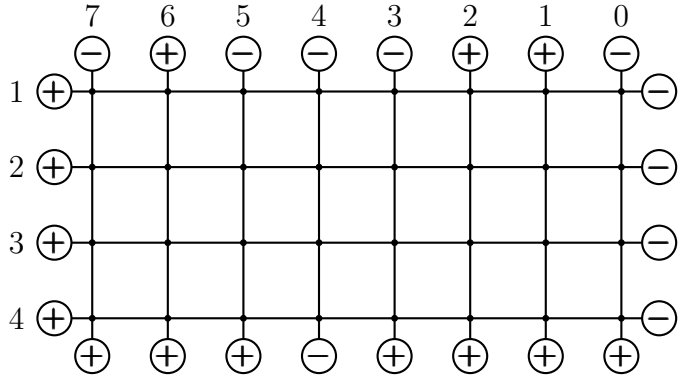
(3.2)



Observe that we have labeled the rows and columns. The columns are numbered starting at 0 in order from right to left.

We will also encounter *skew wall boundary conditions* which have $\oplus$ spins on the left boundary, $\ominus$ spins on the right boundary, but generalizing the extended wall conditions allow $\ominus$ spins on the bottom edge.

(3.3)



2. Gelfand–Tsetlin patterns and states

Let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ be a sequence of nonnegative integers. We say λ is a *partition* if it is weakly decreasing:

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0.$$

We identify two partitions if they differ only by some trailing zeros. Thus $(3, 2)$ and $(3, 2, 0, 0)$ are equal as partitions. The *length* of the partition is the number of nonzero parts. Thus a partition of length $\leqslant n$ may be identified with an element of $\Lambda := \mathbb{Z}^n$. We say λ is a *partition of* k, and write $\lambda \vdash k$ or $k = |\lambda|$ if $\sum \lambda_i = k$. The partition is *strict* if

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n \geqslant 0.$$

A strict partition is the same as a partition into unequal parts.

More generally, let $\lambda = (\lambda_1, \dots, \lambda_n) \in \Lambda$ be any sequence of integers such that $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. Then we call λ a *dominant weight* of length n. Thus a dominant weight λ is a partition of length $\leqslant n$ if and only if $\lambda_n \geqslant 0$ (so all entries are nonnegative).

The reason for this terminology comes from Lie theory. Let $G = \mathrm{GL}(n, \mathbb{C})$, and let T be the subgroup of diagonal matrices. The *weight lattice* of G is the group $X^*(T)$ of rational

characters of T. By definition, a rational character is an algebraic homomorphism from T into the multiplicative group. Such a character has the form

$$\mathbf{t} \mapsto \mathbf{t}^{\lambda} := t_1^{\lambda_1} \cdots t_n^{\lambda_n}, \qquad \mathbf{t} = \begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_n \end{pmatrix},$$

for some $\lambda \in \Lambda$. Thus $X^*(T)$ may be identified with Λ. The weight λ is dominant if and only if $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. We will refer to Λ as the $\mathrm{GL}(n)$ *weight lattice*.

Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\mu = (\mu_1, \ldots, \mu_{n-1})$ be partitions or dominant weights of lengths $\leqslant n$ and $n-1$, respectively. We say that λ and μ *interleave* if

$$\lambda_1 \geqslant \mu_1 \geqslant \lambda_2 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_{n-1} \geqslant \lambda_n. \tag{3.4}$$

We make the same definition if λ and μ are dominant weights, so that their entries are allowed to be negative.

Gelfand–Tsetlin patterns are triangular arrays of integers satisfying certain inequalities. Specifically, a Gelfand–Tsetlin pattern of size n is an array

$$A = \begin{Bmatrix} a_{1,1} & & a_{1,2} & & a_{1,3} & & \cdots & & a_{1,n} \\ & a_{2,1} & & a_{2,2} & & \cdots & & a_{2,n-1} & \\ & & \ddots & & & & \iddots & & \\ & & & & a_{n,1} & & & & \end{Bmatrix}$$

such that the rows are weakly decreasing (so they are dominant weights) that interleave. A Gelfand–Tsetlin pattern is *strict* if every row is a strict partition. If $\lambda = (\lambda_1, \cdots, \lambda_n)$, let $\mathsf{GTP}_n(\lambda)$ be the set of Gelfand-Tsetlin patterns of size n with top row λ, so $a_{1,i} = \lambda_i$.

For example, there are 8 Gelfand—Tsetlin patterns with top row $(2, 1, 0)$. These are:

$$\begin{Bmatrix} 2 & 1 & 0 \\ 1 & 0 & \\ 0 & & \end{Bmatrix}, \quad \begin{Bmatrix} 2 & 1 & 0 \\ 1 & 0 & \\ 1 & & \end{Bmatrix} \quad \begin{Bmatrix} 2 & 1 & 0 \\ 1 & 1 & \\ 1 & & \end{Bmatrix}, \quad \begin{Bmatrix} 2 & 1 & 0 \\ 2 & 0 & \\ 0 & & \end{Bmatrix}$$

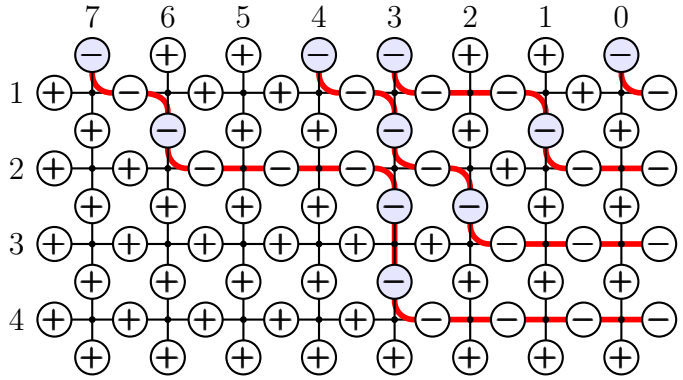
$$\begin{Bmatrix} 2 & 1 & 0 \\ 2 & 0 & \\ 1 & & \end{Bmatrix}, \quad \begin{Bmatrix} 2 & 1 & 0 \\ 2 & 0 & \\ 2 & & \end{Bmatrix} \quad \begin{Bmatrix} 2 & 1 & 0 \\ 2 & 1 & \\ 1 & & \end{Bmatrix}, \quad \begin{Bmatrix} 2 & 1 & 0 \\ 2 & 1 & \\ 2 & & \end{Bmatrix}.$$

These patterns are all strict, except the third one, which we have marked in red.

2.1. States and strict Gelfand–Tsetlin patterns. We will now show that states of a model with extended wall boundary conditions are in bijection with strict Gelfand–Tsetlin patterns of size n, where n is the number of rows of the model. The Gelfand–Tsetlin pattern of a state $\mathfrak{s}$ can be read off from the locations of the $\ominus$ spins on the vertical edges. Let us

illustrate this with the following example:

(3.5)



Note that we have colored every *vertical* edge that has a $\ominus$ spin. The column numbers of these vertical edges can be arranged in an array, thus:

$$\begin{Bmatrix} 7 & & 4 & & 3 & & 0 \\ & 6 & & 3 & & 1 & \\ & & 3 & & 2 & & \\ & & & 3 & & & \end{Bmatrix} \tag{3.6}$$

This is a Gelfand–Tsetlin pattern by the following result.

Proposition 3.2. *Let $\mathfrak{s}$ be a state of the six-vertex model with extended wall boundary conditions and n rows. Above the i-th row, there are exactly $n+1-i$ vertical edges with $\ominus$ spins. Putting the column numbers of these into the i-th row of an array produces a strict Gelfand–Tsetlin pattern* $\mathrm{GTP}(\mathfrak{s})$. *Conversely, given a strict Gelfand–Tsetlin pattern with n rows, provided $N-1$ is larger than every entry of the pattern, there is a unique state $\mathfrak{s}$ corresponding to the pattern in this way.*

Proof. First let us show that there are exactly $n+1-i$ vertical edges with $\ominus$ spins above the i-th row. If $i=1$, then by Lemma 3.1, the number of $\ominus$ spins on the top boundary is exactly n, as required. Now we argue by induction. Assuming there are $n+1-i$ edges above the i-th row with $\ominus$ spins, there are that many paths downward paths to the i-th row; one exits to the right, so $n-i$ paths must exit downward. This means that there are $n-i$ vertical edges carrying $\ominus$ spins above the $i+1$-st row.

It must be shown that the rows interleave. Let $a_{i,1},\dots,a_{i,n+1-i}$ be the columns of the $\ominus$ spins above the i-th row. We must show that

$$a_{i,j} \geqslant a_{i+1,j} \geqslant a_{i,j+1}.$$

The path through the $a_{i,j}$ edge (meaning the edge in column j above the i-th row) is also through the $a_{i+1,j}$ edge; since the paths move down and to the right, $a_{i,j} \geqslant a_{i+1,j}$. To prove that $a_{i+1,j} \geqslant a_{i,j+1}$, note that if this were not true, the path through the $a_{i,j}$ edge would move horizontally on the i-th row past the $a_{i,j+1}$ column. But then it would collide with the path coming downwards through the $a_{i,j+1}$ edge. So this cannot happen.

We have proved that the array $\mathrm{GTP}(\mathfrak{s})$ is a Gelfand–Tsetlin pattern. It is obvious that it is strict.

Conversely, let us suppose that we are given a Gelfand–Tsetlin pattern T of size n. We consider a grid with n rows and N columns, where $N-1$ is greater than the entries in T. Thus $N-1$ is the largest column number in the grid. We may construct a state by putting

$\ominus$ on the vertical edges. The spins on the horizontal edges are then determined by the condition that every vertex is adjacent to an even number of $\ominus$ edges. We leave the reader to convince themselves that this procedure always produces a legal state. $\square$

Definition 3.3. We have already described extended wall boundary conditions on an $N \times n$ grid, but we have not specified precisely the spins on the top boundary. Given a strict partition $\lambda = (\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 \leqslant N$, let us put $\ominus$ spins on the top boundary vertical edges in the λ_i columns for $1 \leqslant i \leqslant n$, and $\oplus$ spins in the remaining columns. As before we put $\oplus$ spins on the left and bottom boundary edges and $\ominus$ spins on the right boundary edges. We will call these boundary conditions the *extended wall boundary conditions of weight* λ.

Corollary 3.4. *Let* $\mathfrak{s}$ *be a state of the extended wall boundary conditions of weight* λ *on an* $n \times N$ *grid. Then* λ *is the top row of* $\mathrm{GTP}(\mathfrak{s})$.

Proof. This is clear, since λ describes the location of the $\ominus$ spins above the first row, but these are the top vertical boundary edges. $\square$

Example 3.5. The extended wall boundary conditions on a 4×8 grid of weight $\lambda = (7, 4, 3, 0)$ are illustrated in (3.2), with a typical state in (3.5). Then $\mathrm{GTP}(\mathfrak{s})$ is given by (3.6) and indeed, the top row of this Gelfand–Tsetlin pattern is λ.

2.2. Gelfand–Tsetlin patterns and tableaux. Gelfand–Tsetlin patterns of size n with top row λ are also in bijection with another important class of mathematical objects, *semistandard Young tableaux* (SSYT) in the alphabet $\{1, 2, \ldots, n\}$.

To define these, recall that the *Young diagram* $\mathrm{YD}(\lambda)$ of a partition λ is a collection of boxes with λ_1 in the first row, λ_2 in the second row, etc. A *semistandard Young tableau* T (SSYT) of shape λ in the alphabet $\{1, 2, \ldots, n\}$ is a filling of $\mathrm{YD}(\lambda)$ with integers $1, \ldots, n$ such that the rows are weakly increasing, and the columns are strictly increasing. The *weight* $\mathrm{wt}(T)$ is $(\mu_1, \ldots, \mu_n)$ where μ_i is the number of i's in T. We denote the set of all semistandard Young tableaux of shape λ in the alphabet $\{1, 2, \ldots, n\}$ by $\mathsf{SSYT}_n(\lambda)$.

Example 3.6. Let $\lambda = (5, 2, 2)$ and $n = 5$. Then

$$T = \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 2 & 2 & 5 \\ \hline 2 & 2 \\ \cline{1-2} 3 & 5 \\ \cline{1-2} \end{array}$$

is a SSYT of shape λ. Its weight is $(2, 4, 1, 0, 2)$.

If $\mathrm{YD}(\mu) \subseteq \mathrm{YD}(\lambda)$, we call λ/μ a *skew partition*. A skew partition λ/μ is a *horizontal strip* if each column of $\mathrm{YD}(\lambda/\mu) := \mathrm{YD}(\lambda) \setminus \mathrm{YD}(\mu)$ contains at most one cell. The condition that λ/μ is a horizontal strip is equivalent to saying that λ and μ interleave.

We may now explain the bijection between the sets $\mathsf{SSYT}_n(\lambda)$ and $\mathsf{GTP}_n(\lambda)$. We may view a SSYT T as a sequence of partitions

$$\varnothing = \lambda^{(0)} \subseteq \lambda^{(1)} \subseteq \cdots \subseteq \lambda^{(n)} = \lambda$$

such that $\lambda^{(i)}/\lambda^{(i-1)}$ is a horizontal strip for $1 \leqslant i \leqslant n$. Namely, $\lambda^{(i)}/\lambda^{(i-1)}$ contains the letters i in T. Generalizing the notion of a semistandard Young tableau, we may also consider a filling of a *skew diagram* YD from the alphabet $\{1, 2, \ldots, n\}$ with the rows weakly

increasing and the columns strictly increasing. We will denote the set of such skew diagrams $\mathsf{SSYT}_n(\lambda/\mu)$. If $\lambda = (5,3,1)$ and $\mu = (2)$, here is a skew tableau of shape λ/μ in $\mathsf{SSYT}_3(\lambda/\mu)$:

$$\begin{array}{|c|c|c|c|c|}\hline \multicolumn{2}{c}{} & 2 & 2 & 3 \\ \hline 1 & 1 & 3 \\ \cline{1-3} 2 \\ \cline{1-1} \end{array}.$$

Example 3.7. The tableau T of Example 3.6 is associated to the sequence of partitions

$$\varnothing \subseteq (2) \subseteq (4,2) \subseteq (4,2,1) \subseteq (4,2,1) \subseteq (5,2,2).$$

Now let A be a Gelfand–Tsetlin pattern of size n. Call the rows of A from top to bottom $\lambda^{(n)}, \lambda^{(n-1)}, \ldots, \lambda^{(1)}$. Since $\lambda^{(i)}$ and $\lambda^{(i-1)}$ interleave, $\lambda^{(i)}/\lambda^{(i-1)}$ forms a horizontal strip. Hence

$$\varnothing = \lambda^{(0)} \subseteq \lambda^{(1)} \subseteq \cdots \subseteq \lambda^{(n)} = \lambda$$

defines a SSYT.

Example 3.8. For T in Example 3.6, the Gelfand–Tsetlin pattern is

$$\left\{\begin{array}{ccccccccc} 5 & & 2 & & 2 & & 0 & & 0 \\ & 4 & & 2 & & 1 & & 0 & \\ & & 4 & & 2 & & 1 & & \\ & & & 4 & & 2 & & & \\ & & & & 2 & & & & \end{array}\right\}.$$

The shape of T is the partition $(5,2,2)$, which we have to pad with zeros since the size $n = 5$ of the Gelfand–Tsetlin pattern is to be the size of the alphabet of T.

3. Parametrized field-free Yang–Baxter equation

There are two main kinds of parametrized Yang–Baxter equations in the six-vertex model. As we will explain in later chapters, these correspond to two different quantum groups, $U_q(\widehat{\mathfrak{gl}}_2)$ and the supersymmetric $U_q(\widehat{\mathfrak{gl}}(1|1))$. They are:

- The field-free six-vertex model, and variants. These are related to $U_q(\widehat{\mathfrak{gl}}_2)$ and will be treated in this section.
- The free-fermionic six-vertex model, and variants. These are related to $U_q(\widehat{\mathfrak{gl}}(1|1))$ and will be treated in the next section.

Let $\Delta \in \mathbb{C}$ be fixed. Let q be found such that $\frac{1}{2}(q + q^{-1}) = \Delta$. We will use the notation $R(a,b,c)$ for the vertex with Boltzmann weights a, b, c, as before. Let G_Δ be the set of (a,b,c) with $a, b \neq 0$ such that

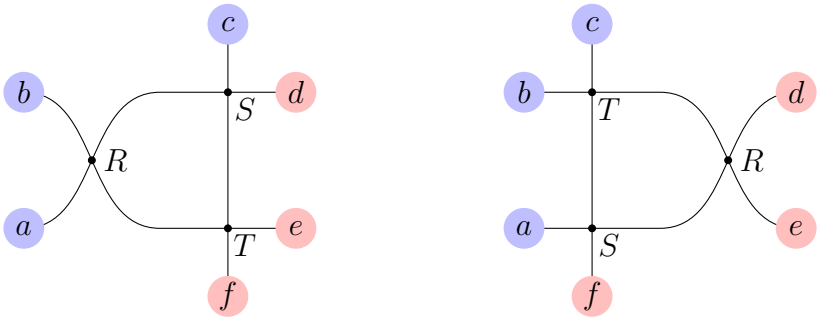
$$\frac{a^2 + b^2 - c^2}{2ab} = \Delta,$$

together with two additional elements $(\pm\Delta, 0, \Delta)$. Eventually we will give G_Δ the structure of a group.

In Chapter 1 we showed that if (a_1, b_1, c_1) and (a_2, b_2, c_2) are in G_Δ, then there exists a third $(a_0, b_0, c_0) \in G_\Delta$ such that if (in the notation of Chapter 1) $R = v(a_0, b_0, c_0)$, $S =$

$v(a_1, b_1, c_1)$ and $T = v(a_2, b_2, c_2)$, then we have a Yang–Baxter equation:

(3.7)



We note that the Yang–Baxter equation is homogeneous in the sense that if any one of (a_i, b_i, c_i) is multiplied by a nonzero constant then the validity of the equation is unchanged. So while R is usually determined by S and T, it is only determined up to a constant multiple.

Now we want to start with R and T and compute S. This will give us our first example of a parametrized Yang–Baxter equation. We begin by noting that G_Δ can be parametrized as follows.

Lemma 3.9. *Let $x \in \mathbb{C}^\times$ and let*

$$(a, b, c) = \left(\tfrac{1}{2}(xq - (xq)^{-1}), \tfrac{1}{2}(x - x^{-1}), \tfrac{1}{2}(q - q^{-1})\right). \tag{3.8}$$

Then $(a, b, c) \in G_\Delta$.

Proof. This is a straightforward calculation. □

Theorem 3.10. *The mapping*

$$R_\Delta \colon \mathbb{C}^\times \longrightarrow \{\textit{field-free Boltzmann weights } (a, b, c)\}$$

is a parametrized Yang–Baxter equation with parameter group $\mathbb{C}^\times$. Here the Boltzmann weights (a, b, c) of $R_\Delta(x)$ are given by (3.8).

Proof. Let R, S and T be field-free vertices with Boltzmann weights $R_\Delta(x)$, $R_\Delta(xy)$ and $R_\Delta(y)$, respectively. The Boltzmann weights are

$$\begin{aligned}
\beta_\Delta(R) &= \left(\frac{1}{2}((xq) - (xq)^{-1}), \frac{1}{2}(x - x^{-1}), \frac{1}{2}(q - q^{-1})\right) := (a, b, c),\\
\beta_\Delta(S) &= \left(\frac{1}{2}(xyq - (xyq)^{-1}), \frac{1}{2}(xy - (xy)^{-1}), \frac{1}{2}(q - q^{-1})\right) := (a', b', c'),\\
\beta_\Delta(T) &= \left(\frac{1}{2}(yq - (yq)^{-1}), \frac{1}{2}(y - y^{-1}), \frac{1}{2}(q - q^{-1})\right) := (a'', b'', c'').
\end{aligned}$$

Checking the parametrized Yang–Baxter equation is now a matter of computation. There are 12 cases of boundary Boltzmann weights that give nontrivial identities, but actually these are redundant and there are only 3 distinct identities. They are:

$$\begin{aligned}
cc'b'' + ba'c'' - ab'c'' &= 0,\\
ac'a'' - bc'b'' - ca'c'' &= 0,\\
cb'a'' - ca'b'' - bc'c'' &= 0.
\end{aligned}$$

These are easily checked. □

Remark 3.11. There are three special cases. If $\Delta = 0$, then we are in the free-fermionic case. The parametrized Yang–Baxter equation in Theorem 3.10 can be embedded in a much larger one with parameter group $\mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(2,\mathbb{C})$, so in this case Theorem 3.10 is true but it is not the whole story.

Remark 3.12. On the other hand, if $\Delta = \pm 1$ then $q = \Delta = \pm 1$ is the unique solution to $\Delta = \frac{1}{2}(q + q^{-1})$. We see from (3.8) that $c = 0$ and $a = \pm b$, so these are very degenerate systems. The values $\Delta = \pm 1$ are phase transition points. See Baxter Baxter 1982, Chapter 8.

Remark 3.13. Another interesting case is $q = e^{2\pi i/6}$. Then we can take $x = q = -(xq)^{-1}$, and all three Boltzmann weights a, b, c are equal. This fact was exploited by Kuperberg 1996 in proving the Alternating Sign Matrix Conjecture.

4. The free-fermionic six-vertex model

Another case where there is solvability is the *free-fermionic case*. Here the relevant Yang–Baxter equation was found (partly) by Korepin around 1981. See Korepin, Bogoliubov, and Izergin 1993a, page 126 with references to earlier literature. Later Brubaker, Bump and Friedberg Brubaker, Bump, and Friedberg 2011a rediscovered this in a slightly more general form and gave applications. They were not aware of Korepin's work until it was called to their attention (by Reshetikhin), but by that time Brubaker, Bump, and Friedberg 2011a was already in print, so Korepin was unfortunately not acknowledged in Brubaker, Bump, and Friedberg 2011a.

We call the six-vertex model vertex v *free-fermionic* if

$$\mathsf{a}_1(v)\mathsf{a}_2(v) + \mathsf{b}_1(v)\mathsf{b}_2(v) = \mathsf{c}_1(v)\mathsf{c}_2(v).$$

We are dropping the field-free condition.

It turns out that *all* free-fermionic weights fit into a parametrized Yang–Baxter equation with parameter group $\Gamma = \mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(1,\mathbb{C})$. This parametrized Yang–Baxter equation was discovered by Korepin (see Korepin, Bogoliubov, and Izergin 1993b page 126, and rediscovered by Brubaker, Bump and Friedberg Brubaker, Bump, and Friedberg 2011a). Let

$$\rho\colon\ \mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(1,\mathbb{C}) \longrightarrow \{\text{free-fermionic vertices}\}$$

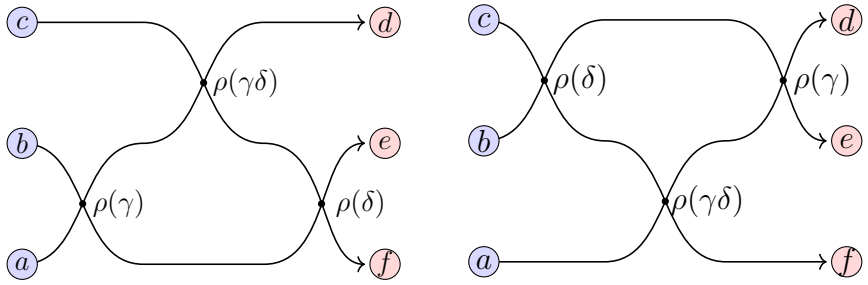
be the map that sends the element

$$\gamma = \left(\begin{pmatrix} a_1 & b_2 \\ -b_1 & a_2 \end{pmatrix}, c_1 \right)$$

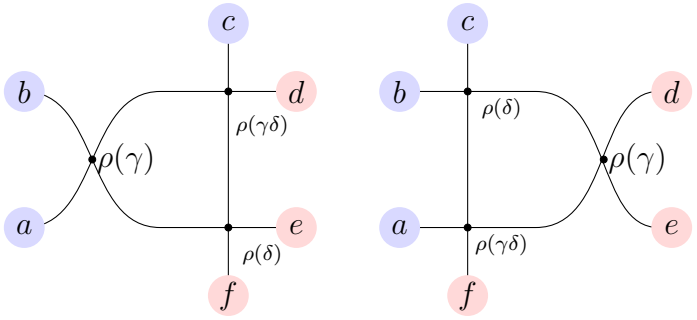
to the vertex with Boltzmann weights $a_1, a_2, b_1, b_2, c_1, c_2$, where $c_2 = (a_1a_2 + b_1b_2)/c_1$.

Theorem 3.14. *The map ρ is a parametrized Yang–Baxter equation with parameter group* $\mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(1,\mathbb{C})$.

The parametrized Yang–Baxter equation can be either of the two forms. We may ask that for all a, b, c, d, e, f the two following partition functions are equal:



Alternatively:



Proof. Let $r = \rho(\gamma)$, $t = \rho(\delta)$ and $s = \rho(\gamma\delta)$, where the product $\gamma\delta$ is just matrix multiplication. Thus

$$\gamma = \left(\begin{pmatrix} a_1(r) & b_2(r) \\ -b_1(r) & a_2(r) \end{pmatrix}, c_1(r) \right), \qquad \delta = \left(\begin{pmatrix} a_1(t) & b_2(t) \\ -b_1(t) & a_2(t) \end{pmatrix}, c_1(t) \right). \tag{3.9}$$

Multiplying γ and δ and remembering that $s = \rho(\gamma\delta)$, the s Boltzmann weights are:

$$c_1(s) = c_1(r)c_1(t), \qquad c_2(s) = c_2(r)c_2(t). \tag{3.10}$$

$$a_1(s) = a_1(r)a_1(t) - b_2(r)b_1(t), \qquad a_2(s) = -b_1(r)b_2(t) + a_2(r)a_2(t), \tag{3.11}$$

$$b_1(s) = b_1(r)a_1(t) + a_2(r)b_1(t), \qquad b_2(s) = a_1(r)b_2(t) + b_2(r)a_2(t). \tag{3.12}$$

Taking $(a, b, c, d, e, f) = (\oplus, \ominus, \oplus, \oplus, \ominus, \oplus)$ gives one equation which we may write

$$b_1(r)a_1(s)b_2(t) + c_2(r)c_1(s)c_2(t) = c_1(r)c_2(s)c_1(t) + b_1(r)a_1(s)b_2(t),$$

which is addressed by assuming (3.10). This also addresses the equation from $(a, b, c, d, e, f) = (\ominus, \oplus, \ominus, \ominus, \oplus, \ominus)$. Substituting the values of $c_i(S)$ from (3.10) throughout, there remain 12 equations, but each is divisible by one or more of $c_1(r)$, $c_2(r)$, $c_1(t)$ or $c_2(t)$, and dividing by these, each equation is repeated and there are only six equations to be satisfied. For example, taking $(a, b, c, d, e, f) = (\oplus, \oplus, \ominus, \ominus, \oplus, \oplus)$ and using the values (3.10) gives an equation that is divisible by both $c_2(r)$ and $c_2(t)$, and dividing by these gives $a_1(r)a_1(t) = b_2(r)b_1(t) + a_1(s)$, which is addressed by (3.11). After imposing (3.11) there are only four nonredundant equations, These simplify a little by using the free-fermionic condition in the form

$$c_2(r) = (a_1(r)a_2(r) + b_1(r)b_1(r))/c_1(r),$$
$$c_2(t) = (a_1(t)a_2(t) + b_1(t)b_1(t))/c_1(t),$$

and then substituting (3.12), all equations are satisfied. It may be checked that the vertex s is also free-fermionic. □

5. A general six-vertex Yang–Baxter equation

This section can be skipped on first reading. But this section provides a satisfactory answer to when, given six-vertex matrices r and t, there exists a solution s to the Yang–Baxter equation $[\![r, s, t]\!] = 0$. Ultimately (Bump and Naprienko 2025) this leads to a *groupoid* parametrized Yang–Baxter equation that accounts for almost all solutions (r, s, t). It is natural to ask whether this is a special case of a more general phenomenon.

The results in this section are based on Bump and Naprienko 2025; Naprienko 2022, and also Brubaker, Bump, and Friedberg 2011a. They are closely related to the results in the last two sections. The main result of Bump and Naprienko 2025 gives a parametrized Yang–Baxter equation in which the parameter object is not a group, but a *groupoid*. We will not explain that result here, however, but we will do some of the groundwork.

Let S be the space of six-vertex vertices r with $c_1(r)$ and $c_2(r)$ both nonzero. Furthermore, let $S^\bullet$ be the subpace where $a_1(r)$, $a_2(r)$, $b_1(r)$, $b_2(r)$ and $c_1(r)$, $c_2(r)$ are all nonzero.

Roughly, we would like to classify all solutions to $[\![r, s, t]\!] = 0$ with $r, s, t \in S$. This, however is a more delicate problem than we will consider, so in this section we will just consider the case where $r, t \in S^\bullet$ but s is in S. Let us define

$$N(r) = a_1(r)a_2(r) + b_1(r)b_2(r) - c_1(r)c_2(r) \tag{3.13}$$

and, generalizing Baxter's Δ from (2.3) in the field free case:

$$\Delta_1(r) = \frac{N(r)}{a_1(r)b_1(r)}, \qquad \Delta_2(r) = \frac{N(r)}{a_2(r)b_2(r)}. \tag{3.14}$$

We are omitting the 2 in the denominator of (2.3). Perhaps we should include it for consistency but it would play no role. Note that N is defined on all of S, but Δ_1 and Δ_2 are only defined on $S^\bullet$. We also define, for $r \in S^\bullet$ another element $r^* \in S$ by

$$a_1(r^*) = \frac{c_1(r)c_2(r) - b_1(r)b_2(r)}{a_1(r)}, \qquad a_2(r^*) = \frac{c_1(r)c_2(r) - b_1(r)b_2(r)}{a_2(r)}, \tag{3.15}$$

$$b_1(r^*) = -b_1(r), \quad b_2(r^*) = -b_2(r), \quad c_1(r^*) = c_2(r), \quad c_2(r^*) = c_1(r).$$

The map $r \to r^*$ has order 2 in a weak sense. Indeed, there is no guarantee that $a_1(r^*)$ or $a_2(r^*)$ is nonzero, but if it is, then both are nonzero and $r^* \in S^\bullet$. In this case $(r^*)^* = r$.

Lemma 3.15. *Suppose $r \in S^\bullet$. Then*

$$N(r^*) = \frac{b_1(r)b_2(r) - c_1(r)c_2(r)}{a_1(r)a_2(r)} N(r), \tag{3.16}$$

and if furthermore $r^* \in S^\bullet$, then

$$\Delta_1(r^*) = \frac{N(r)}{a_2(r)b_1(r)}, \qquad \Delta_2(r^*) = \frac{N(r)}{a_1(r)b_2(r)}. \tag{3.17}$$

Proof. We leave the verification to the reader. □

Remark 3.16. With $r \in S^\bullet$, the vertex r^* might not be in $S^\bullet$, and $a_1(r^*)$ and $a_2(r^*)$ could be zero. Hence $\Delta_1(r^*)$ and $\Delta_2(r^*)$ cannot be defined by (3.14). Indeed, if $a_1(r^*)$ or $a_2(r^*)$ vanishes, then by (3.15) we have $b_1(r)b_2(r) - c_1(r)c_2(r) = 0$, and by (3.16) we therefore have $N(r^*) = 0$. So both the numerator and the denominator in (3.14) for $\Delta_i(r^*)$ vanish. However the numerator and denominator of (3.17) are still nonzero. Therefore $\Delta_i(r^*)$ can still be defined by (3.17), and so $\Delta_i(r^*)$ still make sense if only $r \in S^\bullet$.

It will be useful to package Δ_1 and Δ_2 as a single function, so we define $\Delta : S^\bullet \to \mathbb{C}^2$ by $\Delta(r) = (\Delta_1(r), \Delta_2(r))$.

Theorem 3.17. *Let r and t be in $S^\bullet$. Then a necessary and sufficient condition for there to exist $s \in S$ such that $[\![r, s, t]\!] = 0$ is that $\Delta(r) = \Delta(t^*)$, where $\Delta(t^*)$ is defined even if $t^* \notin S^\bullet$ by Remark 3.16. If this is true, then s is determined up to a constant multiple, and can be chosen so that*

$$c_1(s) = c_1(r)c_1(t), \qquad c_2(s) = c_1(r)c_2(t). \tag{3.18}$$

If $s \in S^\bullet$ then $\Delta(s) = \Delta(t)$ and $\Delta(s^*) = \Delta(r^*)$.

Proof. Let us assume that s exists with $[\![r, s, t]\!] = 0$. We make use of the Yang–Baxter equation in the form (2.7). Different choices $a, b, c, d, e, f \in \{\oplus, \ominus\}$ of give 14 different equations.

Taking $b = e = \ominus$ and $a = c = d = f = \oplus$ gives the equation $c_2(r)c_1(s)c_2(t) = c_1(r)c_2(s)c_1(t)$. The term $b_1(r)a_1(s)b_2(t)$ also appears, but it is on both sides of the equation and can be cancelled. The same identity also comes from the choice $b = e = \oplus$ and $a = c = d = f = \ominus$. Since by assumption the $c_i(r)$, $c_i(t)$ and $c_i(s)$ are all to be nonzero, we can multiply s by a constant to put it in the normalization (3.18).

Thus we may substitute the values in (3.18) of $c_1(s)$ and $c_2(s)$. Then the number of equations is reduced to 12, but actually there are only 6 for the following reason. We find that each of the 12 equations is divisible by one of $c_1(r), c_2(r), c_2(t)$ or $c_2(t)$, and dividing out by these, the equations are now duplicated. To give one example, taking

$$(a, b, c, d, e, f) = (\oplus, \oplus, \ominus, \oplus, \ominus, \oplus)$$

gives the equation $-c_1(r)c_2(s)b_1(t) - b_1(r)a_1(s)c_2(t) + a_1(r)b_1(s)c_2(t) = 0$ but substituting (3.18) the equation is divisible by $-c_2(t)$ and we obtain $c_1(r)c_2(r)b_1(t) + b_1(r)a_1(s) - a_1(r)b_1(s) = 0$. But we can obtain the same identity by taking $(\oplus, \ominus, \oplus, \oplus, \oplus, \ominus)$ and dividing by $c_1(t)$. The six equations are:

$$\begin{aligned}
a_1(r)a_1(t) - b_2(r)b_1(t) &= a_1(s),\\
a_2(r)a_2(t) - b_1(r)b_2(t) &= a_2(s),\\
c_1(r)c_2(r)b_1(t) + b_1(r)a_1(s) - a_1(r)b_1(s) &= 0,\\
c_1(r)c_2(r)b_2(t) + b_2(r)a_2(s) - a_2(r)b_2(s) &= 0,\\
b_2(r)c_1(t)c_2(t) - b_2(s)a_1(t) + a_1(s)b_2(t) &= 0,\\
b_1(r)c_1(t)c_2(t) - b_1(s)a_2(t) + a_2(s)b_1(t) &= 0.
\end{aligned}$$

We may substitute the values for $a_1(s)$ and $a_2(s)$ from the first two equations, and then there are only four equations. Only one term in each depends on s and these may be rearranged as follows:

$$b_1(s) = \frac{a_1(r)b_1(r)a_1(t) - b_1(r)b_2(r)b_1(t) + c_1(r)c_2(r)b_1(t)}{a_1(r)},$$

$$b_1(s) = \frac{a_1(r)a_1(t)b_2(t) - b_2(r)b_1(t)b_2(t) + b_2(r)c_1(t)c_2(t)}{a_1(t)},$$
$$b_2(s) = \frac{a_2(r)b_2(r)a_2(t) - b_1(r)b_2(r)b_2(t) + c_1(r)c_2(r)b_2(t)}{a_2(r)},$$
$$b_2(s) = \frac{a_2(r)a_2(t)b_1(t) - b_1(r)b_1(t)b_2(t) + b_1(r)c_1(t)c_2(t)}{a_2(t)}.$$

So there is a solution if and only if the two expressions for $b_1(s)$ are equal, and the two expressions for $b_2(s)$ are equal. Equating the two expressions for $b_1(s)$ gives and rearranging gives $N(t)a_1(r)b_1(r) = N(r)a_2(t)b_1(t)$ which is equivalent to $\Delta_1(r) = \Delta_1(t^*)$. Similarly the equality of the two expressions for $b_2(s)$ is equivalent to $\Delta_2(r) = \Delta_2(t^*)$. This proves that the existence of a solution s to $[\![r, s, t]\!] = 0$ is equivalent to $\Delta(r) = \Delta(t^*)$.

We leave it to the reader to show that if a solution s exists then $\Delta(s) = \Delta(t)$ and $\Delta(s^*) = \Delta(t^*)$. □

Remark 3.18. Generalizing the notion of a group, a *groupoid* is a set with a partially defined composition law and an "inverse map." Bump and Naprienko 2025 prove that there exists a groupoid $\mathfrak{G}$ with composition law $\star$ and a map $\pi : \mathfrak{G} \longrightarrow S$ such that if $g, h \in \mathfrak{G}$ and $g \star h$ is defined, then $[\![\pi(g), \pi(gh), \pi(h)]\!] = 0$, and in some sense accounts for essentially all solutions to $[\![r, s, t]\!] = 0$ in six-vertex matrices r, s, t. This groupoid contains the group parametrized Yang–Baxter equations in Theorems 3.10 and 3.14 as subgroups. It is natural to ask whether this is a special case of a more general phenomenon.

Exercises

Exercise 3.1. We saw in Section 2 that states of the six-vertex model with extended wall boundary conditions are parametrized by Gelfand–Tsetlin patterns; there we saw that extended wall boundary conditions for an $n \times N$ grid require specification of the boundary spins on the top row of vertical boundary edges, and these may be encoded in a partition λ. For skew wall boundary conditions, we allow $\ominus$ spins on the bottom row of boundary edges, so we have must have another partition μ to describe the bottom boundary conditions.

(i) Show (using paths) that in order for the system to have states, $\mathrm{YD}(\mu) \subseteq \mathrm{YD}(\lambda)$, and $|\lambda| - |\mu| = n$, so λ/μ is a skew partition of size n.

(ii) Assuming that the system has states as in (i), give a bijection between states of the skew wall system and $\mathsf{SSYT}_{N-1}(\lambda/\mu)$.

Exercise 3.2 (Naprienko)**.** This exercise generalizes Examples 2.9 and 2.10 of Chapter 2. Let q_1, q_2 and β be fixed nonzero constants.

(i) Given $z_1, z_2, w \in \mathbb{C}^\times$, define a six-vertex matrix $R = R^{\mathrm{cf}}_{q_1,q_2,\beta}(z_1, z_2, w)$ by:

$\mathtt{a}_1(R)$	$\mathtt{a}_2(R)$	$\mathtt{b}_1(R)$
$q_1z_1 - q_2z_2$	$q_1z_1 - q_2z_2$	$q_1(z_1 - z_2)\beta$
$\mathtt{b}_2(R)$	$\mathtt{c}_1(R)$	$\mathtt{c}_2(R)$
$q_1(z_1 - z_2)\beta^{-1}$	$z_1(q_1 - q_2)w$	$z_2(q_1 - q_2)w^{-1}$

Let $V = \mathbb{C}^2$ as usual in the six-vertex model. Prove that $R^{\mathrm{cf}}_{q_1,q_2,\beta} : (\mathbb{C}^\times)^3 \longrightarrow \mathrm{GL}(V \otimes V)$ is a parametrized Yang–Baxter equation. This means

$$[\![R^{\mathrm{cf}}_{q_1,q_2,\beta}(z_1, z_2, w), R^{\mathrm{cf}}_{q_1,q_2,\beta}(z_1z_1', z_2z_2', ww'), R^{\mathrm{cf}}_{q_1,q_2,\beta}(z_1', z_2', w')]\!] = 0.$$

(ii) Given $z_1, z_2, w \in \mathbb{C}^\times$, define a six-vertex matrix $R = R^{\mathrm{ff}}_{q_1,q_2,\beta}(z_1, z_2, w)$ by:

$\mathtt{a}_1(R)$	$\mathtt{a}_2(R)$	$\mathtt{b}_1(R)$
$q_1 z_1 - q_2 z_2$	$q_1 z_2 - q_2 z_1$	$q_1(z_1 - z_2)\beta$
$\mathtt{b}_2(R)$	$\mathtt{c}_1(R)$	$\mathtt{c}_2(R)$
$q_1(z_1 - z_2)\beta^{-1}$	$z_1(q_1 - q_2)w$	$z_2(q_1 - q_2)w^{-1}$

Prove that

$$[\![R^{\mathrm{ff}}_{q_1,q_2,\beta}(z_1, z_2, w), R^{\mathrm{ff}}_{q_1,q_2,\beta}(z_1 z_1', z_2 z_2', w w'), R^{\mathrm{ff}}_{q_1,q_2,\beta}(z_1', z_2', w')]\!] = 0.$$

(iii) Compute $\Delta_1(R)$ and $\Delta_2(R)$ if $R = R^{\mathrm{cf}}_{q_1,q_2,\beta}$ and if $R = R^{\mathrm{ff}}_{q_1,q_2,\beta}$.

Exercise 3.3. The proof of Theorem 3.17 is incomplete. Finish it by showing that $\Delta(s) = \Delta(r)$ and $\Delta(s^*) = \Delta(r^*)$.

Exercise 3.4. In Theorem 3.17, there is no guarantee that $s \in S^\bullet$, because $\mathtt{a}_1(s)$, $\mathtt{a}_2(s)$, $\mathtt{b}_1(s)$ and $\mathtt{b}_2(s)$ could vanish. But assume that $N(r)$ and $N(t)$ are nonzero. Then prove that $\mathtt{a}_1(s) = 0$ if and only if $\mathtt{a}_2(s) = 0$, and that $\mathtt{b}_1(s) = 0$ if and only if $\mathtt{b}_2(s) = 0$.

Exercise 3.5. In the notation of Theorem 3.17, let $s, t \in S^\bullet$. Prove that a necessary and sufficient condition for there to be $r \in S$ such that $[\![r, s, t]\!] = 0$ is that $\Delta(s) = \Delta(t)$.

Exercise 3.6. If $r \in S^\bullet$ define $\Delta_0(r) = \Delta_1(r)\Delta_2(r)$. If $r, s, t \in S^\bullet$ satisfy $[\![r, s, t]\!] = 0$, prove that $\Delta_0(r) = \Delta_0(s) = \Delta_0(t)$.

Exercise 3.7. Suppose that $r, s, t \in S^\bullet$, and also assume that $r^*, s^*, t^* \in S^\bullet$. If $[\![r, s, t]\!] = 0$, prove that $[\![r^*, t, s]\!] = 0$ and $[\![s, r, t^*]\!] = 0$. Deduce that $[\![t, r^*, s^*]\!] = [\![s^*, t^*, r]\!] = [\![t^*, s^*, r^*]\!] = 0$.

Hint: Regard r as an endomorphism of $V \otimes V$ with matrix (1.1.2). From the fact that r, r^* are both in $S^\bullet$, deduce that r is invertible, and that r^* is a constant multiple of r^{-1}. Then prove that $[\![r^{-1}, t, s]\!] = 0$.

CHAPTER 4

Tokuyama Models

1. Schur polynomials

Schur polynomials are symmetric polynomials which are very important in representation theory and combinatorics. Some useful references are Bump 2013; Bump and Schilling 2017; Macdonald 1995; Stanley 1999. They have direct generalizations that are discussed in Macdonald 1992 and later publications by numerous authors; see Naprienko 2024 for more references to that literature. The free-fermionic six vertex model is a useful framework for Schur polynomials and their generalizations. See **ABPWDomino**; Brubaker, Bump, and Friedberg 2011a; Hamel and King 2007; Naprienko 2024 for treatments from this point of view. This Chapter is based in part on Brubaker, Bump, and Friedberg 2011a.

Let $\lambda = (\lambda_1, \dots, \lambda_r)$ be a partition of length $r \leqslant n$. If $r < n$ we pad λ with 0's so that $\lambda = (\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ has exactly n parts. (This is customary in dealing with partitions.) We will give two definitions of the Schur polynomial s_λ. It will not be obvious that the two definitions are equivalent. We will use a lattice model to prove this.

Let $\mathbf{z} = (z_1, \dots, z_n)$ be indeterminates.

1.1. First definition: determinants.

Define

$$s_\lambda(z_1, \dots, z_n) = \frac{\det(z_j^{\lambda_i+n-i})}{\det(z_j^{n-i})}. \tag{4.1}$$

For example, if $n = 3$,

$$s_\lambda(z_1, z_2, z_3) = \frac{\begin{vmatrix} z_1^{\lambda_1+2} & z_2^{\lambda_1+2} & z_3^{\lambda_1+2} \\ z_1^{\lambda_2+1} & z_2^{\lambda_2+1} & z_3^{\lambda_2+1} \\ z_1^{\lambda_3} & z_2^{\lambda_3} & z_2^{\lambda_3} \end{vmatrix}}{\begin{vmatrix} z_1^2 & z_2^2 & z_3^2 \\ z_1 & z_2 & z_3 \\ 1 & 1 & 1 \end{vmatrix}}.$$

This definition first appeared in Cauchy Cauchy 1815, who defined Schur functions prior to Schur. The denominator is the Vandermonde determinant:

$$\det(z_j^{n-i}) = \prod_{i<j}(z_i - z_j).$$

It will be useful to introduce the vector $\rho = (n-1, n-2, \dots, 0)$ so that the exponents are $(\lambda+\rho)_i = \lambda_i + \rho_i$ and write the numerator as $\det(z_j^{(\lambda+\rho)_i})$.

Lemma 4.1. *The function s_λ is a symmetric polynomial. It is homogeneous of degree $|\lambda| = \sum \lambda_i$.*

Proof. The polynomial ring $\mathbb{C}[z_1, \ldots, z_n]$ is a unique factorization domain. Let us note that the numerator is divisible by every factor $z_i - z_j$ with $i < j$ of the Vandermonde denominator. Indeed, the numerator vanishes when $z_i = z_j$ since two columns of the determinant $\det(z_j^{\lambda_i+n-i})$ are then equal. Thus the numerator is divisible by each factor and therefore by their product since they are coprime. Therefore s_λ is a polynomial. It is symmetric since interchanging z_i and z_j multiplies the numerator and the denominator by -1. The homogeneity is also clear since the numerator and denominator are both homogeneous polynomials. □

The partition λ may be thought of as a dominant weight for the Lie group $\mathrm{GL}(n, \mathbb{C})$. The definition (4.1) is essentially the Weyl character formula, which we now recall.

Let G be an arbitrary complex reductive Lie group, which we will soon specialize to $\mathrm{GL}(n, \mathbb{C})$. Let T be a maximal torus. The group $\Lambda = X^*(T)$ of rational characters is the *weight lattice*. Then Λ contains a root system Φ; let $\{\alpha_1, \ldots, \alpha_r\}$ be the simple positive roots and s_i the corresponding simple reflection in the Weyl group W. By a *Weyl vector* $\rho \in \Lambda$ we mean a $\rho \in \Lambda$ such that $s_i(\rho) = \rho - \alpha_i$ for every positive root.

If $\lambda \in \Lambda$ and $\mathbf{z} \in T$, we will denote by $\mathbf{z}^\lambda$ the application of λ to $\mathbf{z}$. If λ is dominant, it is the highest weight of an irreducible representation, whose character we will denote by χ_λ. The Weyl character formula asserts that

$$\chi_\lambda(\mathbf{z}) = \frac{\sum_{w \in W} (-1)^{\ell(w)} \mathbf{z}^{w(\lambda+\rho)}}{\sum_{w \in W} (-1)^{\ell(w)} \mathbf{z}^{w(\rho)}}. \tag{4.2}$$

Here ℓ is the length function on the Weyl group W.

Now suppose that $G = \mathrm{GL}(n, \mathbb{C})$ and T is the diagonal torus, which we identify with $(\mathbb{C}^\times)^n$ in the obvious way. Then Λ may be identified with $\mathbb{Z}^n$ so that $\mathbf{z}^\lambda = \prod_i z_i^{\lambda_i}$. The Weyl group W is the symmetric group S_n. For the Weyl vector we may take $\rho = (n-1, n-2, \ldots, 0)$. A partition λ is a dominant weight. The numerator and denominator in (4.2) can be compared with the numerator and denominator determinants in (4.1), and we see that $\chi_\lambda(\mathbf{z})$ is the Schur polynomial $s_\lambda(\mathbf{z})$.

1.2. Second definition: tableaux. Recall from Chapter 3 that $\mathsf{SSYT}_n(\lambda)$ is the set of semistandard Young Tableaux (SSYT) of shape λ with entries in the alphabet $\{1, 2, \ldots, n\}$.

The second definition of the Schur function is due to D.E. Littlewood (1938), given by the formula:

$$s_\lambda(z_1, \ldots, z_n) = \sum_{T \in \mathsf{SSYT}_n(\lambda)} \mathbf{z}^{\mathrm{wt}(T)}, \tag{4.3}$$

where the weight $\mathrm{wt}(T)$ of the tableau T is defined in Section 2 of Chapter 2.

It is not obvious from the second definition that the Schur polynomial is symmetric, though that property *does* follow immediately from the first definition. On the other hand, it is obvious from this second definition that the coefficients in the Schur polynomial are nonnegative, a fact is not immediately apparent from the first definition. We will use a lattice model to show that (4.3) is symmetric and equivalent to the first definition.

2. Tokuyama models

There is a formula due to Tokuyama Tokuyama 1988 for the Schur function, or more precisely for

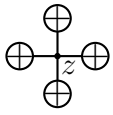
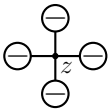
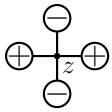
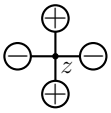
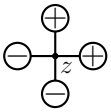
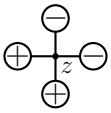
$$\left\{\prod_{i<j}(z_i - qz_j)\right\} s_\lambda(z_1, \dots, z_n) \tag{4.4}$$

as a sum over strict Gelfand–Tsetlin patterns. If $q = 1$, the product is the Vandermonde determinant in the denominator of the first definition, and Tokuyama's formula reduces to the first definition of the Schur polynomial. On the other hand, if $q = 0$, Tokuyama's formula reduces to the combinatorial definition. The special case $q = -1$ has an interpretation in terms of Hall–Littlewood polynomials Tokuyama 1988, Section 3.3.

Tokuyama's original formula can be reformulated as expressing (4.4) as the partition function of a solvable lattice model. This was done by Hamel and King Hamel and King 2007. However they did not use the Yang–Baxter equation. The Yang–Baxter equation was then applied to this problem by Brubaker, Bump and Friedberg Brubaker, Bump, and Friedberg 2011a.

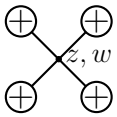
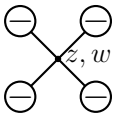
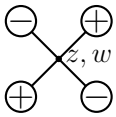
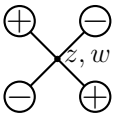
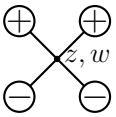
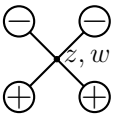
We take the following weights, labeled by a complex number z:

(4.5)

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
					
1	z	$-q$	z	$z(1-q)$	1

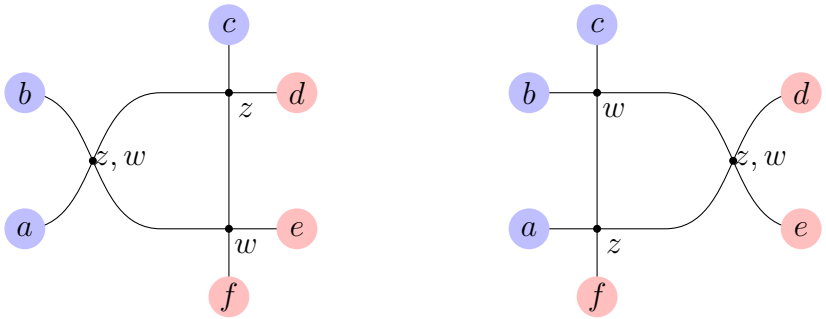
We also take the following R-matrix, labeled by two complex numbers z, w:

(4.6)

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
					
$w - qz$	$z - qw$	$q(z-w)$	$z - w$	$(1-q)z$	$(1-q)w$

Note that this is similar to the R-matrix in Example 2.12 of Chapter 2, though the T-matrix (4.5) is different. The R-matrix can be changed into that of Example 2.12 by the methods of Chapter 5, replacing q by q^2.

Theorem 4.2. *The Yang–Baxter equation is satisfied in that the following two systems are equivalent for all choices of* $a, b, c, d, e, f \in \{\oplus, \ominus\}$*:*



Proof. This RTT equation is a special case of the free-fermionic Yang–Baxter equation in Theorem 3.14 in Chapter 3 with parameter group $\mathrm{GL}(2,\mathbb{C}) \times \mathbb{C}^\times$. To check this, we take r to be the vertex labeled z, w with the Boltzmann weights divided by $(1-q)w$, which does not affect the validity of the Yang–Baxter equation. We take t to be the matrix labeled w. Then γ and δ as in equation (3.9) of Chapter 3 are the following elements of $\mathrm{GL}(2,\mathbb{C}) \times \mathbb{C}^\times$:

$$\left(\begin{pmatrix} \dfrac{w-qz}{(1-q)w} & \dfrac{z-w}{(1-q)w} \\ -\dfrac{q(z-w)}{(1-q)w} & \dfrac{z-qw}{(1-q)w} \end{pmatrix}, \frac{z}{w} \right), \qquad \left(\begin{pmatrix} 1 & w \\ q & w \end{pmatrix}, w(1-q) \right).$$

Multiplying these gives

$$\left(\begin{pmatrix} 1 & z \\ q & z \end{pmatrix}, z(1-q) \right),$$

and the statement follows. □

Now let us explain the models we want to use, called "Gamma Ice" in Brubaker, Bump, and Friedberg 2011a. We will use the extended wall boundary conditions with n rows, and $N > \lambda_1$ columns. We will use the weights (4.5) with the parameter $z = z_i$ in the i-th row. Thus we have defined a system that we will denote $\mathfrak{S}_\lambda(\mathbf{z}; q)$. Let $Z_\lambda(z_1, \dots, z_n; q) = Z_\lambda(\mathbf{z}; q)$ be the corresponding partition function.

Theorem 4.3. *The partition function*

$$Z_\lambda(\mathbf{z}; q) = \prod_{i<j} (z_i - q z_j) S_\lambda(\mathbf{z})$$

where $S_\lambda(\mathbf{z}) = S_\lambda(z_1, \dots, z_n)$ is a symmetric polynomial that is independent of q.

We will give part of the proof in the next section using the train argument and the Yang–Baxter equation. We will then show that it implies the equivalence of the two definitions of the Schur function.

Remark 4.4. The symmetric polynomial S_λ will turn out to be the Schur polynomial s_λ. Eventually we will prove, by specializing the parameter q to 0 and 1 and carefully analyzing the partition function that S_λ satisfies *both* definitions of the Schur polynomial. To avoid confusion, we will not use the notation s_λ until these facts are proved in Theorems 4.12 and 4.18.

3. Proof of Theorem 4.3

Proposition 4.5. *The quotient*

$$S_\lambda(\mathbf{z}; q) = \frac{Z_\lambda(\mathbf{z}; q)}{\prod_{i<j} (z_i - q z_j)} \tag{4.7}$$

is symmetric, that is, invariant under permutations of the z_i.

Proof. We multiply (4.7) by:

$$\prod_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} (z_i - q z_j) \,.$$

This is a symmetric polynomial of degree $n(n-1)$. It consists of the $\frac{1}{2}n(n-1)$ factors in the denominator of (4.7) and $\frac{1}{2}n(n-1)$ other factors, hence it is enough to show that

$$Z_\lambda(\mathbf{z};q)\prod_{i<j}(z_j - qz_i)$$

is symmetric. Let $1 \leqslant k < n$ and let s_k be the "simple reflection" in the symmetric group which interchanges k and $k+1$. These generate the symmetric group, so it is sufficient to show that the last expression is invariant under s_k.

We can pull one factor out and write this as

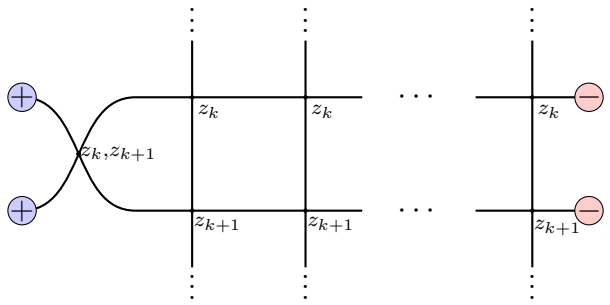
$$Z_\lambda(\mathbf{z};q)(z_{k+1} - qz_k)\left[\prod_{\substack{i<j\\(i,j)\neq(k,k+1)}}(z_j - qz_i).\right]$$

The permutation s_k just permutes the $\frac{1}{2}n(n-1)-1$ factors in brackets. So we may drop these to see that it is sufficient to show that

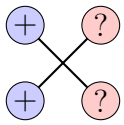
$$Z_\lambda(\mathbf{z};q)(z_{k+1} - qz_k) = Z_\lambda(s_k\mathbf{z};q)(z_k - qz_{k+1}). \tag{4.8}$$

To see this, let us consider the following system. We attach the R-matrix with coordinates z_k, z_{k+1} to the left at the $k, k+1$ rows:

(4.9)



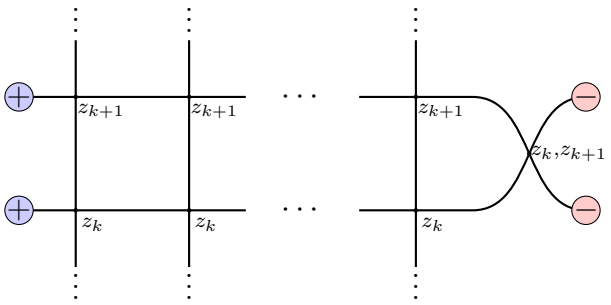
We note that from the Boltzmann weights if the "input" spins are $\oplus, \oplus$ there is only one possibility for the output spins, which must also be $\oplus, \oplus$:



spins marked ? can only $= \oplus$.

The Boltzmann weight of the R-matrix is $z_{k+1} - qz_k$, and so the partition function of the system (4.9) is the left-hand side of (4.8). Using the train argument, this equals the partition

function of



and by the same reasoning, this equals the right-hand side of (4.8). This proves (4.8) and the symmetry of $S_\lambda(\mathbf{z}; q)$ is established. □

Proposition 4.6. *$S_\lambda(\mathbf{z}; q)$ is a polynomial in $z_1, \ldots, z_n$ and q.*

Proof. It is clear that $Z_\lambda(\mathbf{z}; q)$ is a polynomial, since every Boltzmann weight is a polynomial. Rewrite (4.7) as

$$S_\lambda(\mathbf{z}; q) = \frac{\prod_{i>j}(z_i - qz_j) Z_\lambda(\mathbf{z}; q)}{\prod_{i \neq j}(z_i - qz_j)}. \tag{4.10}$$

Both the numerator and the denominator on the right-hand side here are symmetric. In the polynomial ring $\mathbb{C}[z_1, \ldots, z_n, q]$, which is a unique factorization domain, the denominator is a product of coprime polynomials. Hence it is sufficient to show that the numerator is divisible by each. If $i > j$ then it is obvious that the numerator in (4.10) is divisible by $z_i - qz_j$ since it is included as a factor in the product defining the numerator. Because it is symmetric, it is divisible by all factors $z_i - qz_j$ because the symmetric group permutes these transitively. Thus the quotient $S_\lambda(\mathbf{z}; q)$ is a polynomial. □

Lemma 4.7. *Let $\mathfrak{s}$ be a state of the model. The total number of vertices with spin configurations of types $\mathtt{a}_2, \mathtt{b}_1$ and $\mathtt{c}_1$ in the state is $\frac{1}{2}n(n-1)$.*

Proof. A configuration is of type $\mathtt{a}_2, \mathtt{b}_1$ or $\mathtt{c}_1$ if and only if it has a $\ominus$ in the vertical edge below the vertex. We recall the Gelfand–Tsetlin pattern associated to the state in Lemma 1.3 of Chapter 3. There is a $\ominus$ spin on the vertical edge below the vertex in row i and column j if and only if j is one of the entries in the $(i+1)$-th row of the Gelfand–Tsetlin pattern. There are thus $n-1$ configurations of type $\mathtt{a}_2, \mathtt{b}_1$ or $\mathtt{c}_1$ in the first row, $n-2$ in the second row, and so forth, and $\frac{1}{2}n(n-1)$ altogether. □

Proposition 4.8. *$S_\lambda(\mathbf{z}; q)$ is independent of q.*

Proof. The numerator and denominator in (4.7) are both polynomials in $z_1, \ldots, z_n, q$ and the denominator has degree $\frac{1}{2}n(n-1)$ in q. We claim that the numerator has the same degree in q. Reviewing the Boltzmann weights, only configurations of types $\mathtt{b}_1$ and $\mathtt{c}_1$ can contribute a power of q. The number of such patterns is at most $\frac{1}{2}n(n-1)$ by Lemma 4.7.

Since the degree in q of the numerator of (4.7) is at most $\frac{1}{2}n(n-1)$, the degree of the denominator is exactly $\frac{1}{2}n(n-1)$, and the quotient is known to be a polynomial, it has degree 0 in q. Hence it is independent of q. □

Since $S_\lambda(\mathbf{z}; q)$ is independent of q, we may suppress q from the notation and write $S_\lambda(\mathbf{z}; q) = S_\lambda(\mathbf{z})$. We have proved that it is a symmetric polynomial. Later we will show that if $q = 0$, this agrees with the combinatorial definition of $s_\lambda(\mathbf{z})$, and if $q = 1$, it agrees with the Jacobi definition.

Proposition 4.9. *Let $\mathfrak{s}$ be a state of the system, and let*

$$G = \begin{Bmatrix} a_{1,1} & & a_{1,2} & & \cdots & & a_{1,n} \\ & a_{2,1} & & \cdots & & a_{2,n-1} & \\ & & \ddots & & \iddots & & \\ & & & a_{n,1} & & & \end{Bmatrix}$$

be the corresponding strict Gelfand–Tsetlin pattern. Let $A_i = \sum_j a_{i,j}$ be the row sums. Then the Boltzmann weight $\beta(\mathfrak{s})$ equals a polynomial in q times the monomial $\mathbf{z}^\mu$ where

$$\mu = (A_1 - A_2, A_2 - A_3, \ldots, A_n).$$

Proof of Proposition 4.9. To prove the proposition, we note from the Boltzmann weights that $\beta(\mathfrak{s})$ is a polynomial in q times a monomial $\mathbf{z}^\mu$ for some μ. There is a contribution of z_i from every vertex of type $\mathtt{a}_2$, $\mathtt{b}_2$ or $\mathtt{c}_1$. These are precisely the vertices with a $\ominus$ spin to the left of the vertex. Therefore the number of z_i in the product of local Boltzmann weights equals the number μ_i of $\ominus$ spins in the i-th row, not counting the right boundary edge.

We must show that $\mu_i = A_i - A_{i+1}$ (or just A_i if $i = n$). To count the number of $\ominus$ spins on the horizontal edges in the i-th row, not counting the right boundary edge, we enumerate them by the paths. We note that one path enters from the top in the column $a_{i,j}$ and exits at the column $a_{i+1,j}$. There are $a_{i,j} - a_{i+1,j}$ $\ominus$ spins on this edge.

The argument requires minor modification for the last row, in which the last remaining path exits the right and contributes $a_{i,n+1-i}$. We do not need to consider this an exception if we extend the Gelfand–Tsetlin pattern by zero and define $a_{i+1,n+1-i} = 0$. With this convention, $A_{n+1} = 0$.

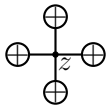
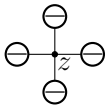
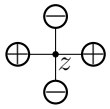
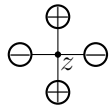
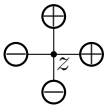
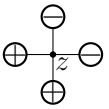
Summing the contributions of all paths,

$$\mu_i = \sum_{j=1}^{n+1-i} a_{i,j} - a_{i+1,j} = A_i - A_{i+1},$$

as required. □

4. Tokuyama Ice: $q = 1$

If either $q = 0$ or $q = 1$, one of the six allowed configurations in the Tokuyama model disappears. In these two cases, there are only five allowed states of spins adjacent to a vertex, and we will call the resulting models *five-vertex models*. In the case $q = 1$, the Boltzmann weights are:

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
					
1	z	-1	z	0	1

We see that there can no longer be any $\mathtt{c}_1$ patterns. This has a profound effect on the paths and on the Gelfand–Tsetlin patterns.

Lemma 4.10. *If G is the Gelfand–Tsetlin pattern of a state having no $\mathtt{c}_1$ patterns, then every row of the Gelfand–Tsetlin pattern is a subset of the row above, obtained by deleting one entry.*

Proof. If the $(i+1)$-st row is *not* obtained from the i-th row by deleting a single entry, then there is an element $a_{i+1,j}$ that is not in the i-th row. Since $a_{i,j} \geqslant a_{i+1,j} \geqslant a_{i,j+1}$ by the definition of a Gelfand–Tsetlin pattern we must have $a_{i,j} > a_{i+1,j} > a_{i,j+1}$. This implies that there is a $\mathtt{c}_1$ pattern in the i-th row at column $a_{i+1,j}$, which is a contradiction. □

Recall that the "Weyl group" W is the symmetric group S_n and we have $\rho = (n-1, n-2, \ldots, 1, 0)$.

Proposition 4.11. *When $q = 1$, we have*

$$Z_\lambda(\mathbf{z}; 1) = \sum_{w \in W} \operatorname{sgn}(w) \mathbf{z}^{w(\lambda+\rho)}. \tag{4.11}$$

Proof. There are $n!$ states $\mathfrak{s}$ that omit $\mathtt{c}_1$ patterns, namely those in which each row is obtained from the previous one by dropping a single entry. By Proposition 4.9, the Boltzmann weight $\beta(\mathfrak{s})$ is $\pm\mathbf{z}^\mu$, where $\mu_i = A_i - A_{i+1}$. By Lemma 4.10, this value $A_i - A_{i+1}$ is some element of the i-th row, hence of the top row $\lambda+\rho$. (The sign $-$ is the number of $\mathtt{b}_1$ patterns.) We may therefore write $\mu = w(\lambda+\rho)$ for some permutation $w \in W$, and $\beta(\mathfrak{s}) = \pm\mathbf{z}^{w(\lambda+\rho)}$, where the sign must be determined.

We have proved in Theorem 4.3 that

$$S_\lambda(\mathbf{z}) = \frac{Z_\lambda(\mathbf{z}; 1)}{\prod_{i<j}(z_i - z_j)} \tag{4.12}$$

is symmetric. The denominator is alternating, that is, it changes sign when an odd permutation is applied. Therefore the numerator $Z_\lambda(\mathbf{z}; 1)$ is also alternating. Now there is one state which has no $\mathtt{b}_1$ patterns: this is the state in which the entry in the i-th row of the Gelfand–Tsetlin pattern G that is dropped is always the first one. For this state, $\beta(\mathfrak{s}) = \mathbf{z}^{\lambda+\rho}$. Therefore $Z_\lambda(\mathbf{z}; 1)$ is of the form $\sum_{w\in W} \pm\mathbf{z}^{w(\lambda+\rho)}$, is known to be alternating, and one of the terms is $\mathbf{z}^{\lambda+\rho}$. Hence the signs of the other terms are determined. This proves (4.11). □

Theorem 4.12. *The symmetric polynomial S_λ agrees with the Schur polynomial s_λ by its first definition.*

Proof. We recognize the numerator and denominator in the ratio (4.12)

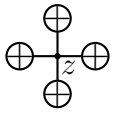
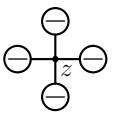
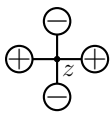
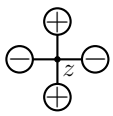
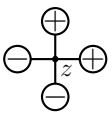
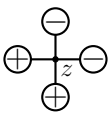
$$S_\lambda(\mathbf{z}) = \frac{\sum_{w\in W} \pm\mathbf{z}^{w(\lambda+\rho)}}{\prod_{i<j}(z_i - z_j)} = \frac{\det(z_j^{\lambda_i+n-i})}{\det(z_j^{n-i})},$$

using the Vandermonde identity. □

5. The crystal limit: $q = 0$

Let us begin with an explanation of the importance of the case $q = 0$. Before the 1980's, an analogy between the representation theory of $\mathrm{GL}(n, \mathbb{C})$ and the theory of semistandard Young tableaux (SSYT) emerged in work of Robinson, Littlewood, Schensted, Knuth, Lascoux and Schützenberger. For example, if λ is a partition, then λ indexes both an irreducible representation $\pi_\lambda^{\mathrm{GL}(n)}$ of $\mathrm{GL}(n, \mathbb{C})$ and the set $\mathcal{B}_\lambda$ of semistandard Young tableaux. The cardinality of $\mathcal{B}_\lambda$ equals the dimension of $\pi_\lambda^{\mathrm{GL}(n)}$, and this is the beginning of a fruitful parallel. Ultimately Kashiwara Kashiwara 1991, in the theory of crystal bases (crystals) gave an explanation for this: the representation $\pi_\lambda^{\mathrm{GL}(n)}$ can be thought of as being in a family of modules of the quantum groups $U_q(\mathfrak{gl}_n)$. These are somewhat complicated objects, but in the "crystal limit" $q \longrightarrow 0$ much of the complexity disappears, and the combinatorial theory remains. The quantum group $U_q(\mathfrak{gl}_n)$ does not, itself, have a limit when $q = 0$, but some of its operations do survive, giving $\mathcal{B}_\lambda$ some extra structure, that of a crystal. We will therefore refer to the case $q \longrightarrow 0$ as the "crystal limit." Crystals are introduced in more depth in Section 8.

When $q = 0$, we have the following Boltzmann weights:

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
					
1	z	0	z	z	1

Now we see that the pattern $\mathtt{b}_1$ no longer appears. This means that every path that comes down to a vertex from the top must bend to the right.

Lemma 4.13. *Let $\mathfrak{s}$ be a state of the system $\mathfrak{S}_\lambda(\mathbf{z}; q)$, and let*

$$G = \begin{Bmatrix} a_{1,1} & & a_{1,2} & & \cdots & & a_{1,n} \\ & a_{2,1} & & \cdots & & a_{2,n-1} & \\ & & \ddots & & \iddots & & \\ & & & a_{n,1} & & & \end{Bmatrix},$$

be the corresponding strict Gelfand–Tsetlin pattern. Then a necessary and sufficient condition that $\mathfrak{s}$ contains no $\mathtt{b}_1$ patterns is that for every i, j we have $a_{i,j} > a_{i+1,j}$.

Proof. In terms of the paths, one path descends from above to the vertex in the i-th row in column $a_{i,j}$ and leaves downwards in the column $a_{i+1,j}$. Thus if $a_{i,j} = a_{i+1,j}$, that means precisely that the vertex in row i and column $a_{i,j}$ produces a $\mathtt{b}_1$ pattern. □

We will call a Gelfand–Tsetlin pattern *left-strict* if its entries satisfy $a_{i,j} > a_{i+1,j} \geqslant a_{i,j+1}$. (The second inequality is part of the definition of a Gelfand–Tsetlin pattern, so the significant assumption is that $a_{i,j} > a_{i+1,j}$.) We see that the states of the five-vertex model $\mathfrak{S}_\lambda(\mathbf{z}; 0)$ are in bijection with the left-strict Gelfand–Tsetlin patterns with top row $\lambda + \rho$.

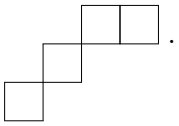
Let us denote by ρ_k the vector $(k-1, k-2, \ldots, 0)$ in $\mathbb{Z}^k$, so that $\rho = \rho_n$ in our previous notation. We can make a Gelfand–Tsetlin pattern with rows $\rho_n, \rho_{n-1}, \ldots, \rho_1$ thus:

$$P = \left\{ \begin{array}{ccccccc} n-1 & & n-2 & & & \cdots & 0 \\ & n-2 & & & \cdots & 0 & \\ & & \ddots & & \iddots & & \\ & & & 0 & & & \end{array} \right\}. \tag{4.13}$$

Lemma 4.14. *The map $G \longrightarrow G - P$ is a bijection between left-strict Gelfand–Tsetlin patterns with top row $\lambda + \rho$ and Gelfand–Tsetlin patterns with top row λ.*

Proof. This is easy to check. □

Let λ, μ be two partitions. Recall that if the Young diagram $\mathrm{YD}(\mu)$ is contained in the Young diagram $\mathrm{YD}(\lambda)$, then λ/μ is called a *skew shape*. Its Young diagram is the set-theoretic difference $\mathrm{YD}(\lambda) - \mathrm{YD}(\mu)$. For example $(5,3,2)/(3,2,1)$ is a skew shape and its diagram is



.

Recall the bijection between Gelfand–Tsetlin patterns with top row λ and semistandard Young tableaux of shape λ described in Chapter 3, Section 2.

Lemma 4.15. *Let G be a Gelfand–Tsetlin patterns with top row λ and T the corresponding semistandard Young tableaux of shape λ. The weight of T is*

$$\mathrm{wt}(T) = (A_n, A_{n-1} - A_n, \ldots, A_2 - A_3, A_1 - A_2),$$

where A_i are the row sums of G.

Proof. Let $\lambda^{(i)}$ be the partitions in the rows of G. By definition, $\mathrm{wt}(T) = (\mu_1, \mu_2, \ldots, \mu_n)$ where μ_i is the number of boxes in T that contain the entry i. These comprise the skew tableau with shape $\lambda^{(n+1-i)}/\lambda^{(n+2-i)}$, and since $|\lambda^{(i)}| = A_i$, we obtain the advertised formula for $\mathrm{wt}(T)$. □

Example 4.16. To illustrate Lemma 4.15, suppose $n = 3$ and

$$G = \left\{ \begin{array}{ccccc} 5 & & 3 & & 1 \\ & 4 & & 1 & \\ & & 3 & & \end{array} \right\} \qquad \text{so} \qquad T = \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 2 & 3 \\ \hline 2 & 3 & 3 \\ \cline{1-3} 3 \\ \cline{1-1} \end{array} .$$

The three skew shapes corresponding to $1, 2, 3$ are

$$(3)/\varnothing, \qquad (4,1)/3, \qquad (5,3,1)/(4,1)\ .$$

So $\mathrm{wt}(T) = (3,2,4)$, which agrees with the formula $(3, 5-3, 9-5)$ from Lemma 4.15

We let w_0 be the "long element" of the Weyl group $W = S_n$, which is the permutation that maps k to $n+1-k$ of $\{1, 2, 3, \ldots, n\}$. If $\mathbf{z} = (z_1, \ldots, z_n) \in (\mathbb{C}^\times)^n$ and $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{Z}^n$, then $w_0\mathbf{z} = (z_n, \ldots, z_1)$ and $w_0\mu = (\mu_n, \ldots, \mu_1)$. Obviously $\mathbf{z}^{w_0\mu} = (w_0\mathbf{z})^\mu$.

Proposition 4.17. *Let $\mathfrak{s}$ be an admissible state of the system $\mathfrak{S}_\lambda(\mathbf{z}; 0)$. Since $\mathfrak{s}$ has no b_1 patterns, the corresponding Gelfand–Tsetlin pattern G is left-strict. Let $G^\circ = G - P$, which is a Gelfand–Tsetlin pattern with top row λ. Let T be the semistandard Young tableau associated with G° as in Lemma 4.15. Then $\beta(\mathfrak{s}) = \mathbf{z}^\rho \cdot (w_0\mathbf{z})^{\mathrm{wt}(T)}$.*

Proof. Since the Boltzmann weights of every vertex can only be 1 or z_i for some i, it is obvious that $\beta(\mathfrak{s})$ is a monomial $\mathbf{z}^\mu$ and we need to compute μ. This is accomplished by Proposition 4.9. Writing $G = P + G^\circ$ the contribution of P is obviously $\mathbf{z}^\rho$, and we must discuss the contribution of G^o but by Lemma 4.15 and Proposition 4.9, this is $\mathbf{z}^{w_0\,\mathrm{wt}(T)} = (w_0\mathbf{z})^{\mathrm{wt}(T)}$. □

Theorem 4.18. *The polynomial $S_\lambda = s_\lambda$ where s_λ is the Schur function defined by the second combinatorial definition.*

Proof. To summarize what we have done so far, culminating in Proposition 4.17, we have seen that every state $\mathfrak{s}$ of $\mathfrak{S}_\lambda(\mathbf{z}; 0)$ has no $\mathfrak{b}_1$ patterns. Such states are parametrized by left-strict Gelfand–Tsetlin patterns with top row $\lambda + \rho$. Each such pattern G can be written as $G^\circ + P$ where G° is a Gelfand–Tsetlin pattern with top row λ. If T is tableau corresponding to G° then $\beta(\mathfrak{s}) = \mathbf{z}^\rho \cdot (w_0\mathbf{z})^{\mathrm{wt}(T)}$. Summing over all states and using the combinatorial definition of the Schur function we obtain

$$Z_\lambda(\mathbf{z}; 0) = \mathbf{z}^\rho\, s_\lambda(w_0\mathbf{z}).$$

On the other hand, we have shown for all q that

$$Z_\lambda(\mathbf{z}; q) = \left(\prod_{i<j}(z_i - qz_j)\right) S_\lambda(\mathbf{z}).$$

When $q = 0$, the product becomes $z_1^{n-1}z_2^{n-2}\cdots = \mathbf{z}^\rho$. Comparing gives

$$S_\lambda(\mathbf{z}) = s_\lambda(w_0\mathbf{z}).$$

We may replace $\mathbf{z}$ by $w_0\mathbf{z}$ and remember that we proved (using the Yang–Baxter equation) that S_λ is symmetric, so $S_\lambda = s_\lambda$. □

Comparing the evaluations of $S_\lambda(\mathbf{z})$ when $q = 1$ and $q = 0$, we have now proved the equivalence of the two definitions of the Schur function.

6. Column parameters

In this section, we consider models whose partition functions are *factorial Schur functions*. The models can be used to develop properties of these, but we only use the models to illustrate Theorem 2.17 of Chapter 2.

In the Tokuyama models, the Boltzmann weights depend on the rows but not the columns. Recall, however, Theorem 2.17 in Chapter 2 which classifies solvable lattice models in which all the Boltzmann weights are controlled by a single parametrized Yang–Baxter equation. In this theorem, we see that the Boltzmann weights can depend on two sets of parameters, the α's which are row parameters, and the β's which are column parameters. Thus we certainly have the option of introducing column parameters to the Tokuyama models.

Macdonald 1992 pointed out that Schur polynomials have a wide set of generalizations, and to prove his point, he gave nine variations. As Macdonald shows, these generalizations exhibit a certain set of properties, such as the Cauchy identity, Jacobi–Trudi identities, both with dual forms, the Pieri rule and the Giambelli formula. Further such variations were found by Okounkov and Olshanski 1998; Okunḱov and Olśhanskiĭ 1997, Okounkov 1998 and Molev 1998.

Lattice model approaches to such generalized Schur functions based on the free-fermionic Yang–Baxter equation were taken by Bump, McNamara, and Nakasuji 2014, Motegi 2017a, Aggarwal, Borodin, Petrov, and Wheeler 2023 and Naprienko 2024. The last two references both treat (differently) four-parameter families of generalized Schur functions. The parametrization in Naprienko 2024 is very elegant.

Macdonald's sixth variation is the *factorial Schur functions*. These were introduced by Biedenharn and Louck 1989 and by Chen and Louck 1993. Schur polynomials, of course, are symmetric functions in a set of variables $\mathbf{z} = (z_1, z_2, \dots)$, and the factorial Schur functions generalize a second set of parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots)$. They arise naturally in algebraic geometry in the torus equivariant cohomology of Grassmannians. Bump, McNamara, and Nakasuji 2014 gave lattice model proofs of many of their properties. In these models, the parameters z_i are associated with the rows of the model, and the parameters α_i are associated with the columns.

Let $G = \mathrm{GL}(2,\mathbb{C}) \otimes \mathbb{C}^\times$ be the parameter group. By Corollary 2.16 and Theorem 2.17 of Chapter 2, if $\gamma_{i,j} \in \mathrm{GL}(2,\mathbb{C}) \times \mathbb{C}^\times$ are the elements corresponding to the Boltzmann weights at each vertex, then $\gamma_{bd}\gamma_{ad}^{-1}$ depends on a and b but is independent of d. We recall that the theorem shows that solvability amounts to writing $\gamma_{i,j} = \phi(i)\psi(j)$ in G, where $\phi : \{1, \dots, n\} \to G$ is a function of the row, and ψ is a function of the column.

Now let us consider the Tokuyama model. Let $\phi(i)$ and $\psi(j) \in G$ be as in Theorem 2.17, so $\gamma_{ij} = \phi(i)\psi(j)$ where $\psi(j) = 1_G$ can be omitted, since the Boltzmann weights do not depend on the column. Thus $\gamma_{ij} = \phi(i)$. The Boltzmann weights are given by the following table:

$\mathtt{a}_1(z_i)$	$\mathtt{a}_2(z_i)$	$\mathtt{b}_1(z_i)$	$\mathtt{b}_2(z_i)$	$\mathtt{c}_1(z_i)$	$\mathtt{c}_2(z_i)$
1	z_i	$-q$	z_i	$(1-q)z_i$	1

Therefore

$$\phi(i) = \left(\begin{pmatrix} \mathtt{a}_1(z_i) & \mathtt{b}_2(z_i) \\ -\mathtt{b}_1(z_i) & \mathtt{a}_2(z_i) \end{pmatrix}, \mathtt{c}_1(z_i) \right) = \left(\begin{pmatrix} 1 & z_i \\ q & z_i \end{pmatrix}, (1-q)z_i \right).$$

Now we may perturb the Tokuyama model by allowing the column parameter $\psi(j)$ to be nontrivial. The choice to give factorial Schur polynomials is

$$\psi(j) = \left(\begin{pmatrix} 1 & \alpha_j \\ & 1 \end{pmatrix}, 1 \right).$$

Then

$$\gamma_{ij} = \phi(i)\psi(j) = \left(\begin{pmatrix} 1 & z_i + \alpha_j \\ q & z_i + q\alpha_j \end{pmatrix}, (1-q)z_i \right).$$

This leads to the following modification of the Tokuyama weights:

$\mathtt{a}_1(z_i, \alpha_j)$	$\mathtt{a}_2(z_i, \alpha_j)$	$\mathtt{b}_1(z_i, \alpha_j)$	$\mathtt{b}_2(z_i, \alpha_j)$	$\mathtt{c}_1(z_i, \alpha_j)$	$\mathtt{c}_2(z_i, \alpha_j)$
1	$z_i + qa_j$	$-q$	$z_i + a_j$	$(1-q)z_i$	1

As with Tokuyama ice, we use the extended wall boundary conditions with the z_i as column parameters. We use z_i in the i-th row, but since our policy is to number the columns $0, 1, 2, \dots$ but the column parameters are $\alpha_1, \alpha_2, \dots$, we put α_j in the $j-1$ column. Let us denote the partition function $Z_\lambda(\mathbf{z}, \boldsymbol{\alpha})$.

Theorem 4.19 (Bump, McNamara, and Nakasuji 2014)**.** *The partition function is*

$$Z_\lambda(\mathbf{z}, \boldsymbol{\alpha}) = \prod_{i<j} (z_i - qz_j)\, s_\lambda(\mathbf{z} \| \boldsymbol{\alpha}; q),$$

where $s_\lambda(\mathbf{z}\|\boldsymbol{\alpha};q)$ is a polynomial that is symmetric in the z_i (though not in the α_i) that is independent of q.

Proof. This is left to the reader (Exercise 4.6). □

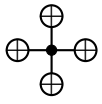
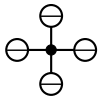
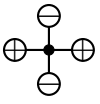
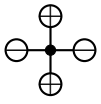
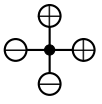
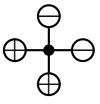
The polynomials in Theorem 4.19 are the *factorial Schur functions* of Macdonald's sixth variation in Macdonald 1992. The notation $s_\lambda(\mathbf{z}\|\boldsymbol{\alpha};q)$ is consistent with Macdonald 1992 but differs from other authors, including Naprienko 2024. See Bump, McNamara, and Nakasuji 2014 for further use of the lattice model representation to develop properties of these polynomials and Aggarwal, Borodin, Petrov, and Wheeler 2023; Naprienko 2024 for more general free-fermionic models that generalize Schur functions.

7. Gamma, Delta and hybrid models

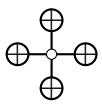
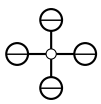
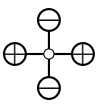
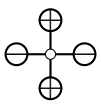
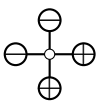
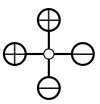
In the last section we introduced column parameters to the free-fermionic Tokuyama models, producing models for factorial Schur functions. We continue to investigate these models, introducing a new phenomenon. The models that we have considered so far in this chapter will be called *Gamma models*; in these the paths move down and to the right. We now mix these with "dual" models that we call *Delta* models in which the paths move down and to the right. We are interested in the relationship between these two types of models, and *hybrid* models in which layers of Gamma "ice" are layered with layers of Delta "ice." Paths move down and to the right on the Gamma layers, and down and to the left on the Delta layers. Gamma and delta layers may be interchanged using the Yang-Baxter equation,

Our main result (Theorem 4.21) is the special case $q = 0$ of a more general result in Chapter 19 of Brubaker, Bump, and Friedberg 2011b or the arxiv version of Brubaker, Bump, and Friedberg 2011a. Other works where layers of Gamma and Delta ice are mixed together include Brubaker, Buciumas, Bump, and Gustafsson 2020b; Gray 2017; Gustafsson and Westerlund 2025; Ivanov 2012; Motegi 2017b; Zhong 2022. We will discuss further analogs in Section 4.

To see this phenomenon in the simplest case, we will take Gamma ice to be the $q = 0$ Tokuyama model. However, we will generalize these slightly by including column parameters, so these are actually models for factorial Schur functions, as in the last section. Thus, as in the last section, supplementing the spectral parameters z_i for each row, for each column (numbered j) we pick a complex number α_j and we use the following Boltzmann weights.

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
					
1	z_i	0	$z_i+\alpha_j$	z_i	1

The Boltzmann weights for the *Delta model* are given as follows.

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{d}_1$	$\mathtt{d}_2$
					
$z_i+\alpha_j$	1	$-\alpha_j$	1	1	z_i

Note that instead of $\mathtt{c}_1$ and $\mathtt{c}_2$ we have two $\mathtt{d}_1$ and $\mathtt{d}_2$ weights. Because of this difference, for the Gamma weights the paths move down and to the right, but for the Delta weights the paths move down and to the left. To distinguish these two types of weights we will use a solid $\bullet$ for Gamma vertices and a white $\circ$ for Delta vertices.

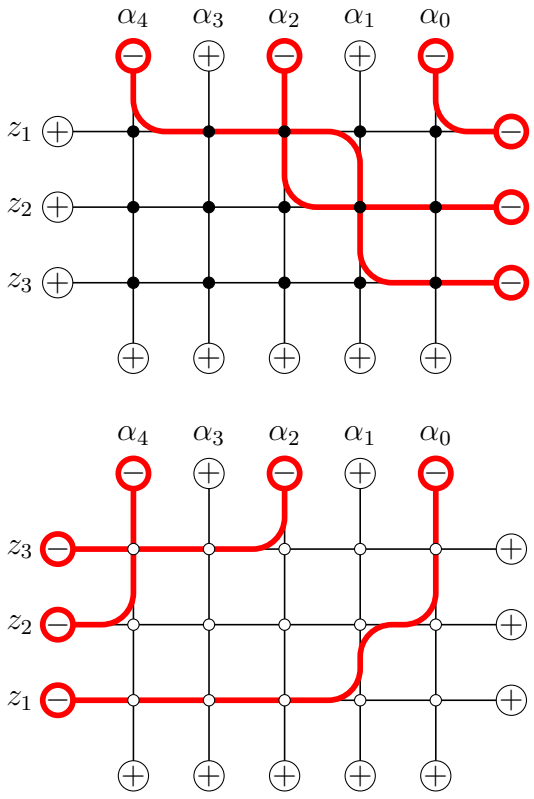
We fix a partition $\lambda = (\lambda_1, \ldots, \lambda_n)$ and spectral parameters $(z_1, \ldots, z_n)$, and column parameters α_j for $j = 0, 1, 2, \ldots$. We have already described the boundary conditions for the Gamma model: we put a $\ominus$ spin at the top boundary edge in columns $\lambda_j + n - j$ and $\oplus$ spins in the remaining columns, $\oplus$ spins on the left and bottom boundary edges, and $\ominus$ spins on the right edges. We use the spectral parameter z_i in the i-th row and the parameter α_j in the j-th column.

For the Delta model, there are two modifications to the Gamma R-matrix. We use the Delta weights, and the top and bottom edge boundary spins are the same as the Gamma R-matrix, but for the Delta system:

- We put $\ominus$ spins on the left edges and $\oplus$ spins on the right and bottom edges; and
- We reverse the order of the parameters in the row, putting z_i in the $n + 1 - i$ row.

Let $\mathfrak{S}^{\Gamma}_{\lambda}(\mathbf{z}; \alpha)$ and $\mathfrak{S}^{\Delta}_{\lambda}(\mathbf{z}; \alpha)$ denote the Gamma and Delta systems as described above. Here $\alpha = (\alpha_0, \alpha_1, \alpha_2, \ldots)$ represents the sequence of column parameters.

Example 4.20. Consider $\lambda = (2, 1, 0)$, so that $\lambda + \rho = (4, 2, 0)$. Here are sample states for the two systems; in the Gamma system we use $\bullet$ to label the vertices. Observe that the Delta state contains a $\mathtt{b}_1$ vertex in column 0, which is allowed provided $\alpha_0 \neq 0$. If $\alpha_0 = 0$, the Boltzmann weight would be zero. For Gamma weights, $\mathtt{b}_1$ states are never allowed.



(4.14)

Let $Z(\mathfrak{S}^{\Gamma}_{\lambda}(\mathbf{z}; \alpha))$ and $Z(\mathfrak{S}^{\Delta}_{\lambda}(\mathbf{z}; \alpha))$ denote the partition functions of the Gamma and Delta systems. The main result of this section is the following theorem. a more general result in

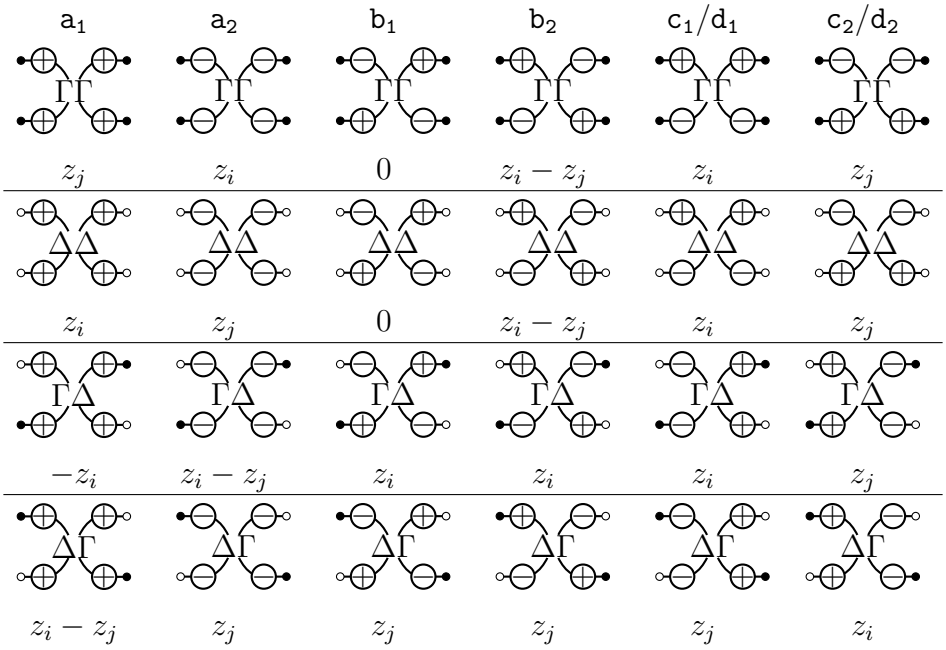
Chapter 19 of Brubaker, Bump, and Friedberg 2011b or the arxiv version of Brubaker, Bump, and Friedberg 2011a. Other works where layers of Gamma and Delta ice are mixed together include Brubaker, Buciumas, Bump, and Gustafsson 2020b; Gray 2017; Gustafsson and Westerlund 2025; Ivanov 2012; Motegi 2017b; Zhong 2022. We will discuss further analogs in Section 4.

Theorem 4.21. *The partition functions are equal:*

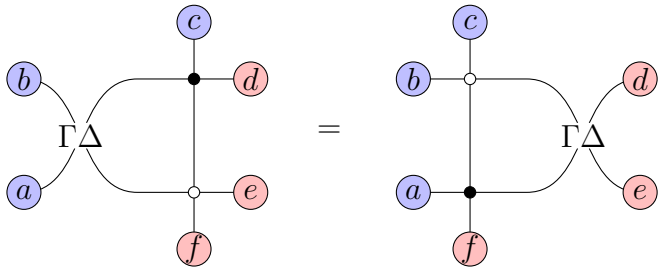
$$Z(\mathfrak{S}^{\Gamma}_{\lambda}(\mathbf{z};\alpha)) = Z(\mathfrak{S}^{\Delta}_{\lambda}(\mathbf{z};\alpha)).$$

Remark 4.22. If the column parameters $\alpha = 0$, then by Theorem 4.18 we have $Z(\mathfrak{S}^{\Gamma}_{\lambda}(\mathbf{z};0)) = \mathbf{z}^{\rho}s_{\lambda}(\mathbf{z})$ in terms of the Schur polynomial $s_{\lambda}(\mathbf{z})$. This may also be proved by similar arguments for $Z(\mathfrak{S}^{\Delta}_{\lambda}(\mathbf{z};0))$, establishing the theorem when $\alpha = 0$. We will give a different argument that is valid for general α below.

Rows of Gamma ice and Delta ice may be freely layered, and Yang–Baxter equations exist to permute them. Note that we have labeled Gamma ice with a black dot • and Delta ice with a white dot ∘. We preserve this notation to indicate the four R-matrices required for these interchanges.

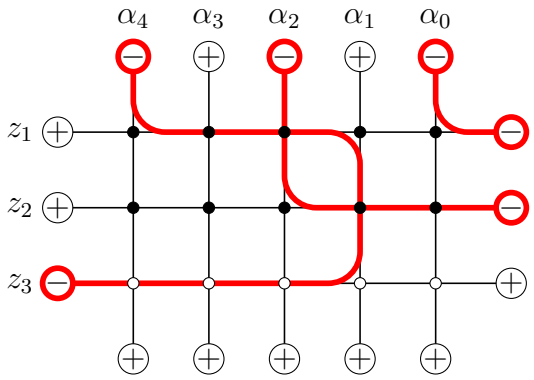


The R-matrix that we have labeled ΓΓ can be recognized as (4.6) with $q = 0$ and $z = z_i$, $w = z_j$. In general if $X, Y \in \{\Gamma, \Delta\}$ then XY represents an R-matrix that interchanges an X vertex with a Y vertex. For example if $XY = \Gamma\Delta$, the Yang–Baxter equation looks like this:

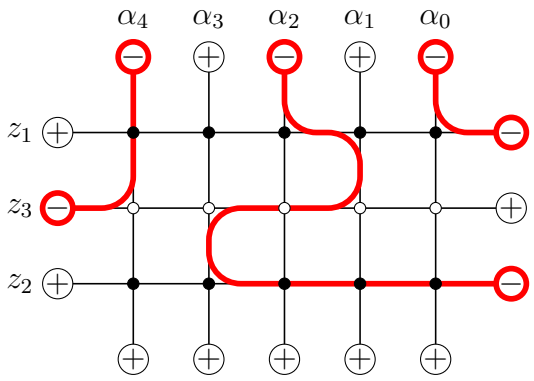


Note that none of these R-matrices depend on α, and this is to be expected in view of the proof of Theorem 2.15 in Chapter 2.

Proof of Theorem 4.21. We begin by noting that we can replace the bottom row of Gamma ice by Delta ice. This entails changing all the $\ominus$ spins of the row to $\oplus$ and vice versa. To see that this works, we note that in the bottom row only $\oplus$ spins can occur limiting the possible states to $\mathtt{a}_1$, $\mathtt{b}_2$ and $\mathtt{c}_2$, and for these three spins (which then become $\mathtt{b}_2$, $\mathtt{a}_1$ and $\mathtt{d}_1$ in the Delta system) this transformation leaves the Boltzmann weights unchanged. This step in the proof can be done state-by-state giving a bijection between the Gamma system and a hybrid system that has one row of Delta ice at the bottom. The Gamma state in Example 4.20 becomes:



Now we wish to move the Delta row with the z_n spectral parameter up. To do this, we attach the $\Gamma\Delta$ R-matrix for and apply the train argument. Attaching the R-matrix at the left multiplies the partition function by z_{n-1}, while detaching it on the right divides the partition function by z_{n-1}, so there is non net change. For this step there is no bijection between states, but here is a typical state for the resulting hybrid system:



We may repeat this process until the Δ row (carrying the parameter z_n) is at the top. Then we have another Γ row, with parameter z_n at the bottom. We may again change this into a row of Δ ice, and apply the train argument to move it up and to the top. When all rows have been transformed into Δ layers, we find that the order of the spectral parameters has been reversed. $\square$

8. Crystals

Crystals are combinatorial analogs of representations of Lie groups. Kashiwara 1991 introduced them as certain limits of modules of quantum groups. Independently, Lusztig 1990 studied canonical bases from a geometric perspective.

Let Λ be the weight lattice of a reductive complex Lie group G. Thus if T is a maximal torus of G we may identify Λ with the group of rational characters of T. If $\mathbf{z} \in T$ and $\lambda \in \Lambda$ we denote by $\mathbf{z}^\lambda$ the application of λ to $\mathbf{z}$. Let $\Phi \subset \Lambda$ be the root system. We will denote by s_i the simple reflections in the Weyl group W, and by $\alpha_i \in \Phi$ the corresponding simple roots for $i \in I$, where I is the index set of the underlying root system.

Following Kashiwara 1991, we now define crystals. We require a set $\mathcal{C}$ with a map $\mathrm{wt} : \mathcal{C} \longrightarrow \Lambda$ and some other structure. Let 0 be an auxiliary element (not in $\mathcal{C}$). For each index $i \in I$ there will be maps $e_i, f_i : \mathcal{C} \longrightarrow \mathcal{C} \cup \{0\}$ such that if $x, y \in \mathcal{C}$ then $f_i(x) = y$ if and only if $e_i(y) = x$, and if this is true then $\mathrm{wt}(y) = \mathrm{wt}(x) - \alpha_i$. Furthermore there will be maps $\varepsilon_i, \varphi_i : \mathcal{C} \longrightarrow \mathbb{Z}$ such that $\varepsilon_i(y) = \varepsilon_i(x) - 1$ and $\varphi_i(y) = \varphi_i(x) + 1$, and $\varphi_i(x) - \varepsilon_i(x) = \langle \alpha_i^\vee, \mathrm{wt}(x) \rangle$. The crystal is *seminormal* if $\varepsilon_i(x) \geqslant 0$ is always the largest integer k such that $e_i^k(x) \neq 0$, and $\varphi_i(x) \geqslant 0$ is always the largest integer m such that $f_i^m(x) \neq 0$.

This is not the most general definition, since Kashiwara allows the maps ε_i and φ_i to take the value $-\infty$ but although this is sometimes useful, such crystals will not appear in this book, so we do not allow it.

If $\mathcal{C}$ and $\mathcal{D}$ are crystals, a tensor product crystal $\mathcal{C} \otimes \mathcal{D}$ is defined. We will follow the conventions in Bump and Schilling 2017 which differ from the conventions in Kashiwara's papers. (See for example Kashiwara 1995 for a useful survey.) The tensor product $\mathcal{C} \otimes \mathcal{D}$ consists of pairs $x \otimes y$ with $x \in \mathcal{C}$ and $y \in \mathcal{D}$ with $\mathrm{wt}(x \otimes y) = \mathrm{wt}(x) + \mathrm{wt}(y)$ and

$$f_i(x \otimes y) = \begin{cases} f_i(x) \otimes y & \text{if } \varphi_i(y) \leqslant \varepsilon_i(x), \\ x \otimes f_i(y) & \text{if } \varphi_i(y) > \varepsilon_i(x), \end{cases}$$

$$e_i(x \otimes y) = \begin{cases} e_i(x) \otimes y & \text{if } \varphi_i(y) < \varepsilon_i(x), \\ x \otimes e_i(y) & \text{if } \varphi_i(y) \geqslant \varepsilon_i(x), \end{cases}$$

$$\varphi_i(x \otimes y) = \max(\varphi_i(x), \varphi_i(x) + \varphi_i(y) - \varepsilon_i(x)),$$

$$\varepsilon_i(x \otimes y) = \max(\varepsilon_i(y), \varepsilon_i(y) + \varepsilon_i(x) - \varphi_i(y)).$$

The tensor product is associative.

If $\mathcal{C}$ is a crystal, we may define its *character*

$$\chi_{\mathcal{C}}(\mathbf{z}) = \sum_{x \in \mathcal{C}} \mathbf{z}^{\mathrm{wt}(x)}.$$

Given a crystal $\mathcal{C}$, the associated *crystal graph* is the directed labeled graph with vertices $\mathcal{C}$ where there is an edge $x \xrightarrow{i} y$ (labeled by the index $i \in I$) if $y = f_i(x)$. Note that we always draw the arrow from x to $f_i(x)$. We say $\mathcal{C}$ is *connected* if its crystal graph is connected. Each connected component is itself a crystal, so $\mathcal{C}$ is the disjoint union of its connected components.

If $u \in \mathcal{C}$ we call u a *highest weight element* (or vector) if $e_i(u) = 0$ for all $i \in I$. We call $\mathcal{C}$ a *highest weight crystal* if it is connected and has a unique highest weight element. Let $\lambda = \mathrm{wt}(u)$; we say λ is the *weight* of the highest weight crystal $\mathcal{C}$.

Lemma 4.23. *If $\mathcal{C}$ is a seminormal highest weight crystal of weight λ then λ is dominant.*

Proof. We must show that $\langle \alpha_i^\vee, \lambda \rangle \geqslant 0$ for every simple coroot $\alpha_i^\vee$. Indeed this equals

$$\langle \alpha_i^\vee, \mathrm{wt}(u) \rangle = \varphi_i(u) - \varepsilon_i(u) = \varphi_i(u) \geqslant 0$$

by seminormality. □

Weyl proved that if λ is a dominant weight then λ is the highest weight of a unique irreducible representation of G. Let $\chi_\lambda(\mathbf{z})$ be its character for $\mathbf{z} \in T$. The characters of irreducible representations are a basis of the W-invariant functions of $\mathbf{z}$.

Theorem 4.24 (Kashiwara, Lusztig, Littelmann)**.** *There is a class of crystals called normal having the following properties. A crystal is normal if and only if every connected component is normal. Every connected normal crystal is a highest weight crystal. For every dominant weight λ, there is a unique connected normal crystal with highest weight λ. The character $\chi_\mathcal{C}$ of a connected normal crystal of highest weight λ equals χ_λ. A tensor product of normal crystals is normal.*

Proof. Kashiwara's construction (see Kashiwara 1991) produces a crystal base for the corresponding representation of a quantum group, with parallel results in the work of Lusztig 1990. Littelmann's different approach (see Littelmann 1995b) uses the theory of "Littelmann paths" in the ambient real vector space of Λ. For a combinatorial proof see Bump and Schilling 2017, where the proof over several chapters begins with Theorem 5.20 and ends with Corollary 13.9. □

8.1. Crystals on tableaux and Gelfand–Tsetlin patterns. Now let us specialize to the case where $G = \mathrm{GL}(n)$ and $\Lambda = \mathbb{Z}^n$. Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition. Then λ is a dominant weight, and the unique normal crystal of shape λ is the crystal of tableaux $\mathcal{B}_\lambda$ defined by Kashiwara and Nakashima 1994. As a set, $\mathcal{B}_\lambda$ is $\mathsf{SSYT}_n(\lambda)$, but we need to describe the crystal operations. The weight operation is already defined: if T is a tableau then $\mathrm{wt}(T) = (\mu_1, \ldots, \mu_n)$ where μ_i is the number of entries equal to i. We must define the maps e_i, f_i, ε_i and φ_i.

First we consider the case where $\lambda = (k) = (k, 0, \ldots, 0)$, so $\mathcal{B}_\lambda$ consists of rows of length k, whose entries are weakly increasing. If $R = \boxed{r_1}\boxed{\cdots}\boxed{r_k} \in \mathcal{B}_{(k)}$ then we define $\varphi_i(R)$ to be the number of i's among the entries r_j, and $\varepsilon_i(R)$ to be the number of $i+1$'s. If $\varphi_i(R) > 0$ then $f_i(R)$ is obtained by replacing the rightmost i by $i+1$. And if $\varepsilon_i(R) > 0$, then $e_i(R)$ is obtained by replacing the leftmost $i+1$ by i. As a special case the crystal $\mathcal{B}_{(0)}$ is defined, with one element having weight 0, and $\mathcal{B}_{(0)} \otimes \mathcal{C} \cong \mathcal{C} \otimes \mathcal{B}_{(0)} \cong \mathcal{C}$ for any crystal $\mathcal{C}$.

We have now completely defined the crystal structure for the crystal of rows $\mathcal{B}_{(k)}$. With this in mind, we have an embedding of $\mathcal{B}_\lambda$ into $\mathcal{B}_{\lambda_n} \otimes \mathcal{B}_{\lambda_{n-1}} \otimes \cdots \otimes \mathcal{B}_{\lambda_1}$ as follows. If T is a semistandard tableau of shape λ, let $R_1, \ldots, R_n$ be the rows of T; then we map T to the "row reading" $\mathrm{RR}(T) \cong R_n \otimes \cdots \otimes R_1$. (We may omit R_j if $\lambda_j = 0$.)

Lemma 4.25. *The image of the map* RR *is a connected component of* $\mathcal{B}_{\lambda_n} \otimes \mathcal{B}_{\lambda_{n-1}} \otimes \cdots \otimes \mathcal{B}_{\lambda_1}$.

Proof. This follows easily from Proposition 3.1 in Bump and Schilling 2017. □

Now we give $\mathcal{B}_\lambda$ the unique crystal structure that makes RR the embedding of a subcrystal in $\mathcal{B}_{\lambda_n} \otimes \mathcal{B}_{\lambda_{n-1}} \otimes \cdots \otimes \mathcal{B}_{\lambda_1}$.

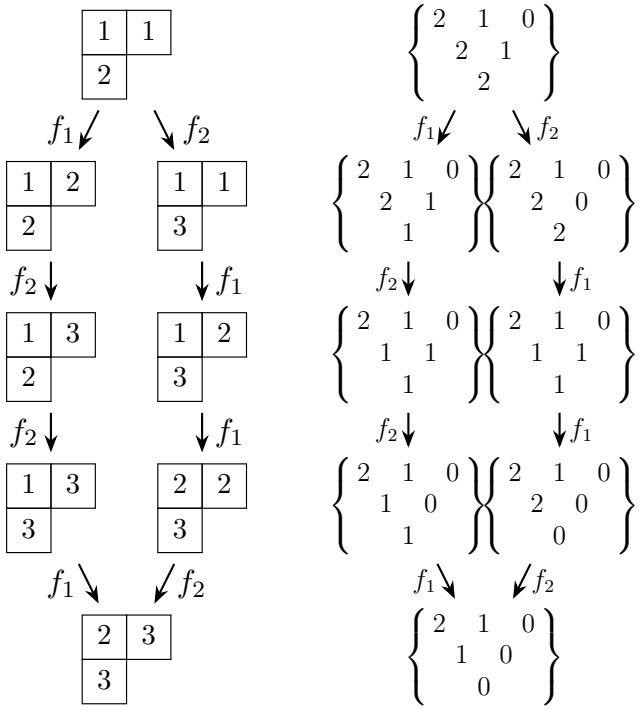


FIGURE 1. The GL(3) crystal $\mathcal{B}_{(2,1)}$, and its corresponding crystal of Gelfand–Tsetlin patterns.

Example 4.26. Let $n = 3$, and let us compute $f_2(T)$ where $T = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}$. If this is nonzero, f_2 will change one of the 2's to a 3, but which one is changed depends on a computation, which we will now perform. (See Bump and Schilling 2017 Section 2.4 for a more systematic approach.) Using the row reading the tableau is interpreted as

$$R_2 \otimes R_1 = x \otimes y, \qquad x = \begin{array}{|c|}\hline 2 \\ \hline\end{array},\ y = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array}$$

We find that $\varphi_2(y) = 1$, $\varepsilon_2(x) = 0$, and by the tensor product rule

$$f_2(T) = x \otimes f_2(y) = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}.$$

We have seen in Chapter 3 that the set $\mathsf{SSYT}_n(\lambda)$ of semistandard Young tableaux of shape λ is in bijection with the set $\mathsf{GTP}_n(\lambda)$ of Gelfand–Tsetlin patterns of size n with top row λ. To recapitulate, let $\lambda^{(i)}$ be the sequence of shapes such that the skew shape $\lambda^{(i)}/\lambda^{(i-1)}$ contains the entries equal to i. Then $\lambda^{(n)}, \ldots, \lambda^{(0)}$ are the rows of the Gelfand–Tsetlin pattern. If $G \in \mathsf{GTP}_n(\lambda)$ and $T \in \mathsf{SSYT}_n(\lambda)$ we write $T = \mathrm{SSYT}(G)$ and $G = \mathrm{GTP}(T)$.

Since $\mathsf{GTP}_n(\lambda)$ is in bijection with the crystal $\mathcal{B}_\lambda$, there is a crystal structure on $\mathsf{GTP}_n(\lambda)$ by transportation of structure. See Kirillov and Berenstein 1995 for much useful information about these crystals. In particular, their discussion of the Schützenberger involution is very important.

Example 4.27. Let $n = 3$ and $\lambda = (2, 1, 0)$. The crystal graphs for $\mathcal{B}_\lambda$ and $\mathsf{GTP}_n(\lambda)$ with corresponding elements drawn in corresponding locations are given in Figure 1.

8.2. Crystals for Tokuyama models. We have seen that for the six-vertex model with extended wall boundary conditions, the states may be parametrized by Gelfand–Tsetlin patterns. Then the crystal operators e_i and f_i become operations on the states. Let us show how this works for the $q = 0$ Tokuyama models.

We will make use of the *Schützenberger involution* on $\mathcal{B}_\lambda$, also called the *Lusztig involution*. It inverts the crystal graph, mapping the highest weight to the lowest weight.

Proposition 4.28. *There is a unique map* $\mathrm{Sch} : \mathcal{B}_\lambda \longrightarrow \mathcal{B}_\lambda$ *such that* $\mathrm{wt}(\mathrm{Sch}(T)) = w_0(\mathrm{wt}(T))$ *for* $T \in \mathcal{B}_\lambda$, *and such that* $\mathrm{Sch} \circ e_i = f_{n-i} \circ \mathrm{Sch}$ *and* $\mathrm{Sch} \circ f_i = e_{n-i} \circ \mathrm{Sch}$.

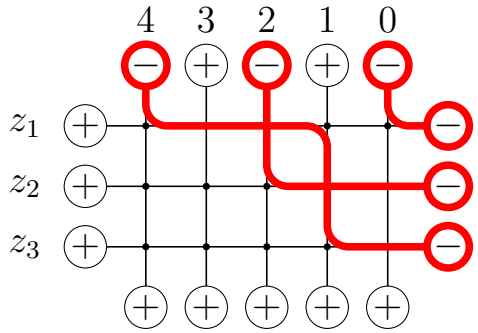
Proof. See Bump and Schilling 2017, Exercise 5.2. □

Let us first consider the $q = 0$ Tokuyama models from Section 5. We will describe a specific bijection between the states of this model and $\mathcal{B}_\lambda$. Let $\mathrm{GTP}(\mathfrak{s})$ denote the associated Gelfand–Tsetlin pattern of a state $\mathfrak{s}$. By Lemma 4.14, we may subtract the pattern P and obtain a Gelfand–Tsetlin pattern $\mathrm{GTP}^\circ(\mathfrak{s}) := \mathrm{GTP}(\mathfrak{s}) - P$. The map GTP° is a bijection between the states of $\mathfrak{S}_\lambda(\mathbf{z}; 0)$ and $\mathsf{GTP}_n(\lambda)$. Then it is beneficial to apply the Schützenberg involution to the corresponding tableau. Thus we define a map $\theta : \mathfrak{S}_\lambda(\mathbf{z}; 0) \longrightarrow \mathcal{B}_\lambda$ by

$$\theta(\mathfrak{s}) = \mathrm{Sch}(\mathrm{SSYT}(\mathrm{GPT}^\circ(\mathfrak{s}))). \tag{4.15}$$

This is a bijection, so $\mathfrak{S}_\lambda(\mathbf{z}; 0)$ becomes a crystal by transportation of structure.

Example 4.29. Let $n = 3$ and $\lambda = (2, 1, 0)$, so $\lambda + \rho = (4, 2, 0)$. We will show how to calculate $\theta(\mathfrak{s})$ for the following state:



We have used the color red to indicate which edges have $\ominus$ spin. We find that

$$\mathrm{GPT}(\mathfrak{s}) = \begin{Bmatrix} 4 & & 2 & & 0 \\ & 2 & & 1 & \\ & & 1 & & \end{Bmatrix}, \qquad \mathrm{GPT}^\circ(\mathfrak{s}) = \begin{Bmatrix} 2 & & 1 & & 0 \\ & 1 & & 1 & \\ & & 1 & & \end{Bmatrix},$$

and so

$$\mathrm{SSYT}(\mathrm{GTP}^\circ(\mathfrak{s})) = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 3 \\ \cline{1-1} \end{array}, \qquad \theta(\mathfrak{s}) = \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 3 \\ \cline{1-1} \end{array}.$$

The crystal structure on $\mathfrak{S}_\lambda(\mathbf{z}; 0)$ can then be computed from the known crystal $\mathcal{B}_\lambda$. Figure 2 shows the case $\lambda = (2, 1, 0)$.

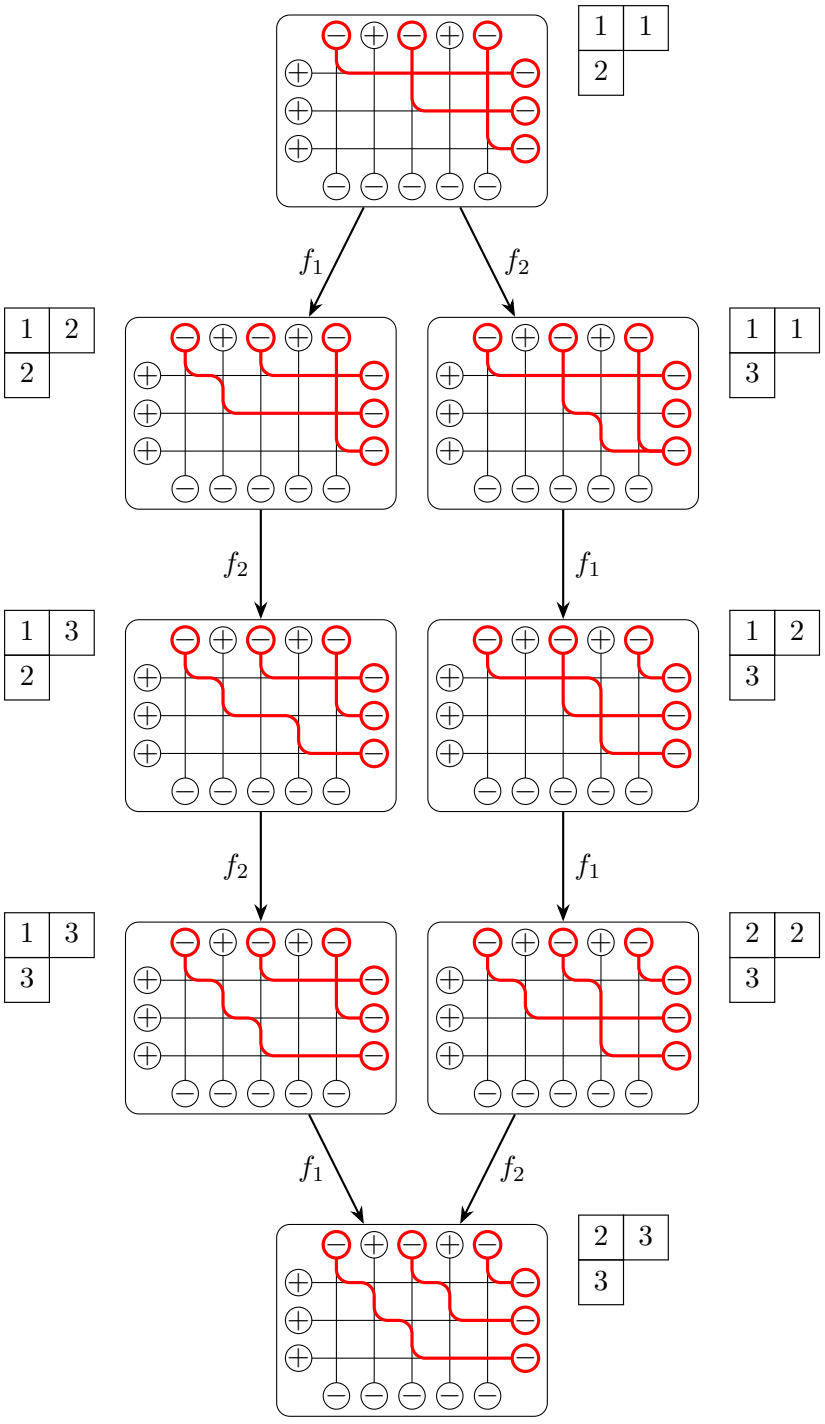


FIGURE 2. The crystal structure on $\mathfrak{S}_\lambda(\mathbf{z};0)$ when $\lambda=(2,1)$ and $n=3$.

Remark 4.30. We have chosen to consider the $q=0$ models here because this is the simplest case, and because for the $q=0$ models we may use the crystal $\mathcal{B}_\lambda$. However, it is also important to consider general q. In (4.15) we defined an embedding of the states of the $q=0$ Tokuyama model into the crystal $\mathcal{B}_\lambda$. If $q\neq 0$, this procedure requires modification,

because if the state contains $\mathtt{b}_1$ vertices, then GTP($\mathfrak{s}$) is not left strict, and we are not able to subtract the Gelfand–Tsetlin pattern P in Lemma 4.14. Therefore we modify the procedure as follows omit this subtraction, replacing (4.15) by

$$\theta(\mathfrak{s}) = \mathrm{Sch}(\mathrm{SSYT}(\mathrm{GPT}(\mathfrak{s}))). \tag{4.16}$$

Here SSYT(GPT($\mathfrak{s}$)) $\in \mathcal{B}_{\lambda+\rho}$, so Sch is the involution of $\mathcal{B}_{\lambda+\rho}$. The image of the map θ is no longer all of the crystal $\mathcal{B}_{\lambda+\rho}$ since there is a constraint on GPT($\mathfrak{s}$), which is now a *strict* Gelfand–Tsetlin pattern. The utility of this embedding, and the use of the Schützenberger involution is seen in Brubaker, Bump, and Friedberg 2011b. Note that $\mathcal{B}_{\lambda+\rho}$ can be embedded in $\mathcal{B}_\lambda \otimes \mathcal{B}_\rho$ and Kim and Lee 2011 proposed (in a slightly different context) that it is better to consider the image of $\theta(\mathfrak{s})$ in $\mathcal{B}_\lambda \otimes \mathcal{B}_\rho$.

8.3. Crystals for Delta ice. Let us consider Delta ice introduced in Section 7. We take the column parameter α_j to be zero. This makes $\mathtt{b}_1$ patterns for Delta ice zero; we will see momentarily that this allows the states to be embedded in $\mathcal{B}_\lambda$.

As with Gamma ice, the column numbers of the vertical edges that carry the lines can be arranged in a Gelfand–Tsetlin pattern. Let us denote this, as before, as GTP($\mathfrak{s}$). Now we need the analog of Lemma 4.13. This is the fact that for Delta ice, a necessary and sufficient condition for there to be no $\mathtt{b}_1$ patterns is that the corresponding Gelfand–Tsetlin pattern is *right strict*, so that $a_{i,j} > a_{i-1,j+1}$. This is proved like Lemma 4.13, remembering that for Delta ice the paths move down and to the left,

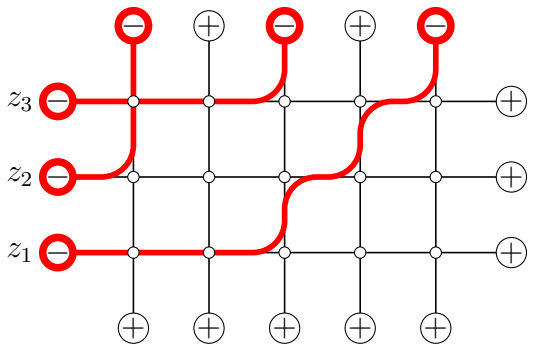
Now since GTP($\mathfrak{s}$) is right strict, with top row $\lambda + \rho$, we may subtract

$$Q = \begin{Bmatrix} n-1 & & n-2 & & \cdots & 0 \\ & n-1 & & \cdots & 1 & \\ & & \ddots & & \iddots & \\ & & & n-1 & & \end{Bmatrix} \tag{4.17}$$

and obtain another Gelfand–Tsetlin pattern, with top row λ. Let $\mathrm{GTP}_\circ(\mathfrak{s})$ denote the "reduced" pattern GTP($\mathfrak{s}$) $- Q$. We may now adapt the map θ in (4.15). In this case we omit the final step of the Schützenberger involution which is not needed for Delta ice. We write

$$\theta_\Delta(\mathfrak{s}) = \mathrm{SSYT}(\mathrm{GPT}_\circ(\mathfrak{s})). \tag{4.18}$$

Example 4.31. We cannot use the state in (4.14), since it contains a $\mathtt{b}_1$ pattern. Let us instead consider



(4.19)

We find that

$$\mathrm{GTP}(\mathfrak{s}) = \begin{pmatrix} 4 & & 2 & & 0 \\ & 4 & & 1 & \\ & & 2 & & \end{pmatrix}, \qquad \mathrm{GTP}(\mathfrak{s}) = \begin{pmatrix} 2 & & 1 & & 0 \\ & 2 & & 0 & \\ & & 0 & & \end{pmatrix},$$

so $\theta_\Delta(\mathfrak{s}) = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 3 \\ \cline{1-1}\end{array}$.

The map θ_Δ is a bijection $\mathfrak{S}_\lambda^\Delta(\mathbf{z};0) \to \mathcal{B}_\lambda$. One may check that

$$\beta(\mathfrak{s}) = \mathbf{z}^\rho \mathbf{z}^{\mathrm{wt}(\theta(\mathfrak{s}))}.$$

Exercises

Exercise 4.1. Compute the partition function for the six-vertex model with the weights as in Section 2 for the partition $(n-1, n-2, \dots, 1, 0)$ and show that it is a q-deformation of the Vandermonde determinant.

The next exercise generalizes Theorem 4.2. It shows that the parameter q can be replaced by a sequence of parameters t_i, so that the Tokuyama model then has two row parameters, z_i and t_i.

Exercise 4.2 (Brubaker, Bump, and Friedberg 2011a)**.** For each $1 \leqslant i \leqslant n$ let z_i and t_i be nonzero complex numbers. In place of the Boltzmann weights (4.5), use the following weights,

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
1	z_i	t_i	z_i	$z_i(1+t_i)$	1

and in place of the R-matrix, use the following weights.

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
$z_j + t_j z_i$	$z_i + t_i z_j$	$t_i z_j - t_j z_i$	$z_i - z_j$	$(1+t_i)z_i$	$(1+t_j)z_j$

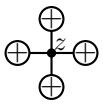
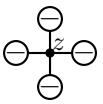
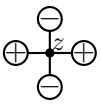
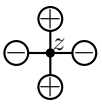
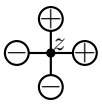
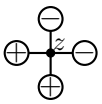
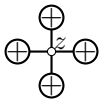
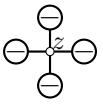
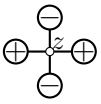
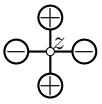
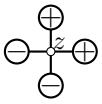
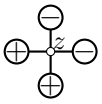
Generalize Theorem 4.2 to prove the Yang–Baxter equation in this generality. Thus we recover Theorem 4.2 by specializing all t_i to $-q$.

Exercise 4.3 (Brubaker, Bump, and Friedberg 2011a)**.** Continuing from the previous exercise, generalize Theorem 4.3 by showing that if we use the Boltzmann weights depending on z_i and t_i from the previous exercise in the i-th row of the grid, the partition function equals

$$\prod_{i<j} (z_i + t_i z_j) s_\lambda(\mathbf{z}).$$

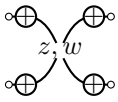
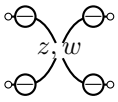
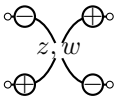
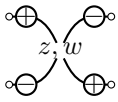
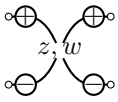
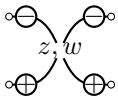
For the next exercises we want to have two types of free-fermionic vertices, which we will call Gamma and Delta. Gamma ice is the Tokuyama ice with Boltzmann weights (4.5), and Delta ice is new to these exercises. They can be used together, and we will distinguish their Boltzmann weights by using a black dot for Gamma ice, and a white dot for Delta ice as in the following table.

(4.20)

	$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
Gamma						
	1	z	$-q$	z	$z(1-q)$	1
Delta						
	1	$-qz$	1	z	$(1-q)z$	1

We will also make use of the following R-matrix:

(4.21)

$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
					
$z-qw$	$w-qz$	$q(z-w)$	$z-w$	$(1-q)z$	$(1-q)w$

Exercise 4.4 (Brubaker, Bump, and Friedberg 2011a)**.** Prove a Yang–Baxter equation using the Delta weights in (4.20) and the R-matrix (4.6).

Exercise 4.5 (Brubaker, Bump, and Friedberg 2011a)**.** Now consider the lattice model with extended wall boundary conditions, using Delta ice. That is, we use the weight (4.20) with instead of (4.5). Evaluate these partition functions.

Exercise 4.6. Complete the proof of Theorem 4.19 by imitating the proof of Theorem 4.3.
Hint: You will need to know that the R-matrix for the factorial models is the same as the R-matrix (4.6). Try to deduce this fact without further calculation from the discussion in this section.

Exercise 4.7. Theorem 2.17 of Chapter 2 implies that the factorial Schur functions satisfy a column Yang–Baxter equation for some R-matrix. Make this explicit.

CHAPTER 5

Modifying R-matrices

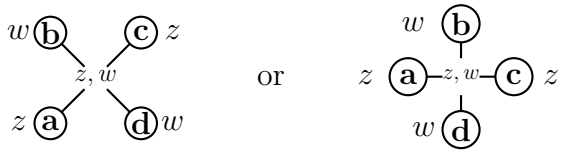
In this chapter, we consider two methods of modifying solutions of the Yang–Baxter equation to get new solutions. We will consider colored models, in which each spin is assigned an attribute that we call a *color*, such as the colored models in Chapters 7 and 8. Borodin and Wheeler Borodin and Wheeler 2022, motivated by probability theory, did much to stimulate the current interest in colored models.

1. Jimbo's R-matrix

In order to have an example in mind, let us use the R-matrix introduced by Jimbo 1986, coming from the quantum group $U_q(\widehat{\mathfrak{sl}}_r)$. (Jimbo also gave R-matrices for standard modules for the other classical Cartan types.) Let us choose r colors. The edge types depend on a complex number z, which we call the *spectral parameter*. For every edge type, the spinsets are the same, a set $\mathcal{C} = \{c_1, \dots, c_r\}$ of *colors*. The colors are ordered so that $c_1 > \cdots > c_r$. The set $\mathcal{C}$ is called the *palette*.

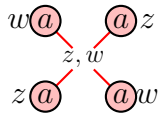
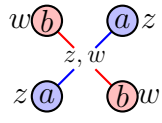
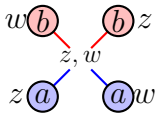
Let z and w be two spectral parameters. We will consider a vertex with adjacent edges labeled $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ as follows:

(5.1)



Thus z is the vertex type at the vertices labeled $\mathbf{a}$ and $\mathbf{c}$, while w is the vertex type at the vertices labeled $\mathbf{b}$ and $\mathbf{c}$. The Boltzmann weights are as follows.

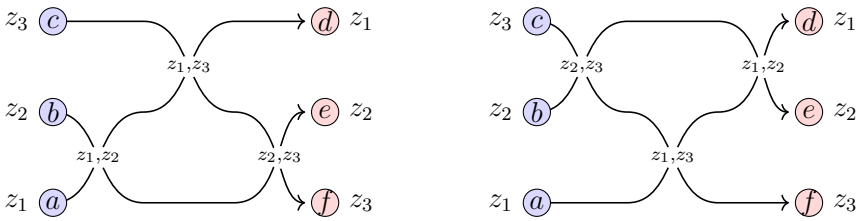
(5.2)

Jimbo's R-matrix		
Type A	Type B	Type C
		
$z - qw$	$q(z-w)$ if $a < b$, $z - w$ if $a > b$,	$(1-q)z$ if $a < b$, $(1-q)w$ if $a > b$,

Jimbo showed that this R-matrix is associated with the quantum group $U_q(\widehat{\mathfrak{gl}}_n)$. To be precise the R-matrix in Jimbo 1986 is not exactly this R-matrix but differs from it by a Drinfeld twist, since it gives both Type B weights the value $\sqrt{q}(z-w)$. But for reference we describe this as *Jimbo's R-matrix*.

We have classified the configurations of spins at a vertex as Type A, Type B or Type C. This is analogous to the classification of the configurations of the six-vertex model as $\mathtt{a}_1$, $\mathtt{a}_2$ (Type A), $\mathtt{b}_1$, $\mathtt{b}_2$ (Type B) and $\mathtt{c}_1$, $\mathtt{c}_2$ (Type C).

Theorem 5.1 (Jimbo 1986)**.** *Jimbo's R-matrix satisfies a parametrized Yang–Baxter equation. If a, b, c, d, e and f are any colors, then the partition functions of the following two systems are equal:*



Proof. Note that the R-matrix only depends on the relative order of the colors a and b in the table (5.2). We observe that it is sufficient to prove this when the number of colors is $\leqslant 3$. Indeed, the colors d, e, f must be a permutation of a, b, c, or else the system has no states. This reduces the number of cases to a reasonably small finite number, which can be checked using a computer or even by hand. In fact, with at most 3 colors and with d, e and f a permutation of a, b, c there are at most $6 \cdot 27 = 162$ possible boundary conditions. □

Theorem 5.1 can be written compactly as follows. Let V be the free vector space on the set $\mathcal{C}$ of colors, and let $R(z, w) : V \otimes V \to V \otimes V$ be the linear R-matrix associated with this linear transformation as in Section 3 in Chapter 2. Then in terms of the Yang–Baxter commutator:

$$[\![R(z_1, z_2), R(z_1, z_3), R(z_2, z_3)]\!] = 0.$$

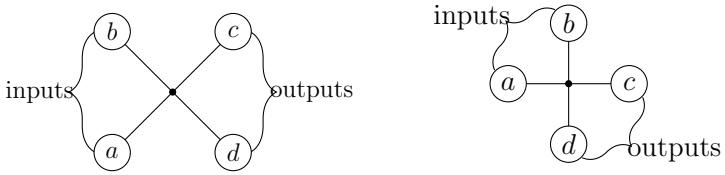
As a notational point, we may denote the vertex (here labeled z, w) alternatively as $\mathbf{v}$, and the Boltzmann weight as

$$\begin{cases} \beta_{\mathbf{v}}(a, a, a, a) & \text{in Type A,} \\ \beta_{\mathbf{v}}(a, b, a, b) & \text{in Type B,} \\ \beta_{\mathbf{v}}(a, b, b, a) & \text{in Type C.} \end{cases}$$

Here a and b are distinct colors, indicated by the colors blue and red, respectively.

1.1. Conservation of Color. In this Chapter, we will explain two methods of modifying Yang–Baxter equations to obtain other Yang–Baxter equations, using Jimbo's R-matrix as an example. Subject to a requirement that we will explain next, the methods apply to colored models, requiring minor modifications for other vertex types, such as for the T-vertices in the bosonic models from Chapter 8.

What is required is a *conservation of color*. That is, let us classify two edges as inputs and two as outputs, thus:



Then the conservation of color requires that the number of inputs of a given color equals the number of outputs. Note that this is true for the Jimbo R-matrix. If we allow bosonic vertices, then these numbers must be counted with multiplicity. If the Boltzmann weights satisfy this conservation of color we will call them *viatic*. We may weaken this condition

slightly to allow $a = b$ and $c = d$, in which case we say the Boltzmann weights are *planoviatic*. Thus the six-vertex model is viatic, while the eight vertex model is planoviatic but not viatic.

If the Boltzmann weights are viatic then the edges of a single color can be organized into paths, and if a path enters a vertex through an input vertex, then it must exit through an output vertex, continuing through the structure, always moving down and to the right. For planoviatic weights, paths also form but they are not forced to move in any particular direction, so they may wander around or form loops.

2. Change of basis

This procedure modifies the Boltzmann weights for vertices of Type C. We may turn the Yang–Baxter equation into a vector Yang–Baxter equation as in Section 3 in Chapter 2. Let $\mathbf{v}$ denote a vertex such as the vertex z, w in (5.1), and let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be the adjacent edges, labeled as in (5.1).

The vector space $V_{\mathbf{e}}$ associated with an edge type $\mathbf{e}$ is then the free vector space on the spinset $\Sigma_{\mathbf{e}}$. Thus with vertices labeled as in (5.1), the Boltzmann weight becomes the matrix, with respect to the bases $\Sigma_{\mathbf{a}} \otimes \Sigma_{\mathbf{b}}$ and $\Sigma_{\mathbf{c}} \otimes \Sigma_{\mathbf{d}}$ of $V_{\mathbf{a}} \otimes V_{\mathbf{b}}$ and $V_{\mathbf{c}} \otimes V_{\mathbf{d}}$ of a linear map $V_{\mathbf{a}} \otimes V_{\mathbf{b}} \longrightarrow V_{\mathbf{c}} \otimes V_{\mathbf{d}}$. Now if we choose another basis $\Sigma'_{\mathbf{e}}$ of $V_{\mathbf{e}}$, we may refer this linear map to the bases $\Sigma'_{\mathbf{a}} \otimes \Sigma'_{\mathbf{b}}$ and $\Sigma'_{\mathbf{c}} \otimes \Sigma'_{\mathbf{d}}$ and obtain a new set of Boltzmann weights for a new Yang–Baxter equation.

A simple way of implementing this is to choose for each vertex map a function $f_{\mathbf{e}} : \Sigma_{\mathbf{e}} \longrightarrow \mathbb{C}^\times$, and to choose $\Sigma'_{\mathbf{e}} = \{x' | x \in \Sigma_{\mathbf{e}}\}$, where we define

$$x' = \frac{1}{f(x)} x.$$

Let us use $\beta_{\mathbf{v}}$ for the original Boltzmann weights, and $\beta'_{\mathbf{v}}$ for the modified ones. Let $a \in \Sigma_{\mathbf{a}}$, $b \in \Sigma_{\mathbf{b}}$, $c \in \Sigma_{\mathbf{c}}$ and $d \in \Sigma_{\mathbf{d}}$. Since the sets $\Sigma_{\mathbf{e}}$ and $\Sigma'_{\mathbf{e}}$ are in bijection, we write $\beta'_{\mathbf{v}}(a, b, c, d)$ instead of $\beta'_{\mathbf{v}}(a', b', c', d')$. Then

$$\beta'_{\mathbf{v}}(a, b, c, d) = \frac{f_{\mathbf{a}}(a) f_{\mathbf{b}}(b)}{f_{\mathbf{c}}(c) f_{\mathbf{d}}(d)} \beta_{\mathbf{v}}(a, b, c, d) = \frac{f_{\mathbf{a}}(a) f_{\mathbf{b}}(b)}{f_{\mathbf{a}}(c) f_{\mathbf{b}}(d)} \beta_{\mathbf{v}}(a, b, c, d) \tag{5.3}$$

because the vertex types $\mathbf{a}$ and $\mathbf{c}$ are the same, as are the vertex types $\mathbf{b}$ and $\mathbf{d}$, so $f_{\mathbf{a}} = f_{\mathbf{c}}$ and $f_{\mathbf{b}} = f_{\mathbf{d}}$. In Types A and B, $a = c$ and $b = d$, so

$$\beta'_{\mathbf{v}}(a, b, c, d) = \beta_{\mathbf{v}}(a, b, c, d)$$

in Types A and B. On the other hand, let us denote $f_{\mathbf{v}} = f_{\mathbf{a}}/f_{\mathbf{b}}$. In Type C we have $a = d$ and $b = c$, so

$$\beta'(a, b, c, d) = f_{\mathbf{v}}(a) f_{\mathbf{v}}(b) \beta(a, b, c, d)$$

in this case.

3. Drinfeld twisting

Drinfeld twisting refers to an operation on Hopf algebras that modifies the comultiplication. If the Hopf algebra is quasitriangular, then it is a source of solutions to the Yang–Baxter equation, and twisting also modifies the R-matrices. In a simple case, we may describe this directly. We will will consider colored systems such as the Jimbo R-matrix. The scheme we describe will modify the Boltzmann weights of Type B, leaving the weights of Type A and Type C unchanged.

If $\mathcal{C}$ is the set of colors, we require a function $\phi : \mathcal{C} \times \mathcal{C} \longrightarrow \mathbb{C}^\times$. We require

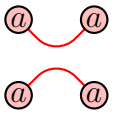
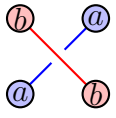
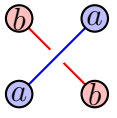
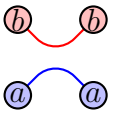
$$\phi(a,b)\phi(b,a) = 1.$$

Now for Type C, we multiply the Boltzmann weight $\beta(a,b,b,a)$ by $\phi(a,b)$. Thus we obtain new Boltzmann weights as follows. With a, b distinct colors:

$$\begin{array}{ll} \beta'_{\mathbf{v}}(a,a,a,a) = \beta_{\mathbf{v}}(a,a,a,a) & \text{Type A} \\ \beta'_{\mathbf{v}}(a,b,a,b) = \phi(a,b)\beta_{\mathbf{v}}(a,b,a,b) & \text{Type B} \\ \beta'_{\mathbf{v}}(a,b,b,a) = \beta_{\mathbf{v}}(a,b,b,a) & \text{Type C} \end{array}$$

Proposition 5.2. *If the weights β satisfy the Yang–Baxter equation, so do the weights β'.*

Proof. Recall that the set $\mathcal{C}$ of colors is assigned an order. By a *colored braid* we mean an Artin braid in which each strand is decorated by a color in $\mathcal{C}$. Let us turn every state $\mathfrak{s}$ of the system (either the left-hand side or the right-hand side) into a "colored braid" of three strands by the following scheme

Type A	Type B_1	Type B_2	Type C
			
	$a < b$	$a > b$	

Thus we have divided Type B into two cases depending on the order of the color, but the larger color always goes over the smaller color. Strands of the same color are not allowed to cross. Let a_1, a_2 and a_3 be the colors on the left (reading from bottom to top) and let b_1, b_2 and b_3 be the colors on the right. In order for there to exist any states, the b_i must be a permutation of the a_i. Thus $b_i = a_{w(i)}$ for a permutation of $1, 2, 3$, and the permutation w is uniquely determined by the requirement that the color a_i is connected to the color b_i by a strand. This is because strands of the same color are not allowed to cross, which implies that w is the smallest permutation (in the Bruhat order) such that $b_i = a_{w(i)}$. Now if $i < j$ and $w(i) > w(j)$ then we say that (i,j) is a descent of w. Since strands of the same color cannot cross, $a_i \neq a_j$ if (i,j) is a descent.

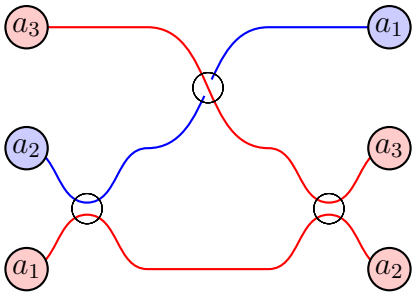
Since alterations in the Boltzmann weights occur when strands cross, we obtain

$$\frac{\beta'(\mathfrak{s})}{\beta(\mathfrak{s})} = \prod_{\text{descents } (i,j)} \phi(a_i, a_j).$$

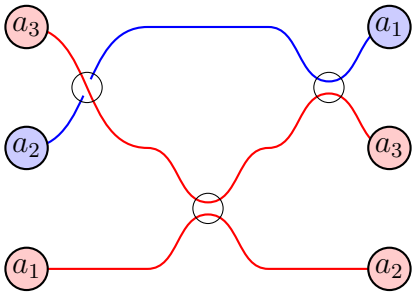
This formula is the same on either the right-hand side or the left-hand side of the Yang–Baxter equation, and so the Yang–Baxter equation for the β weights implies the Yang–Baxter equation for the β' weights.

If (i,j) is not a descent, it is possible for the a_i strand to cross the a_j strand twice, but if this happens, the crossings are in opposite directions and since $\phi(a_i,a_j)\phi(a_j,a_i) = 1$, there is no unwanted contribution.

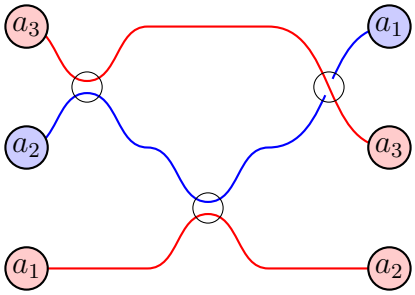
A couple of examples should clarify this proof. In this example, $a_1 = a_3 > a_2$ and w is the permutation $(2,3)$. On the left-hand side there is one state:



We have drawn the vertices as circles for clarity. On the right-hand side there are two states:

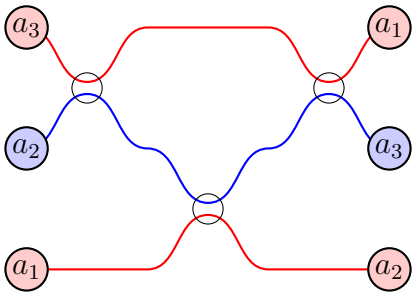


and:

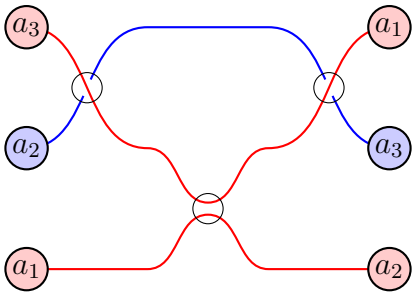


For every one of these states, the blue lines and the red lines cross, so there is a contribution of $\phi(a_2, a_3)$. This is the same for every state on both sides, so the Yang–Baxter equation is preserved.

We give one more example to show that two colored lines may cross twice, with the contributions cancelling. As before we take $a_1 = a_3$, but now we take $b_i = a_i$ for all i, so that the permutation $w = 1$. On the right-hand side, there are two states. In one, there are no crossings:



In the other, there are two crossings, in opposite directions:



For this one, there are two crossings, in opposite senses, and the contributions cancel. □

4. The spherical and antispherical R-matrices

Let us expand Jimbo's R-matrix by adding an additional element $\oplus$ to the spinset. Thus the spinset of any edge consists of $\oplus$ and the r colors $\{c_1, \dots, c_r\}$. The R-matrix is:

Spherical R-matrix			
$w \oplus \ \oplus z$ $z \oplus \ \oplus w$	$w\ c\ \ c\ z$ $z\ c\ \ c\ w$	$w \oplus \ \oplus z$ $z\ c\ \ c\ w$	$w\ d\ \ d\ z$ $z\ c\ \ c\ w$
$z - qw$	$z - qw$	$(1-q)z$	$(1-q)z$ if $c < d$ $(1-q)w$ if $c > d$
$w\ c\ \ c\ z$ $z \oplus \ \oplus w$	$w\ c \ \oplus z$ $z \oplus \ c\ w$	$w \oplus \ c\ z$ $z\ c\ \ \oplus w$	$w\ d\ \ c\ z$ $z\ c\ \ d\ w$
$(1-q)w$	$q(z-w)$	$z - w$	$z - w$ if $c > d$ $q(z-w)$ if $c < d$

We have titled this the *spherical R-matrix* for future reference. The name is justified because of its relationship, to be explained later, to Demazure–Lusztig operators which describe the so-called spherical representation of the affine Hecke algebra.

The spherical R-matrix satisfies the following Yang–Baxter equation. Let V be the free vector space on the expanded set of colors $\mathcal{C} \cup \{\oplus\}$. Then as in Section 3, the spherical R-matrix may be encoded in a linear transformation $R_\circ(z, w; q) = R_\circ(z, w) : V \otimes V \to V \otimes V$.

Theorem 5.3. *If z_1, z_2, z_3 are any spectral parameters, then*

$$[\![R_\circ(z_1, z_2), R_\circ(z_1, z_3), R_\circ(z_2, z_3)]\!] = 0.$$

We will show by two different methods how this Yang–Baxter equation for the spherical R-matrix may be deduced from Theorem 5.1. Both proofs start with Jimbo's R-matrix with the expanded palette $\mathcal{C} \cup \{\oplus\}$. So $\oplus$ is considered to be one of the colors.

First proof.. In the expanded palette $\mathcal{C} \cup \{\oplus\}$, let us take $\oplus$ to be the *smallest* color, so $\oplus < c_r < \cdots < c_1$. This gives the right weights for the spherical R-matrix, except for two that are incorrect:

$\oplus \ \oplus$ $c\ \ c$	$c\ \ c$ $\oplus \ \oplus$
$(1-q)w$	$(1-q)z$

To fix these, we have to switch these Boltzmann weights. We recall that in Jimbo's R-matrix, the edge types depend on a spectral parameter z. Thus we will denote the edge type $\mathbf{e}(z)$. Thus in the notation of Section 2, we have

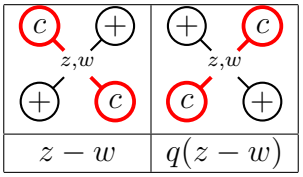
$$\mathbf{a} = \mathbf{c} = \mathbf{e}(z), \qquad \mathbf{b} = \mathbf{d} = \mathbf{e}(w).$$

We can take

$$f_{(z)}(a) = \begin{cases} z & \text{if } a \in \mathcal{C}, \\ 1 & \text{if } a = \oplus. \end{cases}$$

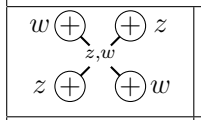
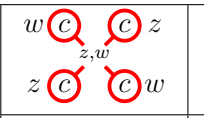
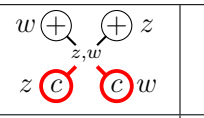
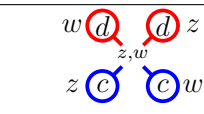
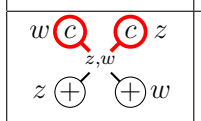
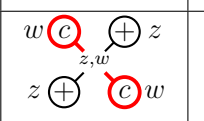
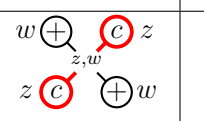
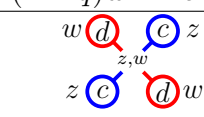
Using (5.3), this has the effect of interchanging these two cases. □

Second proof.. An alternative approach is to take $\oplus$ to be the *largest* color, so $c_1 < \cdots < c_r < \oplus$. Again we get the spherical weights, except that two are wrong. But now they are a different pair:



We can use the method of Drinfeld twisting in Section 3 to move the q to where it should be in this case of the crossing of a $\oplus$ line with a colored line. □

Complementary to the spherical R-matrices are another family that we call *anti-spherical*. Here are the Boltzmann weights:

Antispherical R-matrix			
			
$w - qz$	$z - qw$	$(1-q)z$	$(1-q)z$ if $c < d$ $(1-q)w$ if $c > d$
			
$(1-q)w$	$q(z-w)$	$z - w$	$z - w$ if $c > d$ $q(z-w)$ if $c < d$

Note that this differs from the spherical R-matrix only in the first weight, where all adjacent spins are $\oplus$. It also satisfies the same Yang–Baxter equation. However the Yang–Baxter equation for this R-matrix cannot be deduced from Jimbo's by change of basis and Drinfeld twisting. This is an instance of a *Perk-Schultze equation* Perk and Schultz 1981 equation. Whereas the relevant quantum group for the Spherical R-matrix is $U_q(\widehat{\mathfrak{gl}}(r+1))$, the relevant quantum group for the antispherical R-matrix is the superalgebra $U_q(\widehat{\mathfrak{gl}}(r|1))$.

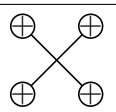
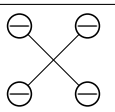
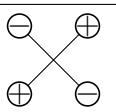
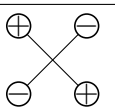
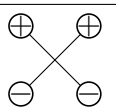
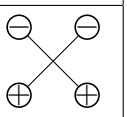
Theorem 5.4. *Let $R_\bullet(z, w; q) = R_\bullet(z, w)$ denote the vector form of the antispherical R-matrix. If z_1, z_2, z_3 are any spectral parameters, then*

$$[\![R_\bullet(z_1, z_2), R_\bullet(z_1, z_3), R_\bullet(z_2, z_3)]\!] = 0.$$

Proof. Although this cannot be deduced from Jimbo's Yang–Baxter equation, it can be proved the same way with a finite amount of computation. That is, since in order for the system to have any solutions, the spins on the right boundary edges must be a permutation of the spins on the left boundary edges, at most 3 colors (not counting $\oplus$) can appear among the boundary edges. There are therefore at most $6 \times 4^3 = 384$ cases to be checked, and this can easily be done using a computer. □

5. Relation with the field-free six-vertex model

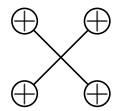
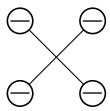
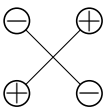
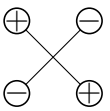
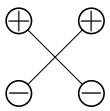
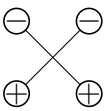
The parametrized Yang–Baxter equation in Theorem 3.10 can also be related to Jimbo's R-matrix. We multiply the parametrized weights by 2, without affecting the Yang–Baxter equation. The Boltzmann weights for this are:

a_1	a_2	b_1	b_2	c_1	c_2
					
$zq-(zq)^{-1}$	$zq-(zq)^{-1}$	$z-z^{-1}$	$z-z^{-1}$	$q-q^{-1}$	$q-q^{-1}$

If $R(z)$ is the corresponding vector R-matrix, then the parametrize Yang–Baxter equation can be written

$$[\![R(z), R(zw), R(w)]\!] = 0. \tag{5.4}$$

It will be convenient to modify the parameters z,w and q in Jimbo's R-matrix, replacing $z \to z^2, w \to w^2$ and $q \to q^{-2}$. We choose the palette to be $\{\oplus, \ominus\}$ with $\ominus < \oplus$. Then the R-matrix is:

a_1	a_2	b_1	b_2	c_1	c_2
					
$z^2-q^{-2}w^2$	$z^2-q^{-2}w^2$	z^2-w^2	$q^{-2}(z^2-w^2)$	$(1-q^{-2})z^2$	$(1-q^{-2})w^2$

Now we multiply every weight by the same factor $qz^{-1}w^{-1}$. Then the Boltzmann weights only depend on the ratio $\zeta := z/w$. They are:

a_1	a_2	b_1	b_2	c_1	c_2
$q\zeta-(q\zeta)^{-1}$	$q\zeta-(q\zeta)^{-1}$	$q(\zeta-\zeta)^{-1})$	$q^{-1}(\zeta-\zeta^{-1})$	$(q-q^{-1})\zeta$	$(q-q^{-1})\zeta^{-1}$

We can write the corresponding R-matrix as $R'(\zeta) = R'(z/w)$. The Yang–Baxter equation has the form:

$$[\![R'(z_1/z_2), R'(z_1/z_3), R'(z_2/z_3)]\!] = 0.$$

This is to be compared with (5.4). Since $(z_1/z_2)(z_2/z_3) = (z_1/z_3)$ it has the right form. The R-matrix $R'(\zeta)$ may be transformed into $R(\zeta)$ with a Drinfeld twist and a change of basis.

Exercises

Exercise 5.1. Check the 384 cases in the proof of Theorem 5.4 via computer.

Exercise 5.2. Check the claim that R-matrix $R'(\zeta)$ can be transformed into $R(\zeta)$ with a Drinfeld twist and a change of basis.

CHAPTER 6

Demazure Operators

In this chapter, we continue the study of colored models and will see many examples of the following phenomenon. We consider models whose boundary conditions depend on a permutation w of a set of colors, describing their locations on the boundary. Thus the partition functions form a family, indexed by w. If $w = 1$, one finds that the system is *monostatic*, meaning that it has only one state, and the partition function is completely known. Supplementing this information, and coming from the Yang–Baxter equation, there are recursion relations that allow one to deduce the partition function for more general w. These recursion formulas are in terms of certain *Demazure operators* that generate a *Hecke algebra*. In this chapter, we introduce these algebraic topics.

To quickly illustrate how Yang–Baxter equations lead to Demazure recursions for partition functions, we will consider a special case in this chapter, where the R-matrix and the T-vertices are all drawn from a single homogeneous parametrized Yang–Baxter equation.

1. Some Lie theory

If G is a complex reductive Lie group, we may associate with G a *Weyl group* W, a *maximal torus* T, a *root system* Φ and a *weight lattice* Λ. Elements of Λ are called *weights*. The weight lattice Λ is the group of rational characters of T, and if $\mathbf{z} \in T$, $\lambda \in \Lambda$, we will denote by $\mathbf{z}^\lambda$ the application of λ to $\mathbf{z}$. The root system Φ is a finite subset of Λ and the group $W = N(T)/T$ acts on T by conjugation, hence on Λ, preserving the root system Φ. Let $\Lambda_\mathbb{R} = \mathbb{R} \otimes \Lambda$ be the ambient real vector space of Λ. It may be endowed with a W-invariant inner product.

Example 6.1. In the majority of our examples, G will be the general linear group, and the reader may specialize to this case if they want, with little loss of continuity. So let us define G, W, T, Φ, Λ when $G = \mathrm{GL}(n, \mathbb{C})$. In this case, the normalizer $N(T)$ is the group of monomial matrices, and the Weyl group $W = N(T)/T$ is the symmetric group S_n. The torus T is $(\mathbb{C}^\times)^n$, which we embed in G via

$$\mathbf{z} = (z_1, \ldots, z_n) \longmapsto \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix}.$$

The *weight lattice* Λ can be identified with $\mathbb{Z}^n$. Then if $\mathbf{z} \in T$ and $\mu = (\mu_1, \ldots, \mu_n) \in \Lambda$, we denote $\mathbf{z}^\mu = z_1^{\mu_1} \cdots z_n^{\mu_n}$. The W-invariant inner product on $\Lambda_\mathbb{R}$ is then just the usual dot product on $\mathbb{R}^n = \Lambda_\mathbb{R}$. Let $\mathbf{e}_i$ be the standard basis vectors in $\mathbb{Z}^n$. Then the root system Φ consists of vectors $\mathbf{e}_i - \mathbf{e}_j$ with $1 \leqslant i, j \leqslant n$ and $i \neq j$.

As usual in Lie theory, the root system Φ can be partitioned into two parts $\Phi = \Phi^+ \cup \Phi^-$ called *positive* and *negative* roots, respectively. Let $\Sigma = \{\alpha_1, \ldots, \alpha_\ell\}$ be the set of *simple roots*, which are the positive roots that cannot be decomposed into other positive roots.

Then W is generated by the simple reflections s_i corresponding to the α_i. For $\mathrm{GL}(n)$, the positive roots are $\mathbf{e}_i - \mathbf{e}_j$ with $1 \leqslant i < j \leqslant n$, and the simple roots are $\alpha_i = \mathbf{e}_i - \mathbf{e}_{i+1}$.

Let Γ be a group with a set $I = \{s_1, \ldots, s_\ell\}$ of generators. We assume the *quadratic relations*

$$s_i^2 = 1. \tag{6.1}$$

Since the s_i have order 2, for every pair s_i, s_j of distinct elements such that the order $n_{i,j}$ of $s_i s_j$ is finite, we also have the *braid relations*

$$s_i s_j s_i \cdots = s_j s_i s_j \cdots \tag{6.2}$$

with $n_{i,j}$ factors on both sides. For example, if $G = \mathrm{GL}(n)$ the braid relations are $s_i s_j s_i = s_j s_i s_j$ if $j = i \pm 1$, and $s_i s_j = s_j s_i$ if $|i - j| > 1$.

With these assumptions, the group Γ is called a *Coxeter group* if the quadratic and braid relations give a presentation of Γ. Concretely, this means that if Ξ is any group containing elements σ_i that also satisfy the quadratic and braid relations, then there is a unique homomorphism $\phi : \Gamma \longrightarrow \Xi$ such that $\phi(s_i) = \sigma_i$.

Theorem 6.2. *The Weyl group W is a Coxeter group.*

Proof. The statement is true for the Weyl group of any Lie group, though we are specializing to the case of the symmetric group. See Bump 2013, Theorem 25.1 or Humphreys 1990, Theorem 19.1. □

Lemma 6.3. *The reflection s_i sends α_i to its negative, and permutes other positive roots. In other words s_i maps $\Phi^+ - \{\alpha_i\}$ to itself.*

Proof. See Bump 2013, Proposition 20.1 (ii) or Bourbaki 2002, Section IV.1.6 Corollary 1. □

Definition 6.4. A Weyl vector is a vector $\rho \in \Lambda$ such that $\rho - s_i(\rho) = \alpha_i$ for simple roots α_i and corresponding simple reflections s_i.

Example 6.5. The weight $\frac{1}{2}\sum_{\Phi^+} \alpha$ is a Weyl vector, as follows from Lemma 6.3. If G is semisimple, this is the unique Weyl vector.

If G is not semisimple, there may be some freedom in choosing the Weyl vector.

Example 6.6. If $G = \mathrm{GL}(n)$, so $W = S_n$ and $\Lambda = \mathbb{Z}^n$ then a useful choice for the Weyl vector is

$$\rho = (n-1, n-2, \ldots, 0). \tag{6.3}$$

Defining ρ to be half the sum of the positive roots would give $\left(\frac{n-1}{2}, \frac{n-3}{2}, \ldots, \frac{1-n}{2}\right)$, and if n is even, this vector has denominators that we can avoid by the choice (6.3). Moreover this choice works well for applications to lattice models.

A factorization $w = s_{i_1} \cdots s_{i_k}$ into simple reflections of shortest possible length k is called a *reduced expression*. The Weyl group has a *length function* $\ell : W \longrightarrow \mathbb{N} = \{0, 1, 2, 3, \ldots\}$. Two possible definitions can be given which are equivalent.

Definition 6.7. The length $\ell(w)$ is the length k of a reduced expression $w = s_{i_1} \cdots s_{i_k}$. Alternatively, $\ell(w)$ is the cardinality of the set

$$\{\alpha \in \Phi^+ \mid w(\alpha) \in \Phi^-\}.$$

Proposition 6.8. *The two definitions of $\ell(w)$ are equivalent. If $w = s_{i_1} \cdots s_{i_k}$ is a reduced expression, then*

$$\{\alpha \in \Phi^+ \mid w(\alpha) \in \Phi^-\} = \{\alpha_{i_k}, s_{i_k}(\alpha_{i_{k-1}}), s_{i_k} s_{i_{k-1}}(\alpha_{i_{k-2}}), \ldots\}.$$

Proof. See Bump 2013, Propositions 20.5 and 20.10 or Bourbaki 2002, Section IV.1.6 Corollary 2. □

Let Λ^* be the dual lattice $\mathrm{Hom}_{\mathbb{Z}}(\Lambda, \mathbb{Z})$. Dual to the root system $\Phi \subset \Lambda$, there is a root system $\Phi^\vee$ in Λ^*, with a bijection $\alpha \mapsto \alpha^\vee$ between Φ and $\Phi^\vee$ such that the simple reflection s_i is the map

$$s_i(\lambda) = \lambda - \langle \lambda, \alpha_i^\vee \rangle \alpha_i.$$

Elements of $\Phi^\vee$ are called *coroots*. If we use the W-invariant inner product to identify $\Lambda_{\mathbb{R}}$ with its dual space, then Λ and Λ^* are both subgroups of $\Lambda_{\mathbb{R}}$ and if we make this identification, we have $\alpha^\vee = \frac{2}{\langle \alpha, \alpha \rangle} \alpha$.

A weight $\lambda \in \Lambda$ is called *dominant* if $\langle \alpha_i^\vee, \lambda \rangle$ is a nonnegative integer. Also, we introduce a partial order on $\Lambda_{\mathbb{R}}$ in which $\mu \preccurlyeq \lambda$ if $\lambda - \mu$ is a sum of positive roots, if we may write $\lambda - \mu = \sum n_i \alpha_i$ where n_i are nonnegative integers. Let (π, V) be an irreducible representation of G, and let χ_π be its character. Restricting χ_π to T, we may write

$$\chi_\pi(\mathbf{z}) = \sum_{\mu \in \Lambda} m_\pi(\mu)\, \mathbf{z}^\mu,$$

and the weight multiplicity function m_π is W-invariant. The μ such that $m_\pi(\mu) > 0$ are called the *weights* of the representation. There is a unique weight λ that is maximal with respect to $\preccurlyeq$, called the *highest weight* of the representation. It is a dominant weight. Conversely, given any dominant weight there is a unique irreducible representation with highest weight λ, so $\pi \leftrightarrow \lambda$ is a bijection between irreducibles and dominant weights.

The function $w \mapsto (-1)^{\ell(w)}$ is a character of W. Denoting $\chi_\pi = \chi_\lambda$ we have the *Weyl character formula*

$$\chi_\lambda(\mathbf{z}) = \prod_{\alpha \in \Phi^+} (1 - \mathbf{z}^{-\alpha})^{-1} \sum_{w \in W} (-1)^{\ell(w)} \mathbf{z}^{w(\lambda+\rho)-\rho}. \tag{6.4}$$

The *Weyl denominator formula* is the identity

$$\mathbf{z}^\rho \prod_{\alpha \in \Phi^+} (1 - \mathbf{z}^{-\alpha}) = \sum_{w \in W} (-1)^{\ell(w)} \mathbf{z}^{w(\rho)}. \tag{6.5}$$

Both (6.4) and (6.5) are true for any choice of Weyl vector ρ. If $G = \mathrm{GL}(n, \mathbb{C})$, then (6.4) is equivalent to (4.4.2).

1.1. Matsumoto's theorem. An expression $w = s_{i_1} \cdots s_{i_k}$ with $k = \ell(w)$ is called *reduced*. There may be many reduced expressions for w. For example if $W = S_4$ and $w_0 = (1,4)(2,3)$ is the longest element, there are 16 reduced expressions for w_0.

Matsumoto's theorem was found independently by Matsumoto 1964 and by Tits in the 1960's. This extremely useful fact says that if $\ell(w) = k$ and if

$$w = s_{i_1} \cdots s_{i_k} = s_{j_1} \cdots s_{j_k}$$

are two reduced expressions, then the equivalence of the two expressions can be proved using only the braid relations (6.2) and not the quadratic relations (6.1). To give an example, if $W = S_4$ then $w_0 = s_1s_2s_1s_3s_2s_1$ and $w_0 = s_3s_2s_3s_1s_2s_3$ are two reduced expression. The braid relations are:

$$s_1s_2s_1 = s_2s_1s_2, \quad s_2s_3s_2 = s_3s_2s_3, \qquad s_1s_3 = s_3s_1.$$

Matsumoto's theorem asserts that we can prove $s_1s_2s_1s_3s_2s_1 = s_3s_2s_3s_1s_2s_3$ using the braid relations and not the quadratic relations. Let us write 121321 instead of $s_1s_2s_1s_3s_2s_1$. Using the braid relations:

$$121321 = 212321 = 213231 = 231213 = 232123 = 323123.$$

To formulate Matsumoto's theorem rigorously, we introduce the *braid group* $B(W)$ of a Coxeter group W. This is the group with generators u_i (in bijection with the s_i) that satisfy the braid relations but not the quadratic relations.

Theorem 6.9 (Matsumoto Matsumoto 1964)**.** *If $s_{i_1} \cdots s_{i_k}$ and $s_{j_1} \cdots s_{j_k}$ are reduced expressions for the same element of W, then the corresponding elements $u_{i_1} \cdots u_{i_k}$ and $u_{j_1} \cdots u_{j_k}$ are equal in the braid group $B(W)$.*

Proof. See Björner and Brenti 2005, Theorem 3.3.1. For a geometric proof assuming W is a Weyl group, see Bump 2013, Theorem 25.2. □

1.2. Hecke algebras. A simple application of Matsumoto's theorem is the construction of Hecke algebras, which are deformations of the group algebra of the Weyl group. We pick a parameter q which may be a complex number or an indeterminate. Now the ring has generators T_i subject to the same braid relations as the simple reflections s_i and the new quadratic relation:

$$T_i^2 = (q-1)T_i + q. \tag{6.6}$$

Let $\mathcal{H}_q = \mathcal{H}_q(W)$ be the algebra generated by the T_i modulo these relations. This is the *Iwahori Hecke algebra of W*.

Proposition 6.10. *If $w = s_{i_1} \cdots s_{i_k}$ is a reduced expression, define $T_w = T_{i_1} \cdots T_{i_k}$. Then T_w depends only on w, not the reduced expression. The T_w ($w \in W$) are a basis of $\mathcal{H}_q$.*

Proof. Let $w = s_{i_1} \cdots s_{i_k}$ and $w = s_{j_1} \cdots s_{j_k}$ be reduced expressions. We want to show $T_{i_1} \cdots T_{i_k} = T_{j_1} \cdots T_{j_k}$, and by Matsumoto's theorem, we are reduced to the case where the two reduced words differ by a single braid group operation. But the T_i satisfy the braid relation. This proves that T_w is well-defined. We omit the verification that they are a vector space basis of $\mathcal{H}_q$. □

1.3. The Bruhat order. Another application of Matsumoto's concerns the definition of the Bruhat order, an important partial order on Weyl group elements. Let $u, v \in W$. Choose a reduced word $v = s_{i_1} \cdots s_{i_k}$. We write $u \leqslant v$ if there is a subsequence $(j_1, \dots, j_\ell)$ of $(i_1, \dots, i_k)$ such that $u = s_{j_1} \cdots s_{j_\ell}$. Because of the following property of Coxeter groups, if such a subsequence exists, it may be arranged so that $u = s_{j_1} \cdots s_{j_\ell}$ is a reduced expression.

Proposition 6.11. *Let $w = s_{i_1} \cdots s_{i_k}$ be a product of k simple reflections such that $\ell(w) < k$. Then it is possible to omit two of the factors and get another reduced expression:*

$$w = s_{i_1} \cdots \widehat{s}_{i_a} \cdots \widehat{s}_{i_b} \cdots s_{i_k},$$

where the "hat" means a factor is omitted, with $1 \leqslant a < b \leqslant k$.

Proof. This is Bump 2013, Proposition 20.4. □

Proposition 6.12. *The definition of the Bruhat order does not depend on the reduced expression $v = s_{i_1} \cdots s_{i_k}$.*

Proof. Matsumoto's theorem shows that it suffices to show there exists such an expression $v = s_{i'_1} \cdots s_{i'_k}$ when $(i_1, \dots, i_k)$ and $(i'_1, \dots, i'_k)$ differ by a braid relation, and this case may be easily handled. □

2. Demazure operators

Let G, T, Λ, W be as in Section 1. Let $\mathcal{O}(T)$ be the ring of rational functions on $\mathcal{O}(T)$. This is isomorphic to the group algebra $\mathbb{C}[\Lambda]$ of the weight lattice; as a ring, it is a Laurent polynomial ring. Let $\mathcal{O}(T_{\mathrm{reg}})$ be the ring obtained from $\mathcal{O}(T)$ by adjoining the reciprocals of polynomials $1 - \mathbf{z}^\alpha$ with $\alpha \in \Phi$. Since $1 - \mathbf{z}^\alpha$ and $1 - \mathbf{z}^{-\alpha}$ differ by the unit $-\mathbf{z}^{-\alpha}$ we need only adjoin $1 - \mathbf{z}^\alpha$ for $\alpha \in \Phi^+$. The ring $\mathcal{O}(T_{\mathrm{reg}})$ can be regarded as the ring of functions that are regular on the set T_{reg} of *regular elements* of T, defined as those $\mathbf{z} \in T$ such that the functions $1 - \mathbf{z}^\alpha$ are all nonvanishing nonvanishing.

For example if $G = \mathrm{GL}(n, \mathbb{C})$ and $T \cong (\mathbb{C}^\times)^n$ is the diagonal torus, then $\mathcal{O}(T)$ is isomorphic to the Laurent polynomial ring $\mathbb{C}[z_1, z_1^{-1}, \dots, z_n, z_n^{-1}]$, and elements are holomorphic functions of $\mathbf{z} = (z_1, \dots, z_n) \in T$. The ring T_{reg} is the set of $\mathbf{z}$ such that the components z_i, z_j are distinct. The ring $\mathcal{O}(T_{\mathrm{reg}})$ is $\mathbb{C}[z_i, z_i^{-1}, (1 - z_i/z_j)^{-1}] = \mathbb{C}[z_i, z_i^{-1}, (z_i - z_j)^{-1}]$.

Definition 6.13. Let $\mathcal{R}$ be the ring generated by $\mathcal{O}(T_{\mathrm{reg}})$ and W, subject to the relations

$$wfw^{-1} = {}^w f, \qquad {}^w f(\mathbf{z}) = f(w^{-1}\mathbf{z}), \qquad w \in W, \qquad f \in \mathcal{O}(T).$$

There is a representation $\mathcal{R} \longrightarrow \mathrm{End}(\mathcal{O}(T_{\mathrm{reg}}))$, in which $f \in \mathcal{O}(T_{\mathrm{reg}})$ acts by multiplication, and an element $w \in W$ acts through the usual action of W on $\mathcal{O}(T_{\mathrm{reg}})$. So we will regard elements of $\mathcal{R}$ as operators on $\mathcal{O}(T)$. *A priori* such an operator takes $\mathcal{O}(T)$ into $\mathcal{O}(T_{\mathrm{reg}})$, since they could introduce denominators of the form $(1 - \mathbf{z}^\alpha)^{-1}$, but we will consider certain *divided difference* operators that take $\mathcal{O}(T)$ into $\mathcal{O}(T)$.

Let s_i be a simple reflection. We define two elements of $\mathcal{R}$ as follows. Let

$$\partial_i = (1 - \mathbf{z}^{-\alpha_i})^{-1}(1 - \mathbf{z}^{-\alpha_i} s_i), \qquad \partial_i^\circ = (\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - s_i).$$

These are *Demazure operators* acting on $\mathcal{O}(T)$. Explicitly

$$\partial_i f(\mathbf{z}) = \frac{f(\mathbf{z}) - \mathbf{z}^{-\alpha_i} f(s_i \mathbf{z})}{1 - \mathbf{z}^{-\alpha_i}}, \qquad \partial_i^\circ f(\mathbf{z}) = \frac{f(\mathbf{z}) - f(s_i \mathbf{z})}{\mathbf{z}^{\alpha_i} - 1}.$$

The relation

$$\partial_i = \partial_i^\circ + 1 \tag{6.7}$$

is easily checked.

Lemma 6.14. *If $f \in \mathcal{O}(T)$, then $\partial_i f, \partial_i^\circ f \in \mathcal{O}(T)$.*

Proof. The operators introduce denominators, since $(\mathbf{z}^{\alpha_i} - 1)^{-1}$ is not in $\mathcal{O}(T)$. However the numerators of both ∂_i and ∂_i° also vanish when $\mathbf{z}^{\alpha_i} = 1$. This is because $\mathbf{z} = s_i\mathbf{z}$, and so $f = s_i(f)$. Therefore the quotients are rational functions without poles. □

Proposition 6.15. *The operators ∂_i° satisfy the same braid relations as the s_i, and the quadratic relation*

$$(\partial_i^\circ)^2 = -\partial_i^\circ. \tag{6.8}$$

Therefore they generate a ring of operators isomorphic to $\mathcal{H}_0$.

Proof. See Bump 2013, Proposition 25.1 for a proof of the braid relations. (The operators ∂_i° are denoted D_i there.) We will check (6.8) as follows.

$$(\partial_i^\circ)^2 = (\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - s_i)(\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - s_i).$$

We expand the first $1 - s_i$ obtaining two terms. The first is $(\mathbf{z}^{\alpha_i} - 1)^{-1}(\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - s_i)$ and the second is

$$-(\mathbf{z}^{\alpha_i} - 1)^{-1}s_i(\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - s_i) = -(\mathbf{z}^{\alpha_i} - 1)^{-1}(\mathbf{z}^{-\alpha_i} - 1)^{-1}s_i(1 - s_i).$$

Combining the two terms gives

$$(\mathbf{z}^{\alpha_i} - 1)^{-1}\left(\frac{1}{\mathbf{z}^{\alpha_i} - 1} + \frac{1}{\mathbf{z}^{-\alpha_i} - 1}\right)(1 - s_i) = -\partial_i^\circ$$

since $\frac{1}{x-1} + \frac{1}{x^{-1}-1} = -1$. □

Proposition 6.16. *The operators ∂_i also satisfy the same braid relations as the s_i, and also the quadratic relation*

$$\partial_i^2 = \partial_i. \tag{6.9}$$

We have $s_i\partial_i = \partial_i$.

Proof. The braid relations are proved in Bump 2013, Proposition 25.3, but note a typo: the wrong font is used for the operators ∂_i in the statement of the theorem. The relations $\partial_i^2 = s_i\partial_i = \partial_i$ are easily checked along the lines of (6.8) in Proposition 6.15. □

Both ∂_i° and ∂_i satisfy the same braid relations as the s_i. Let $w = s_{i_1}\cdots s_{i_k}$ be a reduced decomposition of $w \in W$. Define

$$\partial_w^\circ = \partial_{i_1}^\circ \cdots \partial_{i_k}^\circ, \qquad \partial_w = \partial_{i_1}\cdots\partial_{i_k}.$$

Because both ∂_i° and ∂_i satisfy the same braid relations as the s_i, Matsumoto's theorem shows that ∂_w° and ∂_w are well defined.

Let $\Omega \in \mathcal{R}$ be the operator

$$\Omega = \prod_{\alpha\in\Phi^+}(1 - \mathbf{z}^{-\alpha})^{-1}\sum_{w\in W}(-1)^{\ell(w)}\mathbf{z}^{w(\rho)-\rho}\,w. \tag{6.10}$$

By the Weyl character formula, if λ is a dominant weight

$$\Omega(\mathbf{z}^\lambda) = \chi_\lambda(\mathbf{z})$$

is the character of the irreducible representation with highest weight λ, as a function on $T \subset G$. For $\mathrm{GL}(n)$, $\Omega(\mathbf{z}^\lambda) = s_\lambda$ is a Schur polynomial.

Theorem 6.17 (Demazure)**.** *Let w_0 be the long element of W, and let λ be a dominant weight. Then $\partial_{w_0} = \Omega$. Therefore*

$$\chi_\lambda(\mathbf{z}) = \partial_{w_0} \mathbf{z}^\lambda.$$

Proof. See Bump 2013, Theorem 25.3 for a proof. □

2.1. Demazure characters and Demazure crystals. First, in Theorem 6.19 we clarify the relationship between the operators ∂_w° and ∂_w. If $w \in W$ and if s is a simple reflection, then either $sw > w$ or $sw < w$ (not both). We say that s is a left *ascent* (resp. *descent*) if $sw > w$ (resp. $sw < w$).

Proposition 6.18. *Let $y, w \in W$ and let s be a simple reflection. Assume that $w < sw$ and $y < sy$. Then the following are equivalent:*
(i) $y \leqslant w$;
(ii) $y \leqslant sw$;
(iii) $sy \leqslant sw$.

Proof. See Deodhar 1977, Property Z or Björner and Brenti 2005, Theorem 2.2.7. □

Theorem 6.19. *Let $w \in W$. Then*

$$\partial_w = \sum_{y \leqslant w} \partial_y^\circ. \tag{6.11}$$

Proof. If $w = 1$ this is clear and both sides equal 1. Arguing by induction on $\ell(w)$, we assume (6.11) and prove the same identity with w replaced by sw, where s is any left ascent of w. Using our induction hypothesis

$$\partial_{sw} = \partial_s \partial_w = \partial_s \sum_{y \leqslant w} \partial_y^\circ.$$

Consider a term where $sy < y$. We will argue that $\partial_s \partial_y^\circ = 0$. Indeed, $\partial_y^\circ = \partial_s^\circ \partial_{sy}^\circ$ since $\ell(y) = \ell(sy) + 1$. Now $\partial_s \partial_s^\circ = (1 + \partial_s^\circ)\partial_s^\circ = 0$ by (6.8), and the statement follows. So we may ignore these terms and remembering (6.7) we have

$$\partial_{sw} = (1 + \partial_s^\circ) \sum_{\substack{y \leqslant w \\ y < sy}} \partial_y^\circ = \sum_{\substack{y \leqslant w \\ y < sy}} \partial_y^\circ + \sum_{\substack{y \leqslant w \\ y < sy}} \partial_{sy}^\circ.$$

Now we may apply Proposition 6.18 with $y < sy$ and $w < sw$. Then $y \leqslant w$ if and only if $y \leqslant sw$ if and only if $sy \leqslant sw$, so this equals

$$\sum_{\substack{y \leqslant sw \\ y < sy}} \partial_y^\circ + \sum_{\substack{sy \leqslant sw \\ y < sy}} \partial_{sy}^\circ.$$

In the second term we replace y by sy and obtain

$$\partial_{sw} = \sum_{\substack{y \leqslant sw \\ y < sy}} \partial_y^\circ + \sum_{\substack{y \leqslant sw \\ sy < y}} \partial_y^\circ = \sum_{y \leqslant sw} \partial_y^\circ.$$

□

If $w \in W$ then the function

$$\chi_\lambda(\mathbf{z}; w) := \partial_w \mathbf{z}^\lambda$$

is called a *Demazure character* or *key polynomial*. These may be defined for any Cartan type, and were shown by Demazure 1974 to compute spaces of sections of line bundles on Schubert varieties. If $w = w_0$ then by Theorem 6.17 this is the character of an irreducible representation (i.e. a Schur polynomial if $G = \mathrm{GL}(n)$).

Proposition 6.20. *Let W_λ be the stabilizer in W of λ. If w, w' lie in the same coset wW_λ then $\chi_\lambda(\mathbf{z}; w) = \chi_\lambda(\mathbf{z}; w')$. The coset wW_λ contains a unique element of shortest length.*

Proof. With $\lambda = (\lambda_1, \ldots, \lambda_n)$ a partition, its stabilizer in S_n is the subgroup generated by s_i such that $\lambda_i = \lambda_{i+1}$. A subgroup generated by simple reflections is called a *Young* or *parabolic* subgroup of S_n, and every coset of a parabolic contains a unique element of shortest length.

For the first assertion, since W_λ is parabolic it is sufficient to $\partial_w \mathbf{z}^\lambda = \partial_{w'} \mathbf{z}^\lambda$ if $w' = ws_i$ where $s_i\lambda = \lambda$. Then $\partial_{w'} = \partial_w \partial_i$, and it follows from the definition of ∂_i that if $s_i\lambda = \lambda$ then $\partial_i \mathbf{z}^\lambda = \mathbf{z}^\lambda$. □

Proposition 6.21. *Let W_λ be the stabilizer in W of λ. If w, w' lie in the same coset wW_λ then $\chi_\lambda(\mathbf{z}; w) = \chi_\lambda(\mathbf{z}; w')$. The coset wW_λ contains a unique element of shortest length.*

Proof. With $\lambda = (\lambda_1, \ldots, \lambda_n)$ a partition, its stabilizer in S_n is the subgroup generated by s_i such that $\lambda_i = \lambda_{i+1}$. A subgroup generated by simple reflections is called a *Young* or *parabolic* subgroup of S_n, and every coset of a parabolic contains a unique element of shortest length.

For the first assertion, since W_λ is parabolic it is sufficient to $\partial_w \mathbf{z}^\lambda = \partial_{w'} \mathbf{z}^\lambda$ if $w' = ws_i$ where $s_i\lambda = \lambda$. Then $\partial_{w'} = \partial_w \partial_i$, and it follows from the definition of ∂_i that if $s_i\lambda = \lambda$ then $\partial_i \mathbf{z}^\lambda = \mathbf{z}^\lambda$. □

The related polynomials $\chi^\circ_\lambda(\mathbf{z}; w) := \partial^\circ_w \mathbf{z}^\lambda$ were defined in Lascoux and Schützenberger 1990, who called them *standard bases*. Nowadays following Mason 2009 they are called *Demazure atoms*. By Theorem 6.19 we have

$$\chi_\lambda(\mathbf{z}; w) = \sum_{y \leqslant w} \chi^\circ_\lambda(\mathbf{z}; y). \tag{6.12}$$

These facts can be lifted to statements about crystals. Following Kashiwara 1993; Littelmann 1995a, we may lift the definition of the operator ∂_i to the crystal $\mathcal{B}_\lambda$ of SSYT of shape λ as follows. First note that for any weight μ we have $s_i(\mu) = \mu - k\alpha_i$ and where $k = \langle \alpha_i^\vee, \mu \rangle$, and so

$$\partial_i \mathbf{z}^\mu = \frac{1 - \mathbf{z}^{-\alpha_i(1+k)}}{1 - \mathbf{z}^{-\alpha_i}} = \begin{cases} \mathbf{z}^\mu + \mathbf{z}^{\mu-\alpha_i} + \cdots + \mathbf{z}^{\mu-k\alpha_i} & \text{if } k \geqslant 0, \\ 0 & \text{if } k = -1, \\ -(\mathbf{z}^{\mu+\alpha_i} + \mathbf{z}^{\mu+2\alpha_i} + \cdots + \mathbf{z}^{\mu+(-k-1)\alpha_i}) & \text{if } k < -1, \end{cases}$$

by the geometric series formula. This lifts to the crystal using the operators e_i and f_i. Namely, we may define an endomorphism of the free abelian group $\mathbb{Z}[\mathcal{B}_\lambda]$ by

$$\partial_i v = \begin{cases} v + f_i(v) + \cdots + f_i^k(v) & \text{if } k \geqslant 0, \\ 0 & \text{if } k = -1, \\ -(e_i(v) + \cdots + e_i^{-k-1}(v)) & \text{if } k < -1, \end{cases}$$

for $v \in \mathcal{B}_\lambda$. Now Kashiwara 1993; Littelmann 1995a proved that there exist subsets $\mathcal{B}_\lambda(w)$ called *Demazure crystals* such that if $w = s_{i_1} \cdots s_{i_k}$ is a reduced expression and if T_λ is the highest weight element of $\mathcal{B}_\lambda$ then

$$\partial_{i_1} \cdots \partial_{i_k} T_\lambda = \sum_{T \in \mathcal{B}_\lambda(w)} T.$$

See also Bump and Schilling 2017 Chapter 13.

Moreover there exist subsets $\mathcal{B}^\circ_\lambda(w)$, called *crystal Demazure atoms*, such that

$$\mathcal{B}_\lambda(w) = \bigcup_{y \leqslant w} \mathcal{B}^\circ_\lambda(y).$$

A *key tableau* is one in which every column (except the first) is a subset of the previous column. The weight of a key tableau of shape λ is $w(\lambda)$ for some $w \in W$. Lascoux and Schützenberger 1990 defined a (right) key map from tableaux to key tableaux. For $\mathrm{GL}(n)$ crystals, the existence of the subsets $\mathcal{B}^\circ_\lambda(w)$ is essentially due to Lascoux and Schützenberger, because the subsets $\mathcal{B}^\circ_\lambda(w)$ can be characterized as the set of tableaux whose right key has weight $w(\lambda)$. For other Cartan types, the existence of crystal Demazure atoms is proved in Kashiwara 2002, Section 9.1.

The subsets $\mathcal{B}^\circ_\lambda(w)$ are disjoint. If λ is strongly dominant, they are all nonempty. Otherwise $\mathcal{B}^\circ_\lambda(w)$ is nonempty if and only if w is the shortest element in the coset wW_λ.

2.2. The nil-Hecke algebra. Demazure operators are also commonly called *divided difference operators*. The operators ∂_w and ∂°_w can be defined for any root system, but here we will consider operators D_w that are special to $G = \mathrm{GL}(n)$ and $W = S_n$. Define

$$D_i = z_{i+1}^{-1} \partial^\circ_i, \qquad D_i f(\mathbf{z}) = (z_i - z_{i+1})^{-1}(f(\mathbf{z}) - f(s_i \mathbf{z})).$$

Proposition 6.22.

(i) *The functions D_i satisfy the same braid relations as $s_i \in S_n$. Thus $D_i D_j D_i = D_j D_i D_j$ if $j = i \pm 1$, and $D_i D_j = D_j D_i$ if $|i - j| > 1$.*
(ii) *The operators D_i satisfy the quadratic relation $D_i^2 = 0$.*

Proof. It is convenient to work in the ring $\mathcal{R}$ again, and write $D_i = (z_i - z_{i+1})^{-1}(1 - s_i)$.

Assume $j = i + 1$. Then s_i and s_j generate a subgroup $\langle s_i, s_j \rangle$ isomorphic to S_3. Expanding out $D_i D_{i+1} D_i$ and rearranging, we find

$$D_i D_j D_i = c^{-1} \cdot \sum_{w \in \langle s_i, s_j \rangle} (-1)^{\ell(w)} w,$$

where the constant $c = (z_i - z_{i+1})(z_i - z_{i+2})(z_{i+1} - z_{i+2})$. A small amount of algebra is needed to simplify the coefficients of 1 and s_i. But $D_{i+1} D_i D_{i+1}$ simplifies to the same expression. The proof that $D_i D_j = D_j D_i$ if $|i - j| > 1$ is easy. As for the quadratic relation, D_i^2 equals

$$(z_i - z_{i+1})^{-1}(1 - s_i)(z_i - z_{i+1})^{-1}(1 - s_i) = (z_i - z_{i+1})^{-2}(1 + s_i)(1 - s_i) = 0.$$

□

Since the D_i satisfy the braid relations, we may define D_w for $w \in W$ the same way that ∂_w and ∂°_w were defined.

The algebra generated by the D_i, subject to the braid relations and the quadratic relations $D_i^2 = 0$ is called the *nil-Hecke algebra*. It can be defined for any Cartan type by these defining relations, though the particular operators D_i are only available in Type A. A more subtle approach in Kostant and Kumar 1986 realizes the nil-Hecke algebra in terms of Demazure operators for all Cartan types. The nil-Hecke algebra is important for understanding the cohomology of flag varieties, and the definition of Schubert polynomials.

2.3. Demazure–Lusztig operators. Let q be a parameter. We define the *Demazure–Lusztig operator*

$$\mathcal{L}_i = (\mathbf{z}^{\alpha_i} - 1)^{-1}\Big((1 - s_i) - q(1 - \mathbf{z}^{\alpha_i} s_i)\Big).$$

Similarly, we define the *Demazure–Whittaker operator*

$$\mathcal{T}_i = (\mathbf{z}^{\alpha_i} - 1)^{-1}\Big((1 - s_i) - q(1 - \mathbf{z}^{-\alpha_i} s_i)\Big).$$

Thus if $q = 0$ both operators specialize to the operator ∂_i. We will also consider the conjugation ${}^{\rho}\mathcal{T}_i := \mathbf{z}^{\rho}\mathcal{T}_i\mathbf{z}^{-\rho}$. Because $s_i(\rho) = \rho - \alpha_i$ we have

$${}^{\rho}\mathcal{T}_i = (\mathbf{z}^{\alpha_i} - 1)^{-1}\Big((1 - \mathbf{z}^{\alpha_i} s_i) - q(1 - s_i)\Big).$$

Theorem 6.23 (Lusztig)**.** *Let $\mathfrak{T}_i$ denote either $\mathcal{L}_i$ or $\mathcal{T}_i$, or ${}^{\rho}\mathcal{T}$. The operators $\mathfrak{T}_i$ satisfy the braid and quadratic relations*

$$\mathfrak{T}_i^2 = (q - 1)\mathfrak{T}_i + q,$$

and hence generate an algebra isomorphic to $\mathcal{H}(W)$.

Proof. The quadratic relation may be checked by expanding $\mathcal{L}_i^2$ or $\mathcal{T}_i^2$ in the ring $\mathcal{R}$. The braid relation for $\mathcal{L}_i$ is harder and is proved in Lusztig 1985. For a proof of the quadratic relation for $\mathcal{T}_i$ based on intertwining operators on Whittaker functions for p-adic representations, see Brubaker, Bump, and Licata 2015, Theorem 2. The same argument applied to Iwahori spherical functions would work for $\mathcal{L}_i$. See Brubaker, Bump, and Friedberg 2016. □

Theorem 6.23 implies that we have two representations σ_s and σ_a of $\mathcal{H}_q$ in which $\sigma_s(T_i) = \mathcal{L}_i$, and $\sigma_a(T_i) = \mathcal{T}_i$. But these may be extended to the affine Hecke algebra, as we will now explain.

The (extended) *affine Hecke algebra* $\tilde{\mathcal{H}}_q = \tilde{\mathcal{H}}_q(W)$ is described as follows. It contains a copy of $\mathcal{H}_q$, and also an abelian algebra θ_Λ spanned as a vector space by elements θ_λ ($\lambda \in \Lambda$). These satisfy $\theta_\lambda\theta_\mu = \theta_{\lambda+\mu}$, and the *Bernstein relation*

$$\theta_\lambda T_i - T_i\theta_{s_i\lambda} = (q - 1)\frac{\theta_\lambda - \theta_{s_i\lambda}}{1 - \theta_{-\alpha_i}}.$$

The multiplication map $\theta_\Lambda \times \mathcal{H}_q \to \tilde{\mathcal{H}}_q$ induces a vector space isomorphism $\theta_\Lambda \otimes \mathcal{H}_q \cong \widetilde{\mathcal{H}_q}$.

Now extend σ_s and σ_a to $\tilde{\mathcal{H}}_q$ in which θ_λ acts as multiplication by $\mathbf{z}^{-\lambda}$. The Bernstein relation must be checked. The representations σ_s and σ_a are called (respectively) the *spherical* and *antispherical* representations of the affine Hecke algebra.

We recall the operator Ω defined in (6.10). We will work in the ring $\mathcal{R}_q := \mathbb{C}(q) \otimes \mathcal{R}$, where we have extended the ground field to contain q. (We recall that q can be either a complex number or an indeterminate.) Similarly let $\mathcal{O}_q(T_{\mathrm{reg}}) = \mathbb{C}(q) \otimes \mathcal{O}(T_{\mathrm{reg}})$.

Proposition 6.24. *We have*

$$\sum_{w\in W}\mathcal{L}_w = \Omega\prod_{\alpha\in\Phi^+}(1-q\mathbf{z}^{-\alpha}), \qquad \sum_{w\in W}\mathcal{T}_w = \prod_{\alpha\in\Phi^+}(1-q\mathbf{z}^{-\alpha})\Omega \tag{6.13}$$

in the ring $\mathcal{R}_q$.

Proof. We note the identities

$$1+\mathfrak{L}_i = \partial_i(1-q\mathbf{z}^{-\alpha_i}), \qquad 1+\mathfrak{T}_i = (1-q\mathbf{z}^{-\alpha_i})\partial_i, \tag{6.14}$$

which are easily checked in the ring $\mathcal{R}_q$.

Let us denote

$$\Theta = \sum_{w\in W}\mathcal{L}_w, \qquad \Theta' = \Omega\prod_{\alpha\in\Phi^+}(1-q\mathbf{z}^{-\alpha}).$$

We begin by checking that for $w\in W$ we have $w\Theta=\Theta$ and $w\Theta'=\Theta'$. It is sufficient to prove these when $w=s_i$ is a simple reflection. We have

$$\Theta = (1+\mathcal{L}_i)\sum_{\substack{w\in W\\ s_iw>w}}\mathcal{L}_w = \partial_i(1-q\mathbf{z}^{-\alpha_i})\sum_{\substack{w\in W\\ s_iw>w}}\mathcal{L}_w,$$

so $s_i\Theta=\Theta$ and $s_i\Theta'=\Theta'$ follow from the identities $s_i\partial_i=\partial_i$ and $s_i\Omega=\Omega$.

Now let $\Theta=\sum_{w\in W}\theta_w\cdot w$ and $\Theta'=\sum_{w\in W}\theta'_w\cdot w$ for $\theta_w,\theta_{w'}\in\mathcal{O}_q(T_{\mathrm{reg}})$. Then since $w\Theta=\Theta$ and $w\Theta'=\Theta'$ we have for $y\in W$ the identities $\theta_{yw}=y\theta_w$ and $\theta'_{yw}=y\theta'_w$. This means that if we verify $\theta_w=\theta'_w$ for any particular w, then $\Theta=\Theta'$. We will show that $\theta_{w_0}=\theta'_{w_0}$.

Let $w_0=s_{i_1}\cdots s_{i_N}$ be a reduced expression for the long element. For any root α define

$$\phi(\alpha)=\frac{1-q}{\mathbf{z}^\alpha-1}, \qquad \psi(\alpha)=\frac{1-q\mathbf{z}^\alpha}{1-\mathbf{z}^\alpha}.$$

Then $\mathcal{L}_i=\phi(\alpha_i)+\psi(\alpha_i)s_i$, so

$$\mathcal{L}_w = (\phi(\alpha_{i_1})+\psi(\alpha_{i_1})s_{i_1})\cdots(\phi(\alpha_{i_N})+\psi(\alpha_{i_N})s_{i_N}).$$

Expanding out this product, there is only one way to get a coefficient of w_0 which is to take all the $\psi(\alpha_i)$ factors and none of the $\phi(\alpha_i)$. Thus

$$\theta_{w_0}\cdot w_0 = \psi(\alpha_{i_1})s_{i_1}\psi(\alpha_{i_2})s_{i_2}\cdots\psi(\alpha_{i_N})s_{i_N}.$$

Moving all the s_{i_k} to the right conjugates the $\psi(\alpha_i)$ to produce

$$\theta_{w_0}\cdot w_0 = \psi(\alpha_{i_1})\psi(s_{i_1}\alpha_{i_2})\cdots\psi(s_{i_1}s_{i_2}\cdots s_{i_{N-1}}\alpha_{i_{N-1}})\cdot s_{i_1}\cdots s_{i_N}.$$

Now by Proposition 6.8 applied with $w=w_0^{-1}=s_{i_N}\cdots s_{i_1}$, the arguments of ψ are all positive roots, proving

$$\theta_{w_0} = \prod_{\alpha\in\Phi^+}\psi(\alpha).$$

On the other hand from the definition of Ω and Θ',

$$\theta'_{w_0}\cdot w_0 = \prod_{\alpha\in\Phi^+}(1-\mathbf{z}^{-\alpha})^{-1}\mathbf{z}^{w_0(\rho)-\rho}\quad w_0\prod_{\alpha\in\Phi^+}(1-q\mathbf{z}^{-\alpha}).$$

Since $\rho-w_0(\rho)$ is the sum of the positive roots, we see that $\theta'_{w_0}=\theta_{w_0}$.

We have proved the first identity in (6.13). The second identity can be proved similarly (see Exercise 6.1). □

3. Demazure operators and colored models

We have introduced five types $\partial_i, \partial_i^\circ, D_i, \mathcal{L}_i$ and $\mathcal{T}_i$ of divided difference or Demazure operators in this chapter. We will consider the actions of these on partition functions. Moreover, if $\mathfrak{T}$ is one of these operators we may consider the ρ-conjugated function ${}^\rho\mathfrak{T}_i = \mathbf{z}^\rho \mathfrak{T}_i \mathbf{z}^{-\rho}$. Here are these operators and some of their ρ-conjugates.

$\partial_i = (1-\mathbf{z}^{-\alpha_i})^{-1}(1-\mathbf{z}^{-\alpha_i}s_i)$	$\partial_i^2 = \partial_i$
$\partial_i^\circ = (\mathbf{z}^{\alpha_i}-1)^{-1}(1-s_i)$	$(\partial_i^\circ)^2 = -\partial_i^\circ$
$D_i = (z_i - z_{i+1})^{-1}(1-s_i)$	$D_i^2 = 0$
$\mathcal{L}_i = (\mathbf{z}^{\alpha_i}-1)^{-1}((1-s_i) - q(1-\mathbf{z}^{\alpha_i}s_i))$	$\mathcal{L}_i^2 = (q-1)\mathcal{L}_i + q$
$\mathcal{T}_i = (\mathbf{z}^{\alpha_i}-1)^{-1}((1-s_i) - q(1-\mathbf{z}^{-\alpha_i}s_i))$	$\mathcal{T}_i^2 = (q-1)\mathcal{T}_i + q$
${}^\rho\partial_i = (1-\mathbf{z}^{-\alpha_i})^{-1}(1-s_i)$	
$\partial_i^\circ = (\mathbf{z}^{\alpha_i}-1)^{-1}(1-\mathbf{z}^{\alpha_i}s_i)$	
${}^\rho\mathcal{T}_i = (\mathbf{z}^{\alpha_i}-1)^{-1}((1-\mathbf{z}^{\alpha_i}s_i) - q(1-s_i))$	

These operators all satisfy the braid relations, discussed earlier in this chapter. We are listing the quadratic relations for the unconjugated operators; naturally $\mathfrak{T}_i$ and ${}^\rho\mathfrak{T}_i$ satisfy the same braid and quadratic relations.

Let $\mathfrak{T}_i$ be one of the operators ${}^\rho\partial_i$, ${}^\rho\partial_i^\circ$, D_i, $\mathcal{L}_i$ or ${}^\rho\mathcal{T}_i$. For each of the these cases we will exhibit a family $\{\mathfrak{S}_w\}$ of lattice models indexed by $w \in W$ satisfying a Demazure recursion for the partition functions. If $s_i w > w$, then

$$Z(\mathfrak{S}_{s_i w}) = \mathfrak{T}_i Z(\mathfrak{S}_w). \tag{6.15}$$

This is complemented by the fact that for a particular value of w (often $w = 1$) system $\mathfrak{S}_w$ is monostatic, so that $Z(\mathfrak{S}_w)$ is known. From these facts all partition functions from the family can be evaluated.

3.1. Boundary conditions. We consider a class of colored models that gives examples of all of these operators. In place of the $\ominus$ spin we have a family of n *colors* $c_1, \ldots, c_n$ ordered so that $c_1 \leqslant c_2 \leqslant \cdots \leqslant c_n$. We also have the $\oplus$ spin connoting the absence of color. The spinset of all edges in the model have spinset $\Sigma = \{\oplus, c_1, \ldots, c_n\}$. Although we allow repetitions among the c_i, the $\oplus$ spin is not allowed to be one of the c_i so the cardinality of Σ is at least two. Let V be the free vector space on Σ. For notational reasons, we denote $\oplus$ as v_+ if it is considered to be a vector in V.

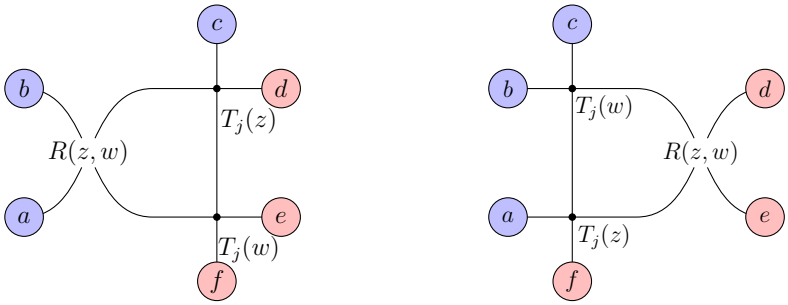
Let $R(z,w)$ be a homogeneous parametrized Yang–Baxter equation. In the RTT equations that leads to the Demazure recurrence (6.15), this is the R-matrix. The T-matrices may or may not be drawn from the same homogenous parametrized Yang–Baxter equation.

Now we consider a grid with n rows, and N columns, with row parameters $\mathbf{z} = (z_1, \ldots, z_n)$. As in Chapter 3 and Chapter 4 we label the rows 1 to n from top to bottom, and the columns from 0 to $N-1$ right to left. At the vertex in row i and column j, we place a vertex $T_j(z_i)$, whose Boltzmann weights need to be specified. It is required that for every j we have a Yang–Baxter RTT equation

$$[\![R(z,w), T_j(z), T_j(w)]\!] = 0 \tag{6.16}$$

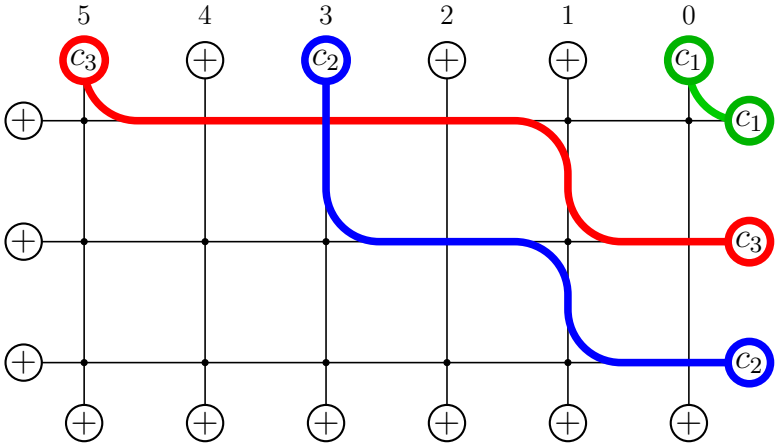
or equivalently, the two following systems must be equivalent. If the T-matrix $T_j(z)$ does not depend on the column j we just write $T_j(z) = T(z)$. As usual, this RTT equation amounts

to the equivalence of the two systems:



The boundary conditions are colored generalizations of the extended wall boundary conditions introduced in Chapter 3. In the examples that we will consider, the Boltzmann weights will be viatic, that is, they will satisfy the conservation of color (Section 1.1). This means that the edges of a given color can be organized into paths, moving downward and to the right. We will consider only models with $\oplus$ spins on the left and bottom boundary edges, so the colored boundary edges are only at the top and right. Thus the colored paths start at the top and end on the right. We allow uncolored $\oplus$ boundary edges at the top, but not on the right.

We call these the *colored extended wall boundary conditions*. If $\lambda = 0$ and $n = N$ (so the grid is square) we will call these boundary conditions *colored domain wall boundary conditions*.



(6.17)

The number N of columns must be at least $\lambda_1 + n - 1$, and the columns are labeled 0 to $N-1$ from right to left (as in Chapters 3 and 4). We may fix a partition $\lambda = (\lambda_1, \dots, \lambda_n)$ to describe the locations of colors on the top boundary edges, putting a color in the columns $\lambda_j + n - j$.

By a *banner* we mean an n-tuple of colors. (These are sometimes also called *flags* but to avoid confusion with flag varieties we will use the term "banner.") The Weyl group acts on banners as follows. For $w \in W$ let

$$w\mathbf{d} = (d_{w^{-1}(1)}, \dots, d_{w^{-1}(n)}).$$

Remark 6.25. Note that this is a *left* action. Indeed, if $\mathbf{d}' = (d'_1, \dots, d'_n) := w_2\mathbf{d}$, then $d'_i = d_{w_2^{-1}(i)}$. So $w_1\mathbf{d}' = (d'_{w_1^{-1}(1)}, \dots, d'_{w_1^{-1}(n)})$ and $d'_{w_1^{-1}(i)} = d_{w_2^{-1}w_1^{-1}(i)} = d_{(w_1w_2)^{-1}(i)}$. Therefore $w_1(w_2\mathbf{d}) = (w_1w_2)\mathbf{d}$. This is basic but a source of possible confusion.

The boundary colors are described by two banners $\mathbf{d} = (d_1, \ldots, d_n)$ and $\mathbf{f} = (f_1, \ldots, f_n)$, where we put the color f_i on the right boundary edge in row i, and d_i on the top boundary edge in column $\lambda_j + n - j$. Now, there is a bijection between the top boundary colors at the top and the right boundary colors, in which corresponding colors are connected by a path. Thus the top boundary colors duplicate those on the right boundary edge, so no more than n can occur, either on the boundary or on any edge in any state.

3.2. Demazure opertors for systems with no repeated colors. We wish to parametrize the system by a Weyl group element w. For simplicity, we will assume in this section that there are no repeated colors, though else in the text we will *not* assume this.

Assumption 6.26. The colors c_i are distinct, so $c_1 < \cdots < c_n$.

We may choose one of two conventions for the Weyl group element w. We consider the top boundary banner $\mathbf{d}$ to be fixed, and suppress it from the notation; we are more interested in how the system depends on $\mathbf{f}$. Let $\mathbf{c}_0 = (c_1, \ldots, c_n)$ be the "standard" banner.

Example 6.27. In the state (6.17), red $= c_3$ is the largest color. The boundary flag $\mathbf{d}$ is $(c_3, c_2, c_1) = w_0\mathbf{c}_0$, and the right boundary flag $\mathbf{f}$ is $(c_3, c_1, c_2) = ww_0\mathbf{c}_0$, where $w = s_1s_2 = (123)$ in cycle notation.

Convention 6.28. We either let $w \in W$ be chosen so that $\mathbf{f} = w\mathbf{c}_0$, or we choose w so that $\mathbf{f} = ww_0\mathbf{c}_0$. Let $\mathfrak{S}_w = \mathfrak{S}_w(\mathbf{z})$ be the resulting system.

Lemma 6.29.

(1) *With the convention that* $\mathbf{f} = w\mathbf{c}_0$*, we have* $s_iw > w$ *if and only if* $f_i < f_{i+1}$.
(2) *With the convention that* $\mathbf{f} = ww_0\mathbf{c}_0$*, we have* $s_iw > w$ *if and only if* $f_i > f_{i+1}$.

Proof. In both cases, switching f_i and f_{i+1} increases the number of inversions in the partition w. □

Remark 6.30. We have chosen the number of rows and the number of colors are both the same number n. This is somewhat arbitrary but taking more colors than rows would not be helpful. The reason is that since clearly only colors in $\mathbf{d}$ can appear on any edge in any state, no more than n colors can appear in the boundary conditions, so we may limit the palette to just the n colors in $\mathbf{d}$.

Subject to a mild "orderly" assumption that we will explain presently, in Theorem 6.34 below we will show that the Demazure recurrence (6.15) is satisfied for a suitable Demazure operator $\mathfrak{T}_i$, depending only on R. We will find R producing the five operators tabulated at the beginning of this subsection. We will also see that the system $\mathfrak{S}_w$ can be evaluated explicitly for one particular choice of w, and the combination of (6.15) and this information determines $Z(\mathfrak{S}_w)$ for all w.

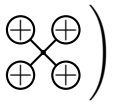
We now impose one more assumption. We say that the Boltzmann weights (already assumed viatic) are *orderly* if they depend only on the orders of the colors, and we make this assumption. In other words, denoting by $\beta_{z,w}$ the Boltzmann weights of $R(z, w)$, we assume that

(6.18) $$c(z, w; c, d) := \beta_{z,w}\left(\begin{array}{cc} c & c \\ d & d \end{array}\right), \qquad b(z, w; c, d) := \beta_{z,w}\left(\begin{array}{cc} d & c \\ c & d \end{array}\right),$$

only depend on the order $c < d$ or $d > c$. Hence there exist functions $c^{\pm}(z,w)$ and $b^{\pm}(z,w)$ such that

$$c(z,w;c,d) = \begin{cases} c^+(z,w) & \text{if } c > d, \\ c^-(z,w) & \text{if } c < d, \end{cases} \qquad b(z,w;c,d) = \begin{cases} b^+(z,w) & \text{if } c > d, \\ b^-(z,w) & \text{if } c < d. \end{cases} \tag{6.19}$$

We also denote

$$a(z,w) = \beta_{z,w} \left(\quad \right).$$


Assumption 6.31. We assume that the weights are orderly.

Lemma 6.32. *Suppose that c and d are distinct colors. Then $b(z,w;c,d)$ is a constant multiple of $z-w$. Furthermore*

$$a(z,z) = c(z,z;c,d).$$

Proof. As part of the definition of a homogeneous parametrized Yang–Baxter equation, the matrix $R(z,z)$ is assumed to be a scalar matrix. Both assertions follow from this fact.

First consider $b(z,w;c,d)$. With respect to the basis of $V \otimes V$ consisting of the spinset $\mathcal{C} \cup \{\oplus\}$, this Boltzmann weight is an off-diagonal entry, so it vanishes when $z = w$. Because it is a homogeneous linear polynomial in z and w, it is a multiple of $z - w$.

Now since $R(z,z)$ is a scalar matrix, write $R(z,z) = cI_{V\otimes V}$. Then $R(z,z)$ multiplies $v_+ \otimes v_+ \in V \otimes V$ by c, where v_+ denotes the image of the spin $\oplus$ in V. Thus $a(z,z) = c$, and similarly $c(z,z;c,d)$ is the same constant c. □

Lemma 6.32 does not guarantee that $b(z,w;c,d)$ is nonzero; indeed, it can vanish. But define operators $\mathfrak{T}_i^+$ and $\mathfrak{T}_i^-$ on $\mathcal{O}(T_{\mathrm{reg}})$ by

$$\mathfrak{T}_i^{\pm} f(\mathbf{z}) = \frac{c^{\pm}(z_{i+1},z_i) f(\mathbf{z}) - a(z_{i+1},z_i) f(s_i \mathbf{z})}{-b^{\pm}(z_{i+1},z_i)} \tag{6.20}$$

when the denominator is nonzero.

Proposition 6.33. *If both operators $\mathfrak{T}_i^+$ and $\mathfrak{T}_i^-$ are defined, they are inverse operators on $\mathcal{O}(T)$.*

Proof. See Exercise 6.2. □

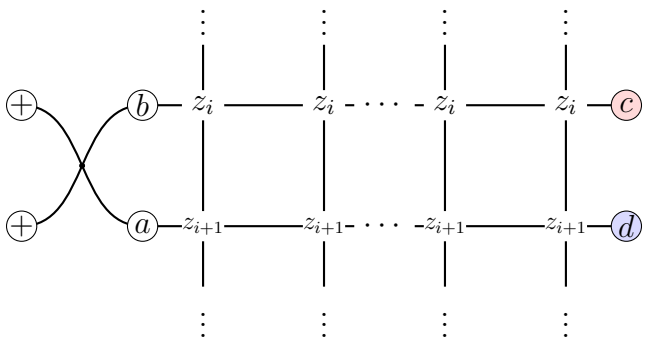
Theorem 6.34. *Assume that the Boltzmann weights are viatic and orderly. Then*

$$Z(\mathfrak{S}_{s_i w}) = \begin{cases} \mathfrak{T}_i^+ Z(\mathfrak{S}_w) & \text{if } f_i > f_{i+1} \text{ and } \mathfrak{T}_i^+ \text{ is defined;} \\ \mathfrak{T}_i^- Z(\mathfrak{S}_w) & \text{if } f_i < f_{i+1} \text{ and } \mathfrak{T}_i^- \text{ is defined.} \end{cases}$$

Proof. Let $c = c_{w^{-1}}(i)$ and $d = c_{w^{-1}}(i+1)$. These are the right edge boundary colors in the i and $i+1$ rows, respectively. Note that $c > d$ if and only if $w^{-1}(i) < w^{-1}(i+1)$, or equivalently $s_i w > w$. The two cases are similar: we assume that $c > d$ or $s_i w > w$.

The R-matrix for the train argument is $R(z_i, z_{i+1})$. The Boltzmann weights are given in Chapter 5, Section 4, with $z = z_1$, $w = z_2$. We attach the R-matrix between the i and $i+1$

rows, and a portion of the configuration looks like this:

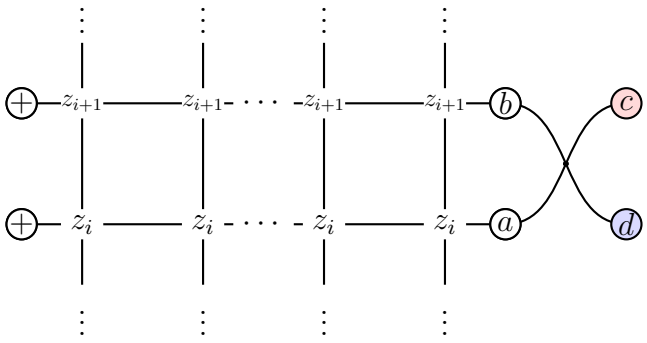


Here $c = w^{-1}(i)$ and $d = w^{-1}(i+1)$.

We are summing over the interior vertices a and b. Since the weights are viatic, the Boltzmann weight of the R-vertex is zero unless $a = b = \oplus$, so the partition function of this system is

$$\beta_{z_i,z_{i+1}} \left(\begin{array}{c} \oplus \quad \oplus \\ \oplus \quad \oplus \end{array} \right) Z(\mathfrak{S}_w(\mathbf{z})) = a(z_i, z_{i+1})\, Z(\mathfrak{S}_{\mathfrak{w}}(\mathbf{z})).$$

Running the train argument produces the system



Now there are two possibilities. If $b = c$ and $a = d$, the contribution is

$$\beta_{z_i,z_{i+1}} \left(\begin{array}{c} c \quad c \\ d \quad d \end{array} \right) Z(\mathbf{S}_w(s_i\mathbf{z})) = c(z_i, z_{i+1}; c, d) Z(\mathbf{S}_w(s_i\mathbf{z})).$$

On the other hand if $b = d$ and $a = c$, then we obtain a contribution

$$\beta_{z_i,z_{i+1}} \left(\begin{array}{c} d \quad c \\ c \quad d \end{array} \right) Z(\mathfrak{S}_{s_iw}(\mathbf{z})) = b(z_i, z_{i+1}; c, d) Z(\mathfrak{S}_{s_iw}(\mathbf{z})).$$

The banner has changed, explaining why w becomes s_iw in this term. So

$$a(z_i, z_{i+1})\, Z(\mathfrak{S}_w(\mathbf{z})) = c(z_i, z_{i+1}; c, d)\, Z(\mathfrak{S}_w(s_i\mathbf{z})) + b(z_i, z_{i+1}; c, d)\, Z(\mathfrak{S}_{s_iw}(s_i\mathbf{z})).$$

Now we replace $\mathbf{z}$ by $s_i\mathbf{z}$ (so $z_i \leftrightarrow z_{i+1}$) and reorganize to get

$$Z(\mathfrak{S}_{s_iw}(\mathbf{z})) = \frac{c(z_{i+1}, z_i; c, d) Z(\mathbf{z}; w) - a(z_{i+1}, z_i)\, Z(\mathfrak{S}_w(s_i\mathbf{z}))}{-b(z_{i+1}, z_i; c, d)} = \mathfrak{T}_i Z(\mathfrak{S}_w(\mathbf{z}))\ .$$

□

To summarize, Theorem 6.34 implies that if $s_i w > w$, then we may write

$$Z(\mathfrak{S}_{s_i w}(\mathbf{z})) = \mathfrak{T}_i Z(\mathfrak{S}_w(\mathbf{z})),$$

where (see Convention 6.28):

$$\mathfrak{T}_i = \begin{cases} \mathfrak{T}_i^+ & \text{if we use the convention } \mathbf{f} = ww_0\mathbf{c}_0, \\ \mathfrak{T}_i^- & \text{if we use the convention } \mathbf{f} = w\mathbf{c}_0, \end{cases}$$

provided that the operator is defined. Similarly if $s_i w < w$ then

$$Z(\mathfrak{S}_{s_i w}(\mathbf{z})) = \mathfrak{T}'_i Z(\mathfrak{S}_w(\mathbf{z})),$$

where

$$\mathfrak{T}'_i = \begin{cases} \mathfrak{T}_i^- & \text{if we use the convention } \mathbf{f} = ww_0\mathbf{c}_0, \\ \mathfrak{T}_i^+ & \text{if we use the convention } \mathbf{f} = w\mathbf{c}_0, \end{cases}$$

again subject to the assumption that the operator is defined. If both $\mathfrak{T}_i$ and $\mathfrak{T}'_i$ are defined, they are inverses.

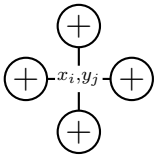
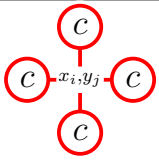
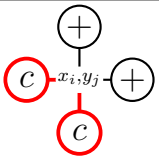
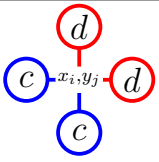
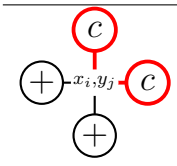
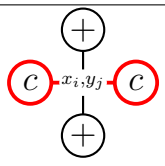
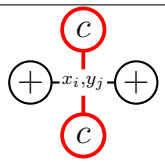
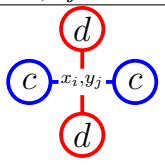
3.3. Examples. The operators $\mathfrak{T}_i^{\pm}$ from the last section really only depend on the R-matrix, not on the T-matrices. We begin by pointing out that if a homogenous parametrized Yang–Baxter equation is given, there are always models with those R-matrices, and for these the Demazure recursions of the previous subsection are valid.

To see this, we may draw the T-matrices from the same parametrized Yang–Baxter equation as the R-matrix. This gives a system with both row and column parameters. We use $\mathbf{x}$ and $\mathbf{y}$ when there are both row and column parameters, so in this paragraph $\mathbf{z} = \mathbf{x}$. We may then take $T_j(z) = R(z, y_j)$. The following identities hold

$$[\![R(x_i, x_{i+1}), T_j(x_i), T_j(x_{i+1})] = [\![R(x_i, x_{i+1}), R(x_i, y_j), R(x_{i+1}, y_j)]\!] = 0,$$

so we have an RTT equation, and we can build solvable lattice models. In conclusion, all we need in order to build solvable lattice models is the homogenous parametrized Yang–Baxter equation, and these will give examples of the theory in the last subsection.

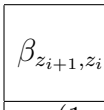
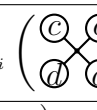
Example 6.35. Let $R = R_\circ$ be the spherical R-matrix from Section 4 of Chapter 5, the Boltzmann weights $R(x_i, y_j)$ at the T-vertex in the i-th row and j-th column are:

T-vertices for the Spherical Model			
			
$x_i - qy_j$	$x_i - qy_j$	$(1-q)x_i$	$(1-q)x_i$ if $c < d$ $(1-q)y_j$ if $c > d$
			
$(1-q)y_j$	$q(x_i - y_j)$	$x_i - y_j$	$x_i - y_j$ if $c > d$ $q(x_i - y_j)$ if $c < d$

We will show that the operator $\mathfrak{T}_i$ in Theorem 6.34 is the Demazure–Lusztig operator $\mathcal{L}_i$.

In this section, we give five examples of homogeneous parametrized colored Yang–Baxter equations that are viatic and orderly. For each of these we compute the operators $\mathfrak{T}_i$ in Theorem 6.34. For definiteness we use the operators $\mathfrak{T}_i = \mathfrak{T}_i^+$, though if we want $\mathfrak{T}_i^-$ we could obtain those also by varying the boundary conditions in Convention 6.28.

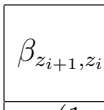
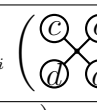
Example 6.36. Consider the spherical model, whose T-weights were considered in Example 6.35. The R-matrix from Chapter 5 gives the following values for $c^+(z_{i+1}, z_i)$, $a(x_{i+1}, x_i)$ and $b^+(x_{i+1}, x_i)$:

β_{z_{i+1},z_i}	β_{z_{i+1},z_i} 	β_{z_{i+1},z_i} 
$(1-q)z_{i+1}$	$(z_{i+1} - qz_i)$	$z_{i+1} - z_i$

Now with $\mathbf{z}^{\alpha_i} = z_i/z_{i+1}$ we can multiply the numerator and denominator of (3.2) by z_{i+1}^{-1} and obtain

$$\mathfrak{T}_i^+ = (\mathbf{z}^{-1} - 1)^{-1}(1 - s_i - q(1 - \mathbf{z}^{\alpha_i} s_i)(\mathbf{z}, w) = \mathcal{L}_i.$$

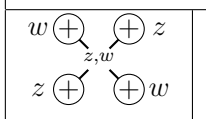
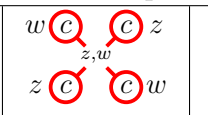
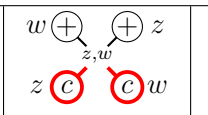
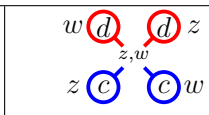
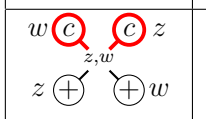
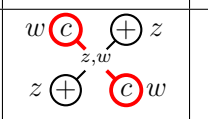
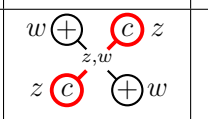
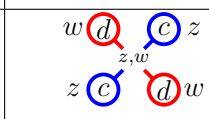
Example 6.37. Next let $R = R_\bullet$ be the antispherical R-matrix from Chapter 5. Here are the relevant Boltzmann weights:

β_{z_{i+1},z_i}	β_{z_{i+1},z_i} 	β_{z_{i+1},z_i} 
$(1-q)z_{i+1}$	$(z_i - qz_{i+1})$	$z_{i+1} - z_i$

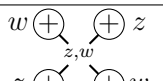
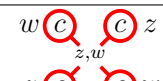
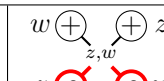
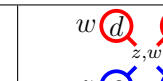
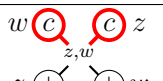
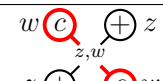
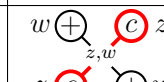
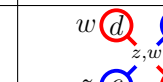
For this example

$$\mathfrak{T}_i = (\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - \mathbf{z}^{\alpha_i} s_i - q(1 - s_i)) = {}^{\rho}\mathcal{T}_i.$$

Next we come to two R-matrices that are very closely related to each other. These will be put to use in Section 4. We call them the *open* and *closed* R-matrices. The open R-matrix is obtained by specializing $q = 0$ in the antispherical model. Here is the R-matrix:

Open R-matrix			
			
w	z	z	z if $c < d$ w if $c > d$
			
w	0	$z - w$	$z - w$ if $c > d$ 0 if $c < d$

The closed R-matrix is nearly identical but differs only in one case (the last case in the first row):

Closed R-matrix			
			
w	z	z	w if $c < d$ z if $c > d$
			
w	0	$z - w$	$z - w$ if $c > d$ 0 if $c < d$

Like the open R-matrix, this is a valid homogeneous parametrized Yang–Baxter equation. However unlike the open R-matrix, the closed R-matrix is not (as far as we know) a specialization of a more general R-matrix depending on q.

Example 6.38. Let us compute the operator for the open model. With $c > d$ we have

$c^+(z_{i+1}, x_i)$	$a(z_{i+1}, z_i)$	$b^+(z_{i+1}, z_i)$
z_{i+1}	z_i	$z_{i+1} - z_i$

Thus (3.2) gives

$$\mathfrak{T}_i = (z_i - z_{i+1})^{-1}(z_{i+1} - z_i s_i) = (\mathbf{z}^{\alpha_i} - 1)^{-1}(1 - \mathbf{z}^{\alpha_i} s_i) = {}^{\rho}\partial_i^{\circ}.$$

The operator $\mathfrak{T}'$ is not defined since $c^-(z_{i+1}, z_i) = 0$, and indeed the operator ∂_i° is not invertible.

Example 6.39. The operator for the closed model is similar, though we will see in Chapter 7 that the subtle difference is important.

$c^+(z_{i+1}, z_i)$	$a(z_{i+1}, z_i)$	$b^+(z_{i+1}, z_i)$
z_i	z_i	$z_{i+1} - z_i$

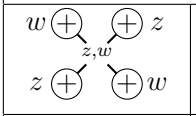
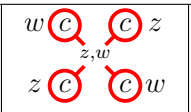
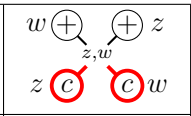
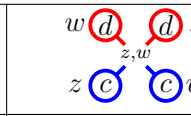
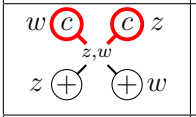
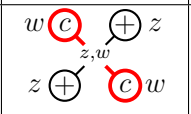
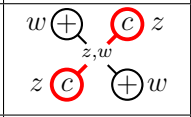
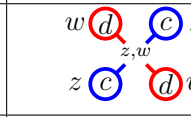
Thus (3.2) gives

$$\mathfrak{T}_i = (z_i - z_{i+1})^{-1}(z_i - z_i s_i) = (1 - \mathbf{z}^{-\alpha_i})^{-1}(1 - s_i) = {}^{\rho}\partial_i.$$

As with the operator $\mathfrak{T}'$ is not defined since $c^-(z_{i+1}, z_i) = 0$, and the operator ∂_i° is also not invertible.

Example 6.40. This model is called the *classical pipedream model*, and we will make further use of this homogeneous parametrized Yang–Baxter equation in Chapter 7. For classical

pipedreams we use the following R-matrix:

Classic Pipedreams R-matrix			
			
1	1	1	1
			
1	0	$z-w$	$z-w$ if $c>d$ 0 if $c<d$

These satisfy a parametrized Yang–Baxter equation (see Exercise 6.3). For consistency with Chapter 7, we use $\mathbf{x}$ instead of $\mathbf{z}$ in this example. Let us compute the corresponding Demazure operator. We have

$c^+(x_{i+1},x_i)$	$a(x_{i+1},x_i)$	$b^+(x_{i+1},x_i)$
1	1	$x_{i+1}-x_i$

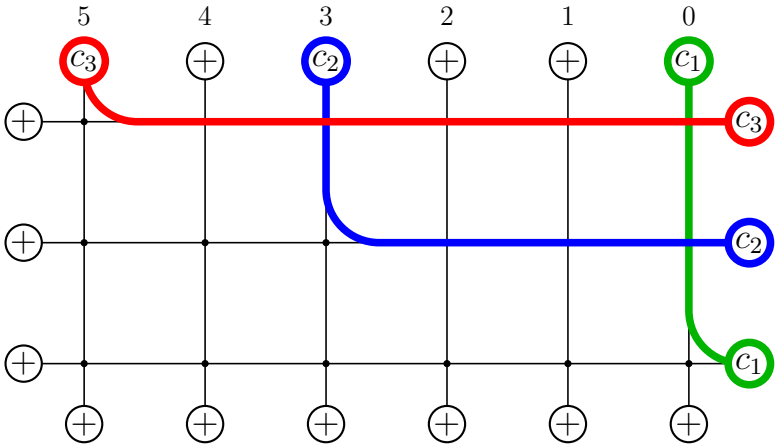
Then

$$\mathfrak{T}_i = (x_i - x_{i+1})^{-1}(1-s_i) = D_i.$$

3.4. Monostatic systems. We recall that a system is called *monostatic* if it has only one state. Let $\mathbf{d}$ and $\mathbf{f}$ be the banners describing the top and right boundary colors.

Proposition 6.41. *If* $\mathbf{d}=\mathbf{f}$ *then the colored extended wall system is monostatic.*

Proof. Since $d_1 = f_1$, the leftmost color on the top row must end up at the right boundary of the top row, and there is only one path connecting these two boundary edges. With this path fixed, there is only one possible path for the color $d_2 = f_2$ and the state looks like this:



(6.21)

□

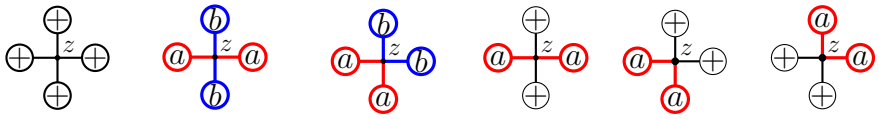
Thus with $\mathbf{d}$ fixed, we have a family of systems $\mathfrak{S}_w$ in which one system is monostatic, and the others are related to it by Demazure recursions. From this information, the partition functions can be computed. Since $\mathbf{f} = w\mathbf{c}_0$, the $w \in W$ for which $\mathfrak{S}_w$ is monostatic is the one for which $\mathbf{d} = w\mathbf{c}_0$.

See Blum 2025 for an interesting and more nuanced situation where there are many monostatic systems, and the partition functions of other systems can be deduced from the monostatic ones using Demazure-like relations coming from the Yang–Baxter equation.

4. The open and closed models

We will apply these principles to the colored versions of the $q = 0$ Tokuyama models called the *open* and *closed* models. The partition functions are, respectively, Demazure atoms and Demazure characters (times $\mathbf{z}^\rho$).

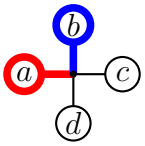
For the open models we use the following Boltzmann weights:



1	z $a \geqslant b$ 0 $a < b$	0 $a > b$ z $a < b$	z	z	1

Then the Yang–Baxter equation is satisfied with the R-matrix given in the previous section. Our convention is that the top boundary banner $\mathbf{d} = w_0\mathbf{c}_0$ and that the right boundary banner is $\mathbf{f} = ww_0\mathbf{c}_0 = w\mathbf{d}$ for $w \in W$. We denote by $\mathfrak{S}^\circ_{\lambda,w} = \mathfrak{S}^\circ_{\lambda,w}(\mathbf{z})$ the system described by these Boltzmann weights and these boundary conditions.

We note that two paths of distinct colors are only allowed to cross in one direction. That is, if a and b are distinct colors, and if paths of the colors a and b meet there, thus:



(6.22)

If $a > b$ then the paths *must* cross at the vertex. On the other hand, if $a < b$ they *may not cross* and instead the paths touch without crossing.

With this in mind, consider what happens when a pair of paths meet several times. The boundary conditions, illustrated in (6.17) guarantee that the first time they meet the color entering the vertex is the larger color. They must therefore cross the first time they meet. The second time they meet, however, the color entering the vertex is the smaller color, and they may not cross. This means that if the two paths meet several times, the first time they meet they will cross, and never again.

Proposition 6.42. *The partition function*

$$Z(\mathfrak{S}^\circ_{\lambda,w}) = \mathbf{z}^\rho \partial^\circ_w \mathbf{z}^\lambda. \tag{6.23}$$

Proof. We will prove this by induction on the length $\ell(w)$. If $w = 1_W$, then the system is monostatic; see (6.21) for the ground state. It is easy to see that there are $\lambda_i + n - i$ vertices that contribute z_i in the i-th row, and thus $Z(\mathfrak{S}^\circ_{\lambda,1}) = \mathbf{z}^\rho$. Now suppose that (6.24) is known for some w, and let $s_i w > w$. By Theorem 6.34 we have

$$Z(\mathfrak{S}^\circ_{\lambda,s_i w}) = \mathfrak{T}^+_i Z(\mathfrak{S}^\circ_{\lambda,w})$$

where by Example 6.38 the operator $\mathfrak{T}^+_i = {}^\rho\partial^\circ_i = \mathbf{z}^\rho \partial^\circ_i \mathbf{z}^{-\rho}$. So

$$Z(\mathfrak{S}^\circ_{\lambda,s_i w}) = \mathbf{z}^\rho \partial^\circ_i \partial^\circ_w \mathbf{z}^\lambda = \mathbf{z}^\rho \partial^\circ_{s_i w} \mathbf{z}^\lambda$$

and (6.24) follows for $s_i w$. □

Now let us consider the question of embedding $\mathfrak{S}^\circ_{\lambda,w}$ into $\mathcal{B}_\lambda$ by adapting (4.15). Given a state $\mathfrak{s}$, we construct a Gelfand–Tsetlin pattern $\mathrm{GTP}(\mathfrak{s})$ whose entries are the columns of the vertical edges carrying a color. This Gelfand–Tsetlin pattern is left strict, as in the $q = 0$ Tokuyama model, so we may again subtract P and write $\mathrm{GTP}^\circ(\mathfrak{s}) = \mathrm{GTP}(\mathfrak{s}) - P$. Then we may again define a map θ_w by

$$\theta_w(\mathfrak{s}) = \mathrm{Sch}(\mathrm{SSYT}(\mathrm{GTP}^\circ(\mathfrak{s})))$$

exactly as in (4.15). Thus $\theta_w : \mathfrak{S}^\circ_{\lambda,w} \longrightarrow \mathcal{B}_\lambda$ is defined. It may be checked that $\beta(\mathfrak{s}) = \mathbf{z}^{\mathrm{wt}(\theta(\mathfrak{s}))}$.

There is one point that we must consider, which is that $\mathrm{GTP}^\circ(\mathfrak{s})$ does not record the color of the edges. Still, we have:

Proposition 6.43. *The maps θ_w are injective. The crystal $\mathcal{B}_\lambda$ is the disjoint union of the images of the maps θ_w.*

Proof. The open Boltzmann weights are "deterministic" in the following sense. Suppose that two colored paths meet at a vertex with input colors a and b as in (6.22). If $a > b$, then the paths must cross at the vertex, meaning that $c = a$ and $d = b$. This is because otherwise, the Boltzmann weight is zero, from the table of open weights. With this in mind, we can reconstruct the colors of all edges from just the information in $\mathrm{GTP}(\mathfrak{s})$, beginning at the top left and moving to the right, then downward row by row. When we are finished, we can read off the banner of right boundary spins, finding that the colored path is is in the image $\theta_w(\mathfrak{S}_{\lambda,w})$ for a unique w. Both statements are clear from these considerations. □

If $\mathfrak{S}_\lambda(\mathbf{z})$ is the $q = 0$ Tokuyama model, this implies that

$$Z(\mathfrak{S}_\lambda(\mathbf{z})) = \sum_w Z(\mathfrak{S}_{\lambda,w}(\mathbf{z})).$$

Using Proposition 6.42 and Theorem 4.18, we obtain

$$s_\lambda(\mathbf{z}) = \sum_{w \in W} \chi^\circ_\lambda(\mathbf{z}; w).$$

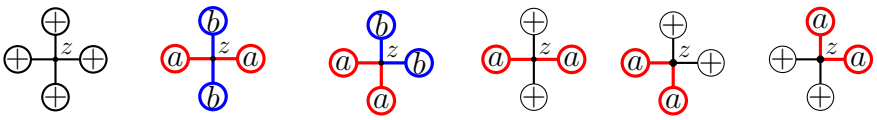
This is the special case of (6.12) where $w = w_0$.

These results can be lifted to the crystal.

Theorem 6.44. *The image of θ_w is the crystal Demazure atom $\mathcal{B}^\circ_\lambda(w)$.*

Proof. Different proofs of this may be found in Brubaker, Buciumas, Bump, and Gustafsson 2021; Yang 2025. □

We now turn to the *closed models*. The Boltzmann weights are very similar to the open models with a minor change:



1	z $a \geqslant b$ 0 $a < b$	z $a > b$ 0 $a < b$	z	z	1

The boundary conditions will be the same as the open model, but there are some interesting differences. We will denote the closed system as $\mathfrak{S}^\bullet_{\lambda,w}$.

Proposition 6.45. *The partition function*

$$Z(\mathfrak{S}^\bullet_{\lambda,w}) = \mathbf{z}^\rho \partial^\bullet_w \mathbf{z}^\lambda. \tag{6.24}$$

Proof. The structure of the proof is the same as Proposition 6.42. The R-matrix is given in Subsection 3.3, and by Theorem 6.34 we have

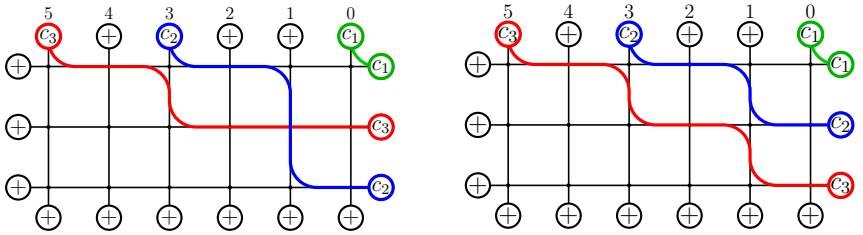
$$Z(\mathfrak{S}^\bullet_{\lambda,s_i w}) = \mathfrak{T}_i^+ Z(\mathfrak{S}^\bullet_{\lambda,w})$$

where by Example 6.39 the operator $\mathfrak{T}_i^+ = {}^\rho\partial_i = \mathbf{z}^\rho \partial_i \mathbf{z}^{-\rho}$. The remaining details are as in Proposition 6.42. □

The closed weights are not deterministic in the same sense that the open weights are. Referring to the configuration (6.22), we see that if $a > b$ then both $c = a$ and $d = b$ or $c = b$ and $d = a$ are legal weights. In other words, if two paths meet at a vertex with the larger color on the left, the paths are free to either cross or not cross.

On the other hand, if $a < b$ then there are no possibilities for c and d. Thus the paths *may not meet* with the smaller color on the left.

These observations have the following consequences. Two paths of different colors may meet several times, and they are not obligated to cross. However if they do cross, the crossing must be the last time they meet, since if they cross earlier, the next time they meet, the smaller color will be on the left, and there will be no legal state. Consider the following two cases.



Both are legal states for closed models $\mathfrak{S}^\bullet_{\lambda,s_1s_2}$ and $\mathfrak{S}^\bullet_{\lambda,w_0}$, with $\lambda = (3,2,0)$. The red and blue lines meet twice, and either cross the last time they meet, or do not cross. Note that these states correspond to the same Gelfand–Tsetlin pattern, showing that the Gelfand–Tsetlin pattern does not determine the state; this reflects the fact that the weights are not locally deterministic.

Yet if we fix w in $\mathfrak{S}^\bullet_{\lambda,w}$, the Gelfand–Tsetlin does determine the state. This does not contradict the last example, because there the two states correspond to different values of w. So there is still an injective map $\theta_w : \mathfrak{S}^\bullet_{\lambda,w} \longrightarrow \mathcal{B}_\lambda$.

Theorem 6.46 (Yang)**.** *The image $\theta_w(\mathfrak{S}^\bullet_{\lambda,w})$ is the Demazure crystal $\mathcal{B}_\lambda(w)$.*

Proof. See Yang 2025 for a proof. □

5. Color merging

Let $\mathfrak{S}_\lambda(\mathbf{z})$ be the $q = 0$ Tokuyama model. In this section we will reprove (6.12) in the form

$$Z(\mathfrak{S}_\lambda) = \sum_{w\in W} Z(\mathfrak{S}^\circ_{\lambda,w}). \tag{6.25}$$

The proof in this section is based on what we call *color merging*. It is called *color blindness* by Borodin and Wheeler. See Aggarwal, Borodin, and Wheeler 2023a; Borodin and Wheeler 2022; Brubaker, Buciumas, Bump, and Gustafsson 2019; Bump and Naprienko 2022 for other examples. Not all Boltzmann weights have this property, but when they do there are consequences of which (6.25) is a typical example.

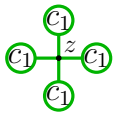
We begin with the observation that both the open and closed models contain the $q = 0$ Tokuyama weights as special cases. That is, if we fix one color c (say $c = c_1$) and restrict the palette to just $\mathcal{P}_1 = \{\oplus, c_1\}$, we recover the $q = 0$ Tokuyama weights with c playing the role of the odd spin $\ominus$. With this in mind, the terms on *both* sides of (6.25) can be regarded as partition functions of the open model. (Earlier in this chapter, we did not allow banners with repeated colors, but in this section we do allow repeated colors on the top and right boundaries.)

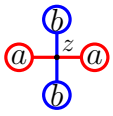
Now we will formulate a more subtle *local lifting property* of the open weights. This will allow us to eliminate one color of the palette $\mathcal{P}_r = \{\oplus, c_1, \dots, c_r\}$ from the boundary conditions, and indeed from all states of the model. Let $\sigma : \mathcal{P}_r \to \mathcal{P}_1$ be the "retraction" map defined by $\sigma(c_i) = c_1$ while $\sigma(\oplus) = \oplus$.

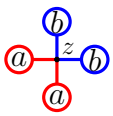
Proposition 6.47 (Local Lifting Property)**.** *Let $a, b \in \mathcal{P}_r$ and let $a', b', c', d' \in \mathcal{P}_1$ such that $\sigma(a) = a'$ and $\sigma(b) = b'$. Then using the weights of the open model,*

$$\beta\left(\begin{array}{ccc} & b' & \\ a' & z & c' \\ & d' & \end{array}\right) = \sum_{\substack{c,d\in\mathcal{P}_n \\ \sigma(c)=c',\sigma(d)=d'}} \beta\left(\begin{array}{ccc} & b & \\ a & z & c \\ & d & \end{array}\right) \tag{6.26}$$

Proof. This can be checked easily by examining the Boltzmann weights. The most interesting case is where a and b are distinct colors, so $a' = b' = c_1$ are not distinct. Now the Boltzman weights are as follows:







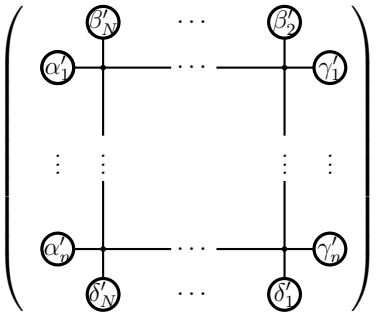
z	z	$a > b$	0	$a > b$
	0	$a < b$	z	$a < b$

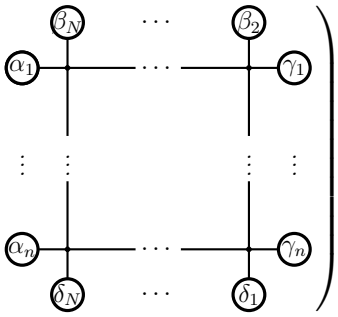
Both sides of the identity (6.26) equal z in this case, whether $a > b$ or $a < b$. All other cases are straightforward. □

The statement of the local property may seem slightly technical, but it is actually very natural, and the example in the proof should clarify its meaning. It has a global generalization, and its significance is that the local property implies the global one.

Theorem 6.48 (Global Lifting Property)**.** *Let $\alpha_i, \beta_i \in \mathcal{P}_n$ and $\alpha'_i, \beta'_i, \gamma'_i, \delta'_i \in \mathcal{P}_1$ be spins such that $\sigma(\alpha_i) = \alpha'_i$ and $\sigma(\beta_i) = \beta'_i$. Then*

(6.27)
$$Z\left(\quad\right) = \sum_{\substack{\gamma_i,\delta_j \\ \sigma(\gamma_i)=\gamma'_i \\ \sigma(\delta_i)=\delta'_i}} Z\left(\quad\right)$$

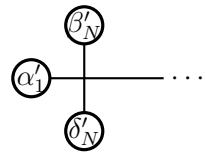


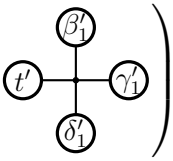


where the summation is over γ_i and δ_i such that $\sigma(\gamma_i) = \gamma'_i$ and $\sigma(\delta_i) = \delta'_i$.

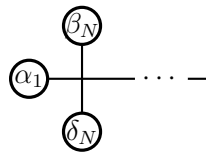
Proof. First we will prove this if $n = 1$, so there is only one row. If also $N = 1$ this is exactly the local lifting property, so assume by induction that $N > 1$ and that the result is true for $N - 1$ columns. Using the definition of the partition function as a sum over the spins of interior edges, and then using the induction hypothesis, the left-hand side in (6.27) equals

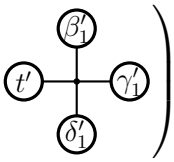
$$\sum_{t' \in \mathcal{P}_1} Z\left(\quad\right) \beta\left(\quad\right) =$$





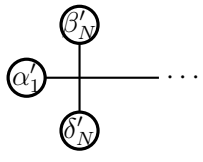
$$\sum_{t' \in \mathcal{P}_1} \sum_{\substack{t,\delta_2,\cdots,\delta_N \in \mathcal{P}_n \\ \sigma(t)=t',\sigma(\delta_i)=\delta'_i}} Z\left(\quad\right) \beta\left(\quad\right)$$

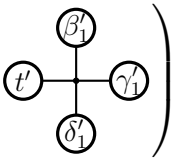




Now interchanging the order of summation gives:

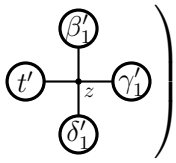
$$\sum_{\substack{\delta_2,\cdots,\delta_N \in \mathcal{P}_n \\ \sigma(\delta_i)=\delta'_i}} \sum_{\substack{t \in \mathcal{P}_n \\ t'=\sigma(t)}} Z\left(\quad\right) \beta\left(\quad\right)$$

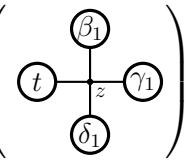




after using the local lifting property in the form

$$\beta\left(\quad\right) = \sum_{\substack{\gamma_1,\delta_1 \in \mathcal{P}_n \\ \sigma(\gamma_1)=\gamma'_1,\sigma(\delta_1)=\delta'_1}} \beta\left(\quad\right)$$





we may reassemble this to obtain the right-hand side of (6.27), which is then proved if the number of rows is 1.

Now very similar arguments go from one row to n. We leave this step to the reader. □

Remark 6.49. Applying this to the open weights and the colored extended wall boundary conditions gives (6.25).

Exercises

Exercise 6.1. Prove the second identity in (6.13).

Exercise 6.2. The goal of this exercise is to prove Proposition 6.33. As part of the definition of a homogeneous parametrized Yang–Baxter equation, it is assumed that $R(x_{i+1}, x_i)R(x_i, x_{i+1})$ is a constant multiple of the identity matrix. Show that this implies identities

$$\begin{aligned} &a(x_i, x_{i+1})a(x_{i+1}, x_i) = \\ &b(x_i, x_{i+1}; d, c)b(x_{i+1}, x_i; c, d) + c(x_i, x_{i+1}; c, d)c(x_{i+1}, x_i; c, d), \end{aligned}$$

$$c(x_i, x_{i+1}; d, c)b(x_{i+1}, x_i; c, d) + b(x_i, x_{i+1}; c, d)c(x_{i+1}, x_i; c, d) = 0,$$

and use these to prove that $\mathfrak{T}_i^+$ and $\mathfrak{T}_i^-$ are inverse operators.

Exercise 6.3. Show that the classic pipedreams R-matrix satisfies a homogeneous parametrized Yang–Baxter equation.

Hints: Note that including $\oplus$ as a color smaller than all the other colors the Boltzmann weights remain orderly, that is, the Boltzmann weights only depend on the order of the colors. Deduce that only three colors (possibly including $\oplus$) can occur in any configuration of the Yang–Baxter equation, resulting in a relatively small number of cases that can be checked by hand or with a computer.

CHAPTER 7

Schubert Polynomials

Schubert polynomials originate in the cohomology of flag varieties. Grothendieck polynomials are generalizations related to K-theory. These geometric polynomials have descriptions called *pipedreams* that fit nicely into the theory of solvable lattice models. In this chapter, we discuss these pipedream models, and related crystal structures on pipedreams.

1. Schubert polynomials: Definition and history

Schubert polynomials constitute a family of multivariate polynomials that represent cohomology classes of Schubert varieties in the flag manifold, playing a central role in algebraic geometry, combinatorics, and representation theory. Introduced by Lascoux and Schützenberger 1982, they generalize the concept of Schur polynomials (see Chapter 4) and provide a combinatorial tool for understanding the geometry of flag varieties. Schubert polynomials are labeled by permutations and encode intersection numbers and degeneracy loci in terms of polynomial expressions. Their rich algebraic structure connects to symmetric functions, Demazure operators (see Chapter 6), and the geometry of Grassmannians, making them a key object in modern algebraic combinatorics.

1.1. Geometry. A *complete flag* in $V = \mathbb{C}^n$ is a sequence of subspaces

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{n-1} \subset V_n = V,$$

such that $\dim V_i = i$ for all $i = 0, 1, \ldots, n$. Let $\mathcal{F}\ell(n)$ denote the variety of complete flags in V.

Let $G = GL_n(\mathbb{C})$ and B be the Borel subgroup of upper triangular matrices in G. Then G/B is the vector space of formal linear combinations of cosets of B with the action of G given by left multiplication. We can pick the matrices with the following property as the coset representatives of the cosets in G/B. The rightmost nonzero entry in each row is 1 and is the first nonzero entry in its column. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of $\mathbb{C}^n$, where $\mathbf{e}_i$ has 1 in position i and zeros everywhere else. The Borel subgroup B stabilizes the *standard flag*

$$E_\bullet = (\{0\} = E_0 \subset E_1 \subset \cdots \subset E_n = \mathbb{C}^n), \quad \text{with } E_i = \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_i\}.$$

There is a left action of G on $\mathcal{F}\ell(n)$ given by

$$g\,(V_0 \subset V_1 \subset \cdots \subset V_n) = (gV_0 \subset gV_1 \subset \cdots \subset gV_n) \qquad \text{for all } g \in G.$$

Hence $\mathcal{F}\ell(n)$ can be identified with the quotient G/B via

$$G/B \leftrightarrow \mathcal{F}\ell(n)$$
$$gB \leftrightarrow gE_\bullet.$$

Fix a reference flag

$$F_\bullet = (\{0\} = F_0 \subset F_1 \subset \cdots \subset F_n = V), \quad \text{with } \dim F_i = i.$$

For each permutation $w \in S_n$, the corresponding *Schubert cell* $X_w^\circ \subset \mathcal{F}\ell(n)$ is defined as

$$X_w^\circ = \left\{ \tilde{F}_\bullet \in \mathcal{F}\ell(n) \;\middle|\; \dim(\tilde{F}_i \cap F_j) = \#\{k \leqslant i \mid w(k) \leqslant j\} \text{ for all } i, j \right\}.$$

Their Zariski closures are called *Schubert varieties*, denoted by X_w. Schubert varieties are intimately related to the Bruhat order introduced in Chapter 6, Section 1.3. In particular, $X_v \subseteq X_w$ if and only if $v \leqslant w$ in Bruhat order and

$$X_w = \bigcup_{v \leqslant w} X_v^\circ.$$

In the bijection of $\mathcal{F}\ell(n)$ with G/B, the open Schubert cell X_w° corresponds to the double coset BwB/B. It is an affine space of complex dimension $\ell(w)$, or real dimension $2\ell(w)$. This gives a cellular decomposition of $\mathcal{F}\ell(n)$ showing that the Schubert classes $[X_w]$ are a basis of the cohomology ring. Since they are algebraic, the cohomology ring coincides with the Chow ring of algebraic cycles. Because there is cohomology only in even dimensions, $H^*(\mathcal{F}\ell(n))$ is therefore commutative. See Fulton 1997 and Brion 2005 for further details.

Under the *Borel isomorphism* (see Borel 1953) the cohomology of G/B is identified with the polynomial ring in n variables

$$H^*(G/B) \cong \mathbb{Z}[x_1, \ldots, x_n]/\mathcal{I}^{S_n}$$

modulo the ideal $\mathcal{I}^{S_n}$ generated by the nonconstant symmetric polynomials. The *Schubert polynomials* $\mathfrak{S}_w$ indexed by $w \in S_n$ are polynomials in $\mathbb{Z}[x_1, \ldots, x_n]$ which map to Schubert classes under the surjective ring homomorphism

$$\begin{aligned} \mathbb{Z}[x_1, \ldots, x_n] &\to H^*(G/B) \\ \mathfrak{S}_w &\mapsto [X_{w_0 w}], \end{aligned}$$

where $w_0 = n\ n-1\ \ldots 1$ is the long permutation and $(w_0 w)(i) = n+1-w(i)$. The Schubert polynomials play the analogous role in the cohomology of flag varieties as Schur polynomials play in the cohomology of the Grassmannian.

1.2. Recursive definition. Recall the *divided difference operator* D_i from Chapter 6, Section 3 defined on polynomials $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ as follows. Let $s_i = (i, i+1)$ be the adjacent transposition in the symmetric group S_n. Then

$$D_i f(x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n) - s_i f(x_1, \ldots, x_n)}{x_i - x_{i+1}},$$

where $s_i f(x_1, \ldots, x_n)$ is the polynomial obtained by swapping x_i and x_{i+1} in $f(x_1, \ldots, x_n)$.

The *Schubert polynomial* $\mathfrak{S}_w := \mathfrak{S}_w(x_1, \ldots, x_n)$ corresponding to a permutation $w \in S_n$ is defined recursively as follows:

(1) For the longest permutation $w_0 = n\,(n-1) \cdots 1$ in one-line notation, define

$$\mathfrak{S}_{w_0} := \mathbf{x}^\rho = x_1^{n-1} x_2^{n-2} \cdots x_{n-1},$$

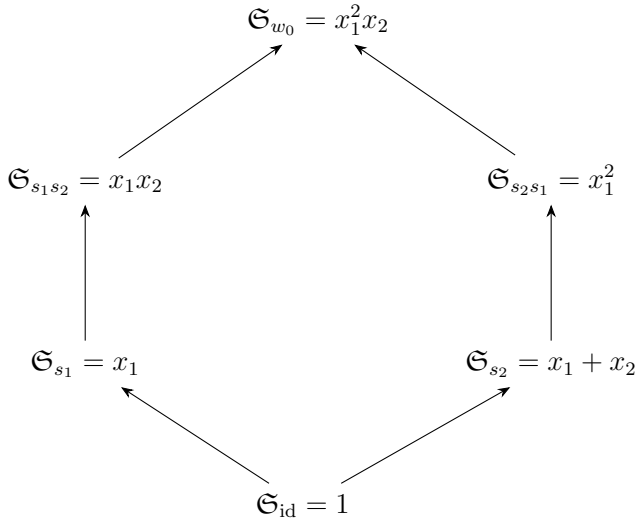
where $\rho = (n-1, n-2, \ldots, 1, 0)$ as in (6.3).

(2) For any $w \in S_n$ and s_i such that $\ell(ws_i) < \ell(w)$ (i.e., s_i reduces the length), define

$$\mathfrak{S}_{ws_i} = D_i \mathfrak{S}_w. \tag{7.1}$$

This recursive definition produces a well-defined polynomial $\mathfrak{S}_w$ for every permutation $w \in S_n$ by the Matsumoto theorem (see Section 1.1) since the divided difference operators obey the same braid relations as the symmetric group.

Example 7.1. Let S_3 denote the symmetric group on 3 elements. The Schubert polynomials $\mathfrak{S}_w$ for $w \in S_3$ are given as follows, where the permutations are given by their reduced expressions organized in weak order:



A generalization called *double Schubert polynomials* were introduced by Macdonald 1991a,b. Double Schubert polynomials are polynomials in two sets of variables $\mathfrak{S}_w(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)$ indexed by a permutation $w \in S_n$, which become the usual Schubert polynomials when all the variables y_i are 0. The double Schubert polynomial $\mathfrak{S}_w(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)$ can also be defined recursively as follows:

(1) We have

$$\mathfrak{S}_{w_0}(x_1, x_2, \ldots x_n, y_1, y_2, \ldots, y_n) = \prod_{i+j \leqslant n} (x_i - y_j),$$

where w_0 is the long element in S_n.

(2) For any $w \in S_n$ and s_i such that $\ell(ws_i) < \ell(w)$ (or alternatively $w(i) > w(i+1)$), define

$$D_i \mathfrak{S}_w = \mathfrak{S}_{ws_i},$$

where the D_i operator acts on the $\mathbf{x}$ variables.

There is a formula which expresses the double Schubert polynomials in terms of Schubert polynomials

$$\mathfrak{S}_w(\mathbf{x}, \mathbf{y}) = \sum_{\substack{w = v^{-1}u \\ \ell(w) = \ell(u) + \ell(v)}} \mathfrak{S}_u(\mathbf{x}) \mathfrak{S}_v(-\mathbf{y}). \tag{7.2}$$

For a proof see Macdonald 1991a, equation (6.3).

2. Combinatorics of Schubert polynomials

In the early 1990s, Stanley was the first to conjecture the monomial expansion for the Schubert polynomial $\mathfrak{S}_w$ in terms of compatible sequences. Compatible sequences appeared in his study of the number of reduced expressions for w_0 in his celebrated paper Stanley 1984. Billey, Jockusch, and Stanley 1993 proved the monomial formula using recurrences. Fomin and Stanley 1994 gave a different proof using nil-Coxeter algebras. An analysis in terms

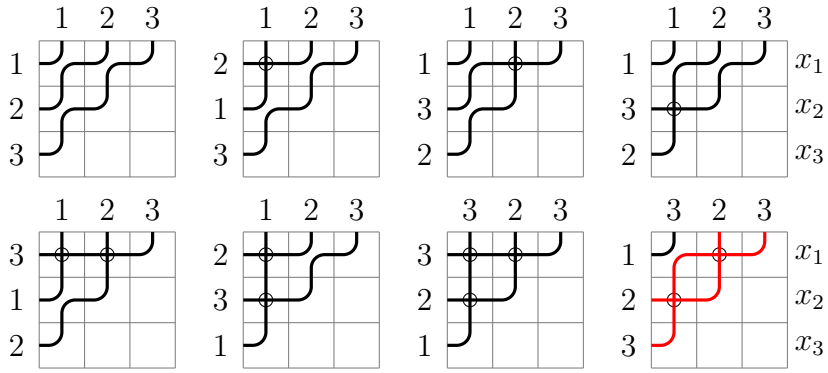
of the Yang–Baxter equation was given by Fomin and Kirillov 1996. Combinatorial models using reduced pipedreams was first given by Bergeron and Billey 1993 (though they called them rc-graphs). The term *pipedreams* was coined by Knutson and Miller 2005. We will discuss another more recent model in terms of reduced bumpless pipedreams by Lam, Lee, and Shimozono 2021 in the next section. For this reason, we use the terminology *classical pipedreams* for the pipedreams of Bergeron and Billey 1993; Knutson and Miller 2005.

2.1. Classical pipedreams. Classical pipedreams are configurations of lines called *pipes* through an $n \times n$ grid, that may sometimes cross. In our convention the pipes connect to the top and left edges of the grid. Pipes move down and to the left from the top. They may cross or bend around each other. The following configurations are allowed:

These may be thought of as tiles that may be assembled to make the pipedream. The first tile is called the "crossing" and the second the "bump." The last three partially empty two tiles are needed because there will be regions in the grid that do not contain any pipes. For boundary conditions, we assume that every top boundary edge is connected to a left boundary edge by a pipe but that no pipes exit to the right or bottom. These boundary conditions actually rule out the last tile, which doesn't occur but which should be allowed for reasons related to the Yang-Baxter equation.

Here are the possible classic pipedreams for $n = 3$:

(7.3)



We may associate a permutation $w(p)$ with a pipedream p. To define this, we label the columns 1 to n from left to right; then the pipes carry these numbers to the left boundary, where we can read off the permutation in one-line notation. In the $n = 3$ pipedreams, for example, the permutations in one-line notation are $(123) = 1_W$, $(213) = s_1$, $(132) = s_2$ and $(132) = s_2$ for the top row, and $(312) = s_2 s_1$, $(231) = s_1 s_2$, $(321) = w_0$ and $(123) = 1_W$ for the second row.

A pipe is called *reduced* if no pair of lines crosses more than once. For $w \in W$ let $\mathcal{RP}(w)$ be the set of reduced pipes p with $w(p) = w$. The last pipedream is nonreduced since the two red lines cross twice. We are mainly interested in the reduced pipedreams.

If p is a reduced pipedream let $\mathbf{x}^{\mathsf{wt}(p)}$ be the monomial in $x_1, \cdots, x_n$ in which the exponent of x_i is the number of crossings in the i-th row. In the examples, we've circled the crossings, from which the weights of the pipedreams can be seen to be 1, x_1, x_1 and x_2 for the first row, and x_1^2, $x_1 x_2$ and $x_1^2 x_2$ for the first three patterns of the second row. (We are excluding the nonreduced pattern.)

Later in this section we will prove that

$$\sum_{p \in \mathcal{RP}(w)} \mathbf{x}^{\mathsf{wt}(p)} = \mathfrak{S}_w(\mathbf{x}). \tag{7.4}$$

This can be confirmed for the S_3 Weyl group by comparing the values from the pipedreams with Example 7.1.

Example 7.2. If $w = s_2 = (132)$ in one-line notation, then $\mathcal{RP}w$ consists of the last two pipedreams in the top row of (7.3). These have weights x_2 and x_1, consistent with $\mathfrak{S}_{s_2} = x_1 + x_2$. For S_3, this is the only permutation with more than one reduced pipedream.

Equation (7.4) has an easy generalization to double Schubert polynomials. Let us define

$$(\mathbf{x}|\mathbf{y})^{\mathsf{wt}(p)} = \prod_{\text{crossing at } (i,j)} (x_i - y_j).$$

Then

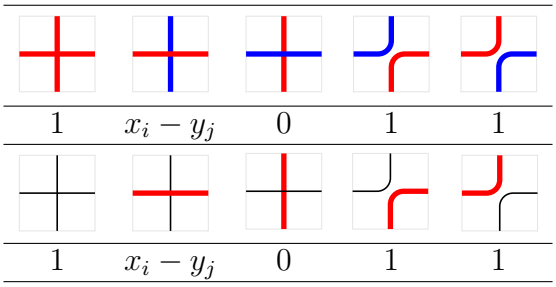
$$\sum_{p \in \mathcal{RP}(w)} (\mathbf{x}|\mathbf{y})^{\mathsf{wt}(p)} = \mathfrak{S}_w(\mathbf{x}; \mathbf{y}). \tag{7.5}$$

Of course (7.4) is the special case $\mathbf{y} = 0$.

Before we prove (7.5), let us explain how the addition of color to the pipes makes pipedreams into solvable lattice models. Thus far we have defined classic reduced pipedreams as they appear in much literature such as Knutson and Miller 2005, but more recently they have been treated as states of solvable lattice models. For example, the models that we will discuss in this section are special cases of the colored models in Brubaker, Frechette, et al. 2023, and the colored lattice model versions of the bumpless pipedreams that we will consider below in Section 2.4 are special cases of lattice models considered in Buciumas and Scrimshaw 2022a.

We choose n colors $c_1, \cdots, c_n$ and an "uncolor" $\oplus$ ordered so that $\oplus < c_1 < \cdots < c_n$. We now describe a model that resembles those of Chapter 6, though differs from those by a mirror reflection (left to right) and also in that the Weyl group element w follows a different convention. We will number the rows of the grid 1 to n from top to bottom, and the columns 1 to n from left to right. We have seen before that the lattice model point of view often gives us the option of introducing column parameters. In this case we will obtain double Schubert polynomials as the partition functions. So let $\mathbf{y} = (y_1, \cdots, y_n)$ be a second set of parameters. We make use of the following Boltzmann weights at the i-th row and the j-th column:

(7.6)



We are using Red and Blue to denote two distinct colors with red(•) > blue(•).

For boundary conditions, we put color c_i at the top boundary in column i, and with $w \in W$ fixed we put color $c_{w(i)}$ at the left boundary in row i. (In Chapter 6, we would put $c_{w^{-1}(i)}$ at the right boundary in row i.) On the right and bottom boundary edges we put $\oplus$.

With our boundary conditions the first pattern does not appear, since there will be no more than one path of any given color. However we include it since it does appear in the Yang-Baxter equation.

An important point is that two colors can only cross in one direction, because the third pattern is given weight zero. A consequence of this is that two colored paths may not cross twice. We have already seen this phenomenon with the open and closed models in Chapter 6. If two paths of colors $c_k \leqslant c_l$ both appear, since the larger color starts on the top boundary to the right of the smaller, a crossing will be legal and will contribute a factor if $x_i - y_j$ to the Boltzmann weight. The second time they meet, the will be in the third pattern in (7.6), and the Boltzmann weight is zero. This is similar to the fact that in the open and closed models, two paths may not cross more than once, for exactly the same kind of reason.

Let $\mathfrak{P}_w(\mathbf{x};\mathbf{y})$ denote the colored model that we have described. We will call it the *colored classical pipedream model*. The terminology is justified by the following observation.

Proposition 7.3. *Let $w \in W$. Then*

$$\sum_{p \in \mathcal{RP}(w)} (\mathbf{x}|\mathbf{y})^{\mathsf{wt}(p)} = Z(\mathfrak{P}_w(\mathbf{x};\mathbf{y})).$$

Proof. Let us consider a pipedream $p \in \mathcal{RP}(w)$. We may obtain a state $\mathbf{s}$ of the model $\mathfrak{P}$ described above by simply coloring the paths. It is clear that

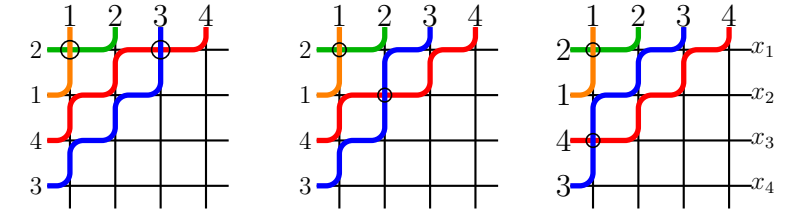
$$(\mathbf{x}|\mathbf{y})^{\mathsf{wt}(p)} = \beta(\mathbf{s})$$

since both sides are products of terms $x_i - y_j$ over the crossing points (i,j). □

Example 7.4. The partition function of $\mathfrak{P}_w$ where $w = s_1 s_3 = 2143 \in S_4$ is

$$(x_1 - y_1)(x_1 - y_3) + (x_1 - y_1)(x_2 - y_2) + (x_1 - y_1)(x_3 - y_1). \tag{7.7}$$

The three terms are associated with the three classical pipedreams:



$(x_1 - y_1)(x_1 - y_3)$ $(x_1 - y_1)(x_2 - y_2)$ $(x_1 - y_1)(x_3 - y_1)$

where

$$\text{red } (\bullet) > \text{blue } (\bullet) > \text{green } (\bullet) > \text{orange } (\bullet).$$

The expression (7.7) equals the double Schubert polynomial $\mathfrak{S}_w$.

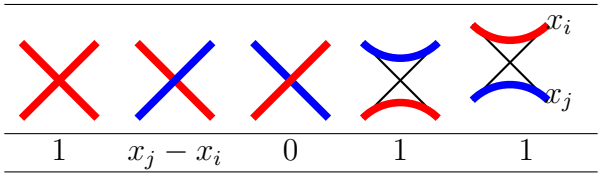
Theorem 7.5. *The partition function of $\mathfrak{P}_w(\mathbf{x};\mathbf{y})$ is the double Schubert polynomial $\mathfrak{S}_w(\mathbf{x};\mathbf{y})$.*

Note that this implies (7.5) and hence (7.4).

Before we can prove this, we will need a Yang-Baxter equation. The R-matrix is essentially the parametrized Yang–Baxter equation from Example 6.40. We use the following

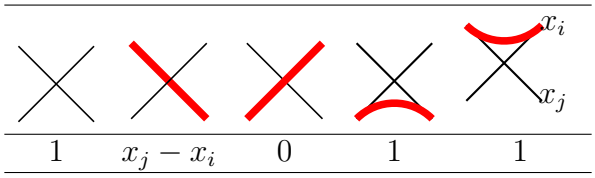
Boltzmann weights, where it is understood that red(•) > blue(•). As with the T-vertices discussed above, we allow the possibility that one color in this chart is the uncolor $\oplus$, which is ordered smaller than the other colors. This gives us a parmetrized Yang-Baxter equation. It is related to the Yang-Baxter equation from Example 6.40, by easy transformations (switch the order of the colors and the signs of z and w).

(7.8)



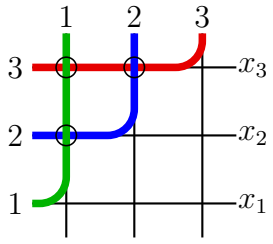
We extend this to all colors, including $\oplus$ which (we recall) is the smallest color.

(7.9)



The boundary conditions for the pipedream models are similar to the colored domain wall boundary conditions, which are the special case of the extended wall boundary conditions of Chapter 6 in which $\lambda = 0$. But the pipedream models are the mirror images of the colored domain wall boundary conditions from Chapter 6.

Proof of Theorem 7.5. For $w = w_0$, there is only one configuration for the classical pipedream model: the largest color goes to row 1, the second largest color goes to row 2 and so on. Since colors can only cross once, this forces all paths. For $n = 3$ the configuration looks as follows:



There are crossings at all positions (i, j) in row i and column j for $j \leqslant n - i$. Hence the weight for this configuration is $\mathbf{x}^\rho = x_1^{n-1} x_2^{n-2} \cdots x_{n-1}$. This shows that $\mathfrak{S}_{w_0}(\mathbf{x}) = Z_{w_0}(\mathbf{x})$.

Computing the Yang–Baxter equation as in the proof of Theorem 6.34 in Chapter 6. Attach the R-matrix (7.8) with (i, j) taken to be $(i + 1, i)$ to the *right* of the grid, run the train argument to obtain

$$Z_w(\mathbf{x}) = (x_i - x_{i+1}) Z_{ws_i}(s_i\mathbf{x}) + Z_w(s_i\mathbf{x}).$$

Replacing $\mathbf{x}$ by $s_i\mathbf{x}$ gives

$$Z_w(s_i\mathbf{x}) = (x_{i+1} - x_i) Z_{ws_i}(\mathbf{x}) + Z_w(\mathbf{x}),$$

which yields the recursion

$$Z_{ws_i}(\mathbf{x}) = D_i Z_w(\mathbf{x}).$$

Comparing with (7.1) shows that both Z_w and $\mathfrak{S}_w$ satisfy the same recursion and hence must be equal. □

2.2. Compatible sequences. Reduced pipedreams can also be encoded as compatible sequences Billey, Jockusch, and Stanley 1993. A *compatible sequence* for a permutation $w \in S_n$ is a pair $(\mathbf{a}, \mathbf{r})$ of sequences $\mathbf{a} = (a_1, \ldots, a_\ell)$ and $\mathbf{r} = (r_1, \ldots, r_\ell)$ such that $\mathbf{a}$ is a reduced word for w and $\mathbf{r}$ is $\mathbf{a}$-compatible. A sequence $\mathbf{r}$ is $\mathbf{a}$-compatible if

$$\begin{aligned} &(1) \quad r_i \leqslant r_j \quad \text{for } 1 \leqslant i < j \leqslant \ell, \\ &(2) \quad r_k \leqslant a_k \quad \text{for all } 1 \leqslant k \leqslant \ell, \\ &(3) \quad a_k < a_{k+1} \quad \text{implies } r_k < r_{k+1} \text{ for all } 1 \leqslant k < \ell. \end{aligned} \tag{7.10}$$

The weight of a compatible sequence $(\mathbf{a}, \mathbf{r})$ is $\mathsf{wt}(\mathbf{a}, \mathbf{r}) = (\mathsf{wt}_1, \ldots, \mathsf{wt}_{n-1})$, where wt_i is the number of letters i in $\mathbf{r}$. Denote by $\mathcal{RC}(w)$ the set of compatible sequences for w.

There is a weight preserving bijection between reduced classical pipedreams and compatible sequences

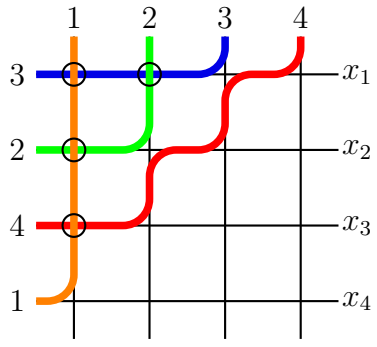
$$\varphi \colon \mathcal{RP}(w) \to \mathcal{RC}(w)$$

defined as follows. Denote by $(i_1, j_1), \ldots, (i_\ell, j_\ell)$ the coordinates of the crosses in the reduced classical pipedream $p \in \mathcal{RP}(w)$, where i_k is the row index and j_k is the column index with rows labeled from top to bottom and columns labeled from left to right. Note that a classical pipedream is uniquely specified by the position of the crosses. Order these coordinates according to the following order. We say $(i, j) < (i', j')$ if either $i < i'$ or $i = i'$ and $j > j'$. Set $\mathbf{r} = (i_1, \ldots, i_\ell)$ and $\mathbf{a} = (a_1, \ldots, a_\ell)$, where $a_k = i_k + j_k - 1$.

Proposition 7.6. *The map φ is well-defined. That is, for $p \in \mathcal{RP}(w)$, we have $\varphi(p) \in \mathcal{RC}(w)$.*

Proof. Note that by definition due to the ordering we have $r_i \leqslant r_j$ for $i < j$ which is condition (1) in (7.10). Also $r_k = i_k \leqslant i_k + j_k - 1 = a_k$ since $j_k \geqslant 1$, so that (2) in (7.10) holds. Assume $r_k = r_{k+1}$. Then $i_k = i_{k+1}$, so that by the ordering we have $j_k > j_{k+1}$. But this implies that $a_k = i_k + j_k - 1 > i_{k+1} + i_{k+1} - 1 = a_{k+1}$. Hence (3) in (7.10) holds. Finally, we need to show that $\mathbf{a}$ is a reduced word for w. This follows from the fact that the paths are wavy diagonal lines moving southeast along the trajectories $i + j =$ constant (where i is the row index and j is the column index), which are allowed to cross, but never to recross since the classical pipedreams are reduced. Hence $(\mathbf{a}, \mathbf{r}) \in \mathcal{RC}(w)$. □

Example 7.7. Let $w = s_1 s_2 s_1 s_3 \in S_4$, so that $w = 3241$ in one-line notation. The following is a reduced classical pipedream in $\mathcal{RP}(w)$



where

red (•) > blue (•) > green (•) > orange (•).

The ordered list of the coordinates of the crosses is

$$(1,2),(1,1),(2,1),(3,1)$$

so that $\mathbf{r} = (1,1,2,3)$ and $\mathbf{a} = (2,1,2,3)$.

We may also interpret a compatible sequence as a *decreasing factorization* of a reduced word for w. Recall that $\mathbf{a} = (a_1, \ldots, a_\ell)$ is a reduced word for w. By the conditions in (7.10), the values $a_k > a_{k+1}$ are decreasing whenever $r_k = r_{k+1}$. Hence viewing all a_k with corresponding $r_k = i$ as part of the i-th decreasing factor of the reduced word for w, we may interpret $(\mathbf{a}, \mathbf{r})$ as a decreasing factorization.

Example 7.8. Continuing Example 7.7, we find that

$$(\mathbf{a}, \mathbf{r}) = ((2,1,2,3),(1,1,2,3))$$

corresponds to the decreasing factorization $(21)(2)(3)$.

2.3. Chute moves and crystal operators. Bergeron and Billey 1993 defined *chute moves* on classical pipedreams. Recall that a classical pipedream is uniquely specified by the position of the crosses. A chute move on rows i and $i+1$ for columns c_1 to c_2 (with $c_1 < c_2$) changes a configuration of crosses as follows

$$\begin{array}{cccccccccccccc} & & c_1 & & & & & c_2 & & c_1 & & & & & c_2 \\ & i & \cdot & + & + & + & + & + & \mapsto & \cdot & + & + & + & + & \cdot \\ & i+1 & \cdot & + & + & + & + & \cdot & & + & + & + & + & + & \cdot \end{array} \tag{7.11}$$

Lenart 2004 considered a subset of these chute moves. Lenart's chute moves on classical pipedreams were rediscovered in Gold, Milićević, and Sun 2024a,b as Demazure crystal operators by showing that the weight preserving bijection of Section 2.2 intertwines the crystal operators of Assaf and Schilling 2018 on decreasing factorizations with the chute moves.

To define the crystal chute moves of Gold, Milićević, and Sun 2024a,b; Lenart 2004, we need to define a pairing process that pairs some of the crosses in row i with some of the crosses in row $i+1$. This pairing process matches crosses in row i one at time, from right to left.

Definition 7.9. Given a reduced classical pipedream p for a permutation in S_n, we fix a row index $i \in [n]$. Denote the rightmost cross in row i by c. (Since crosses only occur in boxes (i,j) such that $i + j \leqslant n$, p has no crosses in row n.) We define a pairing process on row $1 \leqslant i < n$ of p as follows:

(1) Look for an unpaired cross c_+ in row $i+1$ whose column index is greater than or equal to that of c, so that c_+ lies below and weakly to the right of c in the diagram p. If there are multiple such c_+, choose the leftmost c_+.
   (a) If such a c_+ exists, we say that c and c_+ are paired.
   (b) If no such c_+ exists, we say that c is unpaired.

(2) Denote by c' the cross in row i which is both closest to c and lies to the left of c.
   (a) If such a c' exists, we reset $c := c'$ and start again from step (1).
   (b) If no such c' exists, the pairing process on row i is complete.

The crystal operators acting on a reduced classical pipedream p depend on the set P_+ of coordinates of the crosses in p.

Definition 7.10. Define *(*lowering) crystal chute moves

$$f_i \colon \mathcal{RP}(w) \to \mathcal{RP}(w) \sqcup \{\emptyset\} \qquad \text{for } w \in S_n \text{ and } 1 \leqslant i < n$$

as follows. Let $p \in \mathcal{RP}(w)$. Perform the pairing process of Definition 7.9 on row i of p.

(1) If all crosses in row i are paired, set $f_i p = \emptyset$.
(2) Otherwise, denote by $(i, j) \in P_+$ the leftmost unpaired cross in row i.
   (a) If $(i, k) \in P_+$ for all $1 \leqslant k \leqslant j$, set $f_i p = \emptyset$.
   (b) Otherwise, $f_i p$ is given by the chute move of (7.11) with $c_2 = j$. More formally, define $m \in \mathbb{N}$ such that:
      (i) $(i, j-m), (i+1, j-m) \notin P_+$ and
      (ii) $(i, j-k), (i+1, j-k) \in P_+$ for all $1 \leqslant k < m$.
   Define a new classical pipedream by

$$f_i p = p \setminus \{(i, j)\} \cup \{(i+1, j-m)\}.$$

Example 7.11. Consider the classical pipedream of Example 7.7. For $i = 1$, the crosses in positions $(1, 1)$ and $(2, 1)$ are paired. The cross with coordinates $(1, 2)$ is unpaired. We are in case (ii) (a) of Definition 7.10 and hence $f_1 p = \emptyset$.

Similarly, let p_1 and p_2 be the classical pipedreams on the left and right of (**??**), respectively. Then $f_1 p_1 = p_2$.

The operators f_i are called lowering operators since they lower the weight of a reduced classical pipedream by a simple root α_i. That is, if $f_i p \neq \emptyset$ for $p \in \mathcal{RP}(w)$, then

$$\mathsf{wt}(f_i p) = \mathsf{wt}(p) - \alpha_i.$$

The lowering crystal chute moves have partial inverses called raising crystals chute moves which we define next.

Definition 7.12. Define *(*raising) crystal chute moves

$$e_i \colon \mathcal{RP}(w) \to \mathcal{RP}(w) \sqcup \{\emptyset\} \qquad \text{for } w \in S_n \text{ and } 1 \leqslant i < n$$

as follows. Let $p \in \mathcal{RP}(w)$, and perform the pairing process of Definition 7.9 on row i of p.

(1) If all crosses in row $i+1$ are paired, set $e_i p = \emptyset$.
(2) Otherwise, let $(i+1, j) \in P_+$ be the rightmost most unpaired cross in row $i+1$. Let $q > j$ be minimal such that $(i+1, q) \notin P_+$. Then

$$e_i p = p \setminus \{(i+1, j)\} \cup \{(i, q)\}. \tag{7.12}$$

A reduced pipedream $p \in \mathcal{RP}(w)$ is called *highest weight* if $e_i p = \emptyset$ for all $1 \leqslant i < n$.

One of the main results of Assaf and Schilling 2018 (translated to classical pipedreams in Gold, Milićević, and Sun 2024a,b) is that $\mathcal{RP}(w)$ with the crystal chute moves f_i and e_i decomposes into a union of Demazure crystals. To state the results, we first need to review an algorithm to extract a permutation from a highest weight pipedream (see Gold, Milićević, and Sun 2024b, Algorithm 6.1).

Definition 7.13 (Algorithm 6.1 Gold, Milićević, and Sun 2024b)**.** For $w \in S_n$, let $p \in \mathcal{RP}(w)$ be a highest weight element. On the set of crosses P_+ of p perform the following:

(1) Shift all crosses in row i to the right by $i - 1$.

(2) For each row, beginning in the lowest row, move the leftmost cross say in position (r, c) down to position (c, c), so that its row and column index match. Fix these crosses.
Set $\ell = 2$.
(3) (a) Beginning with the lowest row containing unfixed crosses, consider the leftmost unfixed cross. Move that cross down to the lowest possible row, remaining in its current column, such that:
(i) The cross does not move through other crosses;
(ii) The cross is the ℓ-th cross from the left in its new row; and
(iii) The cross does not have any previously fixed crosses to its right in the new row.
(b) Fix this moved cross.
(c) Repeat steps 3a and 3b once within each row until all rows with unfixed crosses have been considered.
(4) Increment ℓ by 1 and repeat step 3.

Once all crosses in the diagram are fixed, the algorithm terminates. Denote the resulting diagram by $\tilde{p}$.

The *truncating permutation* π_p is the shortest permutation such that $\mathsf{wt}(\tilde{p}) = \pi_p(\mathsf{wt}(p))$.

Theorem 7.14 (Assaf and Schilling 2018; Gold, Milićević, and Sun 2024a,b)**.** *Given $w \in S_n$, the operators e_i and f_i for $1 \leqslant i < n$ from Definitions 7.10 and 7.12 define a type A_{n-1} Demazure crystal on $\mathcal{RP}(w)$ as follows:*

$$\mathcal{RP}(w) \cong \bigcup_{\substack{p \in \mathcal{RP}(w) \\ e_i p = \emptyset, \forall 1 \leqslant i < n}} B_{\pi_p}(\mathsf{wt}(p)), \tag{7.13}$$

where the truncating permutation π_p is given in Definition 7.13.

Example 7.15. Let $w = (1, 2, 5, 4, 3) \in S_5$. A reduced expression for this permutation is $w = s_3 s_4 s_3$. The Schubert polynomial indexed by w can be computed to be $\mathfrak{S}_{s_3 s_4 s_3}(x_1, x_2, x_3) = s_{(2,1)}(x_1, x_2, x_3)$. The classical pipedreams are arranged as a crystal in Figure 1.

The crosses for the highest weight element in Figure 1 are given by

$$p = \begin{array}{ccccc} \cdot & \cdot & + & + & \cdot \\ \cdot & \cdot & + & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{array}.$$

After applying the algorithm of Definition 7.13, we obtain

$$\begin{array}{ccccc} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & + & + & \cdot \\ \cdot & \cdot & \cdot & + & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

so that $\pi_p = s_2 s_3 s_1 s_2$.

Note that the Schubert polynomial of Example 7.15 is indeed a symmetric function. It turns out that this is always the case in the stable limit as we will see in Section 3. The

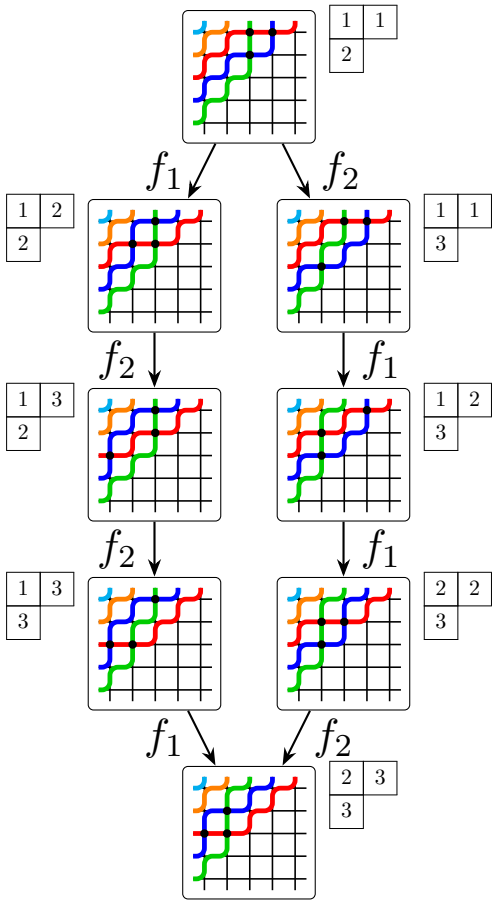


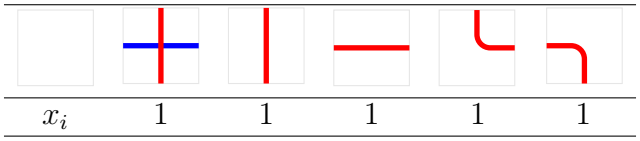
FIGURE 1. The classical pipedreams for the permutation $w = (1, 2, 5, 4, 3) \in S_5$ arranged as a crystal $B_3(2, 1)$.

stable limit of Schubert polynomial $\mathfrak{S}_w$ for $w \in S_n$ is defined as (see Macdonald 1991a,b)

$$\sum_{m\to\infty} \mathfrak{S}_{1^m\times w}(x_1, x_2, \ldots, x_{n+m}).$$

Remark 7.16. Note that instead of considering row chute moves, one could have also used columns chute moves for an analogous column crystal structure.

2.4. Bumpless pipedreams. In 2021, Lam, Lee, and Shimozono 2021 introduced another combinatorial model for (double) Schubert polynomials in terms of *bumpless pipedreams*. In this model the following set of tiles is allowed



The first tile is empty with spectral parameter x_i associated to row i, and the second tile is a cross where the colors are ordered as

$$\text{red } (\bullet) > \text{blue } (\bullet).$$

The remaining tiles are vertical, horizontal, and corner tiles. Note that unlike for classical pipedreams, there are no bump tiles, hence the name bumpless pipedreams.

Remark 7.17. Note that the bumpless pipedream tiles are in bijection with the six states of the six-vertex model introduced in Chapter 5.

The *bumpless pipedream model* is given by a square grid with n rows and columns, with columns labeled by $1, \ldots, n$ from left to right at the bottom, and the rows are labeled by the one-line notation of $w \in S_n$ from bottom to top on the right. The colors on the top boundary are ordered with the largest color to the left.

Pipes move from the top down and to the right. Note that with this scheme, two pipes may cross once but they cannot cross a second time.

Remark 7.18. Our convention differs from Lam, Lee, and Shimozono 2021 in that rows are reflected top to bottom. This convention is more suitable for the study of hybrid pipedreams (see Section 4) with the conventions for classical pipedreams of Section 2.1.

Denote by $\mathcal{BP}(w)$ the set of reduced bumpless pipedreams with right boundary given by w. Define

$$Z_w^{\mathcal{BP}}(\mathbf{x}) = \sum_{p \in \mathcal{BP}(w)} x^{\mathsf{wt}(p)}$$

to be the partition function of the bumpless pipedream model.

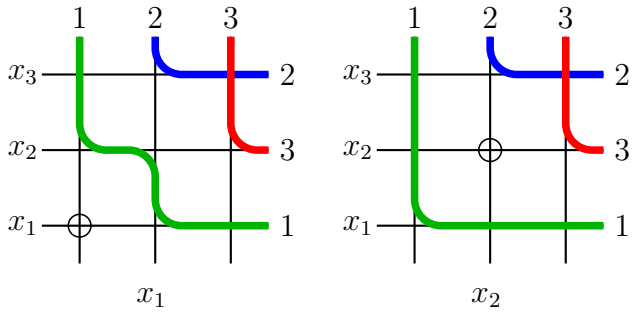
Theorem 7.19 (Lam, Lee, and Shimozono 2021)**.** *The Schubert polynomial is the partition function of the bumpless pipedream model*

$$\mathfrak{S}_w(\mathbf{x}) = Z_w^{\mathcal{BP}}(\mathbf{x}).$$

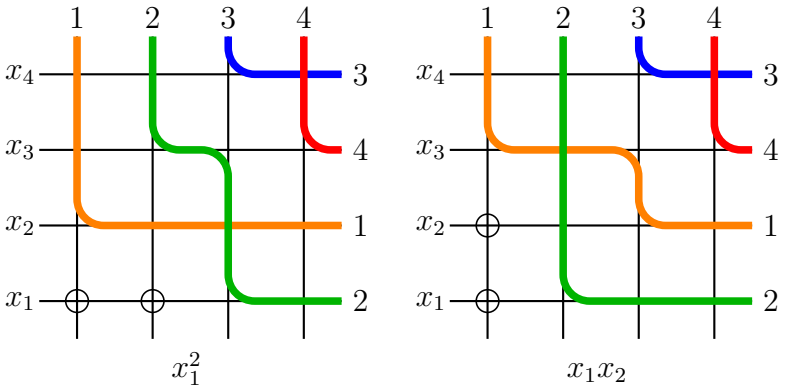
Example 7.20. In Example **??**, we computed $\mathfrak{S}_{s_2}(x_1, x_2, x_3) = x_1 + x_2$ using classical pipedreams. Let us now compute this Schubert polynomial using bumpless pipedreams. As before, we use the ordering of colors

$$\text{red } (\bullet) > \text{blue } (\bullet) > \text{green } (\bullet).$$

The two bumpless pipedream configurations in this model with weight x_1 and x_2, respectively:

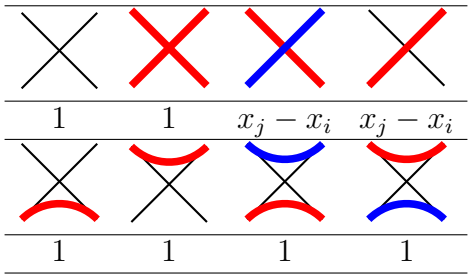


Example 7.21. Let us now compute the Schubert polynomial $\mathfrak{S}_{s_1 s_3}(x_1, x_2, x_3, x_4) = x_1^2 + x_1x_2 + x_1x_3$ from Example 7.4 using bumpless pipedreams:

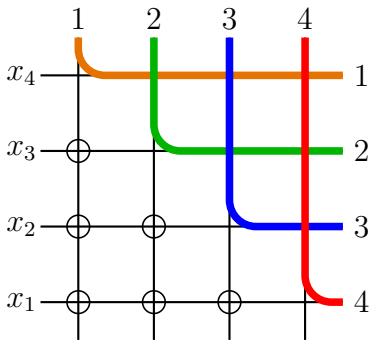


For bumpless pipedreams, the R-matrix is given by

(7.14)



Proof of Theorem 7.19. For $w = w_0$, there is only one configuration for the bumpless pipedream model, with $n - i$ empty cells in row i. The case $n = 4$ is shown below:



Hence the weight for this configuration is $\mathbf{x}^\rho = x_1^{n-1}x_2^{n-2}\cdots x_{n-1}$. This shows that $\mathfrak{S}_{w_0}(\mathbf{x}) = Z_{w_0}^{\mathcal{BP}}(\mathbf{x})$.

Using the train argument it can be shown that $Z_w^{\mathcal{BP}}(\mathbf{x})$ satisfies the same recursion as the Schubert polynomial (see Exercise 7.5)

$$Z_{ws_i}^{\mathcal{BP}}(\mathbf{x}) = D_i Z_w^{\mathcal{BP}}(\mathbf{x}).$$

□

3. Stanley symmetric functions

The Stanley symmetric function F_w indexed by a permutation $w \in S_n$ was introduced by Stanley 1984 to enumerate the reduced decompositions of w. The Stanley symmetric functions are related to Schubert polynomials through a stable limit

$$F_w(\mathbf{x}) = \lim_{s\to\infty} \mathfrak{S}_{1^s\times w}(\mathbf{x}),$$

where $1^s \times w$ is the permutation $12\ldots s(w_1+s)\ldots(w_n+s)$ in one-line notation.

Stanley 1984 conjectured that the coefficients c_w^λ in the Schur expansion of F_w

$$F_w(\mathbf{x}) = \sum_\lambda c_w^\lambda \; s_\lambda(\mathbf{x})$$

are nonnegative integers. This was shown by Edelman and Greene 1987 through an insertion algorithm from reduced words of w to a pair of tableaux now known as the Edelman–Greene insertion. This algorithm shows that c_w^λ is equal to the number of increasing tableaux of shape λ whose row reading words are reduced words of w^{-1}. See also Fomin and Greene 1998; Lascoux and Schützenberger 1985.

The Edelman–Greene coefficients c_w^λ can also be understood by counting highest weight elements in a crystal. As mentioned in Section 2.2, reduced classical pipedreams are equivalent to compatible sequences which in turn can be interpreted as decreasing factorizations of reduced words of w. Denote by $\mathcal{W}_w$ the set of all decreasing factorizations of $w \in S_n$. More precisely, these are all factorizations $w^k w^{k-1}\cdots w^1$ such that $w = w^k w^{k-1}\cdots w^1$ with $\ell(w) = \ell(w^1)+\cdots+\ell(w^k)$ and each w^i is decreasing. Then from the results in Section 2.2 one can deduce that

$$F_w(\mathbf{x}) = \sum_{w^k\cdots w^1\in\mathcal{W}_w} x_1^{\ell(w^1)}\cdots x_k^{\ell(w^k)}.$$

In Morse and Schilling 2016, a crystal structure on $\mathcal{W}_w$ was defined (see also Bump and Schilling 2017, Chapter 10). Through the correspondence between decreasing factorizations and reduced classical pipedreams, this crystal is related to the crystal of Section 2.3.

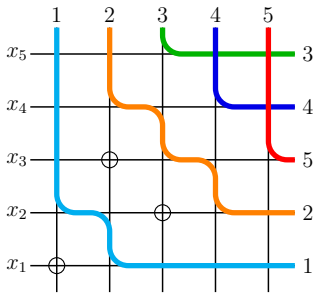
In this section, we define an analogous crystal for Stanley symmetric polynomials in $s+1$ variables on bumpless pipedreams. Lam, Lee, and Shimozono 2021 defined *EG-pipedreams* as the bumpless pipedreams, where all the empty boxes are in the southwest corner (in our convention) and form a partition. The partition is called the *shape* of the EG-pipedream.

Example 7.22. The first two bumpless pipedreams in Example 7.21 are EG-pipedreams with shapes (2) and (1, 1), respectively. The last bumpless pipedream in this example is not an EG-pipedream since not all empty cells are in the southwest corner.

By Lam, Lee, and Shimozono 2021, Theorem 5.14, the Edelman–Greene coefficient c_w^λ is equal to the number of EG-pipedreams of w of shape λ.

The EG-pipedreams correspond to the highest weight elements in the crystal. In general, a bumpless pipedream for the permutation $1^s \times w$ with empty boxes only in the bottom $s+1$ rows can be associated to a semistandard Young tableau T of shape λ as follows. The shape λ is obtained by sliding all empty boxes along their diagonals $d = c - r$, where r is the row index (from the bottom) and c is the column index (from the left), towards their southwest corners. The entries of the semistandard Young tableau T on the diagonal d are the row indices of the empty boxes on diagonal d in the bumpless pipedream.

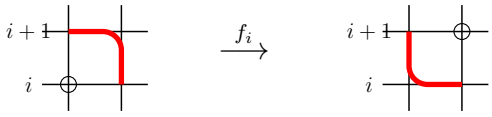
Example 7.23. Consider the bumpless pipedream



for the permutation $(1, 2, 5, 4, 3) = 1^2 \times w$ for $w = (3, 2, 1)$. It has an empty cell on diagonal 0 and row 1, diagonal 1 and row 2, and diagonal -1 and row 3. This bumpless pipedream corresponds to the semistandard Young tableau $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$, which has an entry 1 on diagonal 0, entry 2 on diagonal 1, and entry 3 on diagonal -1.

We would like to emphasize that the association of a semistandard Young tableaux to a bumpless pipedream of $1^s \times w$ with empty boxes only in the bottom $s + 1$ rows is not a bijection; for example there can be multiple EG-pipedreams of the same shape λ, which would all be associated to the same highest weight semistandard Young tableau of shape λ. For a fixed configuration of the strands $s + 1, s + 2, \ldots, s + n$, the map is however bijective.

We have discussed in Chapter 8 how to define crystal operators f_i and e_i for $1 \leqslant i \leqslant s$ on semistandard Young tableaux of shape λ in the alphabet $\{1, 2, \ldots, s + 1\}$. The above map from certain bumpless pipedreams for $1^s \times w$ to semistandard Young tableaux thus induces a crystal structure on these bumpless pipedreams by keeping the stands $s + 1, \ldots, s + n$ fixed. Note that the first s strands for the bumpless pipedreams associated to $1^s \times w$ do not cross. Recall that an f_i crystal operator on a semistandard Young tableaux locally changes a letter i to a letter $i + 1$. Hence locally an f_i crystal operator on a bumpless pipedreams for $1^s \times w$ acts as a droop move



An example is given in Figure 2.

4. Hybrid pipedreams

Both the classic and bumpless pipedream models represent double Schubert polynomials. Yet the models seem rather different. Knutson and Udell 2023 observed that one can mix

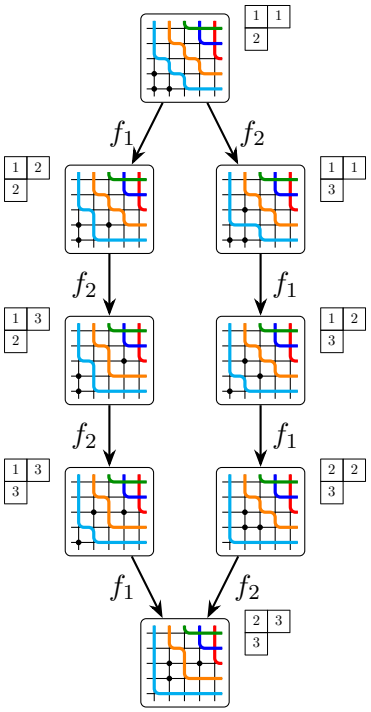
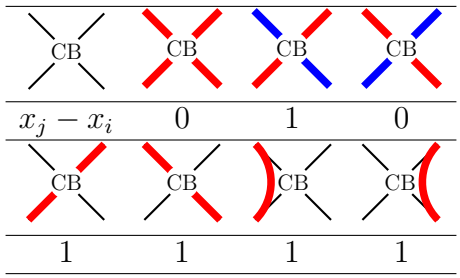


FIGURE 2. The crystal $B_3(2,1)$ on the bumpless pipedreams for $w = (1,2,5,4,3)$. The corresponding semistandard Young tableaux are also indicated.

layers of classic and bumpless pipe freely to obtain interesting hybrid models. These are analogous to the hybrid models for Gamma and Delta ice that we considered in Chapter 7, and indeed the structure of the proof of Theorem 7.25 here is very similar to the proof of Theorem 4.21.

In this section we consider the Knutson–Udell hybrid models. A difference is that we will replace their Theorem 3 by an application of the Yang–Baxter equation.

The version of the Yang–Baxter equation that we require uses the following R-matrix:

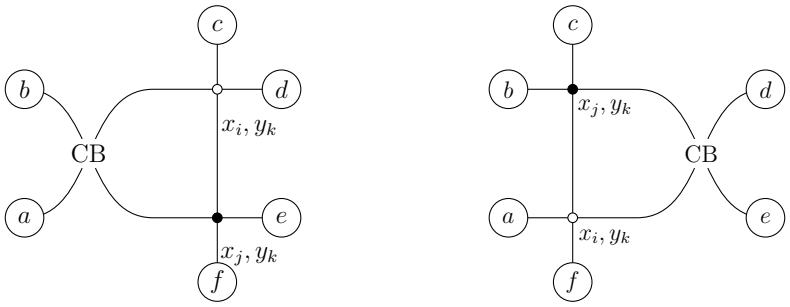


This assumes that red is the larger color. The CB R-matrix depends on two parameters x_i and x_j which we suppress from the notation. Its R-matrix resembles bumpless pipedreams in that if two paths of the same color meet, they must cross, and crossings are only allowed in one direction.

Because we will be mixing classic and bumpless pipedreams, we will use the following notation. A vertex of classic pipedream is marked with a white dot $\circ$, while a vertex of bumpless pipedream is marked with a black dot $\bullet$. We also decorate the vertices with two parameters, x_i and y_k for some k. In the applications, the subscript k of y_k is the column

number. But since the rows move up and down, the parameter i in x_i is not necessarily the row number. As we have already seen, in the pure classic pipedream models, the parameter in the i-th row is x_i; in the pure bumpless pipedream models, the parameter in the i-th row is x_{n+1-i}. In the hybrid models, the parameter in the i-th row can be any one of $x_1, \ldots, x_n$.

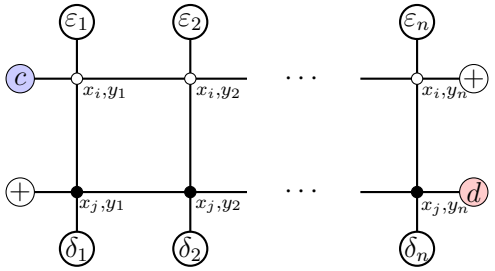
The Yang–Baxter equation then has the following form:



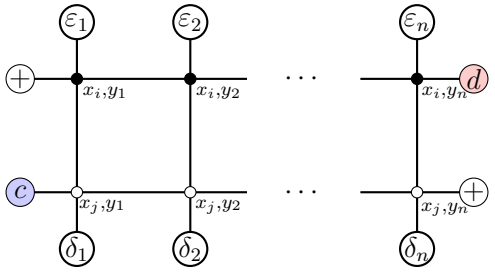
This is valid for any selection of boundary spins (colors or $\oplus$) a–f. To prove this, at most 3 colors can appear in any legal configuration, so all cases can easily be checked using a computer program or otherwise.

Proposition 7.24. *Let c and d be colors (possibly equal). Then the two following systems are equivalent for any choice of the boundary spins:*

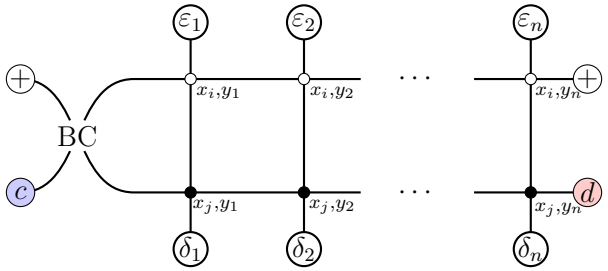
(7.15)



(7.16)



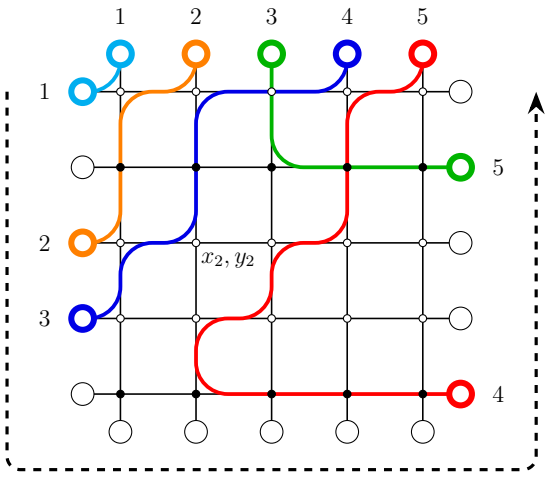
Proof. We attach the CB R-matrix as follows:



Note that there is only one legal configuration for the R-matrix, so this equals (7.15). Now we may run the train argument and detach the R-matrix to obtain (7.16). □

We may now describe the Knutson–Udell mixed models, which contain the classical pipedream and bumpless pipedream models as special cases. Columns are labeled 1 to n. Every row is either a row of a classic or bumpless pipedream. For the boundary conditions, the top boundary colors are, as with the classic and bumpless pipedream models, $c_1, \ldots, c_n$ in ascending order from left to right. Each classic row gets one of these colors at its left boundary, and $\oplus$ signifying no color at the right; and each bumpless row gets one of these colors at its right boundary, with $\oplus$ on the left.

The rows and columns are labeled as shown.



The columns are labeled 1–n from left to right. For the rows, the rule is more complicated: we visit the boundary edges as indicated by the arrow, first labeling the classic rows on the left edge from top to bottom (but skipping the bumpless rows), then labeling the remaining bumpless rows from bottom to top.

We use the row and column parameters x_i in the i-th row, labeled this way, and y_j in the j-th column. The Boltzmann weights are then as in Subsection 2.1 for the classical rows, and as in Subsection 2.4 for the bumpless rows. We have shown a sample state that is legal in both the classic and bumpless rows, and labeled the vertex in row 2 (by the numbering scheme just described) and column 2.

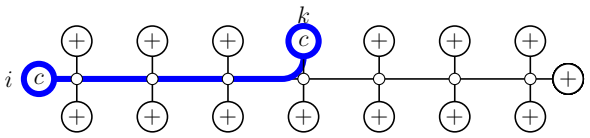
Let w be the permutation such that $c_{w(i)}$ is the color of the colored boundary edge in the i-th row. In this example, the permutation is (12453) in 1-line notation. The model is determined by the permutation w, with one other piece of information, which is the sequence of classic and bumpless rows. We can indicate this information in a *code*, which is a sequence of n letters B or C indicating the species of the rows ordered from top to bottom. So for the above example, the code is $CBCCB$ indicating that we have rows of classic, then bumpless pipes stacked in that order. We may indicate the hybrid model as $\mathfrak{P}_w(\gamma)$, where γ is the code.

Theorem 7.25 (Knutson-Udell)**.** *Let w be a permutation and γ any code. Then the partition function $Z(\mathfrak{P}_w(\gamma)) = \mathfrak{S}_w(\mathbf{x}; \mathbf{y})$.*

Proof. If $\gamma = C^n$ or B^n this is the classic or bumpless pipedream model, and we have already shown that the partition function is the double Schubert polynomial. So the essence of the statement is that the models $\mathfrak{P}_w(\gamma)$ all have the same partition function, independent of the code. In this way we also obtain a new proof that the classic and bumpless representations of $\mathfrak{S}_w(\mathbf{x}; \mathbf{y})$ represent the same value.

It follows from Proposition 7.24 that we may interchange B and C layers. We need to supplement this fact with another fact: we may change the bottommost layer from C to B without changing the partition function. Let us see why this is true. Let γ be a code ending in C, and let γ' be the code obtained by changing the C to be B. We will describe a weight preserving bijection between the states of $\mathfrak{P}_w(\gamma)$ and $\mathfrak{P}_w(\gamma')$.

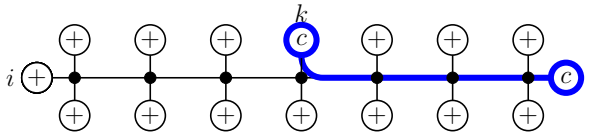
Depending on the code γ, this row is numbered i for some value, according to the scheme described above. It carries a unique pipe that descends in some column numbered k for some value.



The vertices in this row in columns 1 to $k-1$ contribute:

$$\prod_{j=1}^{k-1}(x_i - y_j). \tag{7.17}$$

We may modify this state to obtain a state of $\mathfrak{P}_w(\gamma')$ by reversing the direction of the path:



Again, the vertices in the bottom row to the left of column k contribute the same factor (7.17), and these two corresponding states of $\mathfrak{P}_w(\gamma)$ and $\mathfrak{P}_w(\gamma')$ have the same weight. Summing over all states, the partition functions are equal.

Now we may see why all codes γ give the same partition function. We are allowed to do the following operations on the code: either interchange a B and a C, or change the last letter from C to B or vice versa. □

Exercises

Exercise 7.1. Show that the dimension of the Schubert cell X_w° is given by

$$\dim X_w^\circ = \binom{n}{2} - \ell(w),$$

where $\ell(w)$ is the length of the permutation $w \in S_n$.

Exercise 7.2. Prove the formula (7.2) for double Schubert polynomials.

Exercise 7.3. Show that with the R-matrix in (7.8) for classical pipedreams, the Yang–Baxter equation is satisfied.

Exercise 7.4. Show that with the R-matrix in (7.14) for bumpless pipedreams, the Yang–Baxter equation is satisfied.

Exercise 7.5. Use the train argument to show that

$$Z_{ws_i}^{\mathcal{BP}}(\mathbf{x}) = D_i Z_w^{\mathcal{BP}}(\mathbf{x}).$$

CHAPTER 13

The Fermionic Fock Space

In Chapter 12 we saw that in the field-free six-vertex model there is a Hamiltonian H and also a commuting family of six-vertex model row transfer matrices T_θ acting on a Hilbert space, which in that case was $\mathcal{H} = \otimes^N \mathbb{C}^2$. The main theorem is that H commutes with T_θ, which was proved by showing that $H = (T_\theta^{-1} T_\theta')|_{\theta=\chi} + cI_{\mathcal{H}}$ for a suitable constant c. This result was proved by Baxter, in the greater generality of the eight-vertex model.

For the free-fermionic six-vertex model, there is a similar result, due to Brubaker and Schultz Brubaker and Schultz 2018. In the proof we will follow Brubaker, Buciumas, Bump, and Gustafsson 2020b, where a more general result is proved. (The models in Brubaker, Buciumas, Bump, and Gustafsson 2020b may be regarded as generalizations of the result in Brubaker and Schultz 2018 to a colored model.) In this free-fermionic case there is a Hamiltonian operator H and a row transfer matrix T, and the result is now in the form $e^H = T$. But the conclusion is the same: the Hamiltonian H commutes with the row transfer matrix T.

The identity $e^H = T$ can be thought of as an expansion of T in terms of operators J_k which move particles right or left to lower or higher energy levels. If $k > 0$, then J_k moves the particle right to a lower energy level, and if $k < 0$ it moves the particle to the left. There are correspondingly two versions of both the Hamiltonian and the row transfer matrix.

1. The fermionic Fock space

The *fermionic Fock space* was invented by Dirac in the theory of the electron. The electron is described by the *Dirac equation*, which we will not discuss, except to mention that the energy levels are quantized, and there are solutions of arbitrary negative energy. This seems unphysical, since a particle could radiate an arbitrarily large amount of energy by falling to lower and lower energy levels.

But Dirac proposed a solution to this. Since the Dirac equation is linear, solutions can exist in superposition. The electron is a fermion, subject to the Pauli exclusion principle, meaning that no two electrons can occupy the same state. Dirac's proposal was that all sufficiently large negative energy level states are occupied, and all sufficiently large positive energy levels are unoccupied.

Mathematically, the states are vectors in a Hilbert space that is called the *fermionic Fock space* $\mathfrak{F}$, which we will now describe. This is based on another Hilbert space that we call V, with basis vectors u_i ($i \in \mathbb{Z}$). Each u_i represents a particle with a definite energy level equal to i. Let us fix $m \in \mathbb{Z}$ and consider a sequence $\mathbf{j} = (j_m, j_{m-1}, \ldots)$ where $j_m > j_{m-1} > \cdots$ and $j_k = k$ for k sufficiently negative. Define the *charge m fermionic Fock space*, denoted $\mathfrak{F}_m$ to be the free vector space on formal symbols

$$|\mathbf{j}\rangle := |\mathbf{j}\rangle_m = u_{j_m} \wedge u_{j_{m-1}} \wedge \cdots, \qquad \mathbf{j} = (j_m, j_{m-1}, j_{m-2}, \ldots). \tag{13.1}$$

The Fock space $\mathfrak{F}$ resembles the exterior algebra $\bigwedge V$, except that the basis vectors are *infinite* wedges (called *semi-infinite monomials*).

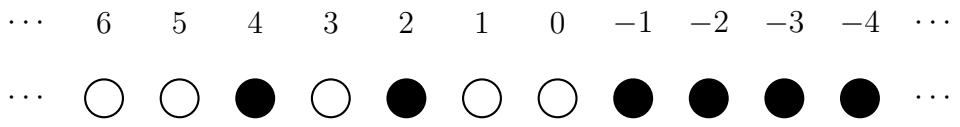
We extend the notation $\xi_{\mathbf{j}}$ to sequences $\mathbf{j} = (j_m, j_{m-1}, \ldots)$, where $j_k = k$ for k sufficiently negative, dropping the assumption that the sequence is strictly decreasing, by the usual rules for $\wedge$ in the exterior algebra. Thus $|\mathbf{j}\rangle = 0$ if $j_k = j_l$ for any distinct $k, l < m$. And interchanging two adjacent indices changes the sign of $|\mathbf{j}\rangle$.

We can visualize the vector $|\mathbf{j}\rangle$ by a *Maya diagram* in which sites numbered by integers are filled with stones. If the site n equals j_k for some k, the site is *occupied*, otherwise it is *unoccupied*. We put a black stone at the occupied sites, and a white stone at the unoccupied sites.

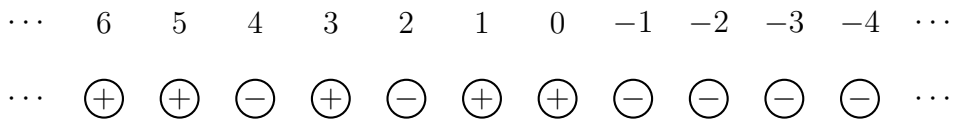
For example, if $\mathbf{j} = (4, 2, -1, -2, -3, -4, \ldots)$ and hence

$$|\mathbf{j}\rangle = u_4 \wedge u_2 \wedge u_{-1} \wedge u_{-2} \wedge u_{-3} \wedge u_{-4} \wedge \cdots$$

then the Maya diagram looks like this:



The main point is that every sufficiently negative site is occupied, and every sufficiently positive site is unoccupied. Although Maya diagrams are traditional (originating in soliton theory with M. Sato and his collaborators), because we want to relate this story to the six-vertex model as we have been, we prefer to use $-$ and $+$ for the occupied and unoccupied sites respectively. Hence the Maya diagram looks as follows:



For this state the charge is $m = 1$.

If $j_k = k$ for all $k \leqslant m$, we obtain the charge m *vacuum vector* for which we have an alternative notation

$$|\varnothing\rangle_m = u_m \wedge u_{m-1} \wedge \cdots .$$

In general we may define the *energy* of $|\mathbf{j}\rangle_m$ to be $\sum_{k \leqslant m}(j_k - k)$. This is a finite sum. The vacuum is the unique semi-infinite monomial in $\mathfrak{F}_m$ of energy 0.

2. The row transfer matrix $T_\Delta(z;q)$

We will describe a version of the free-fermionic six-vertex model that we call *Delta ice*. The grid is of infinite width, and the Boltzmann weights in each row depends on a parameter $z \in \mathbb{C}^\times$.

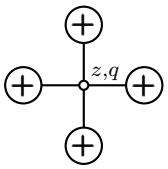
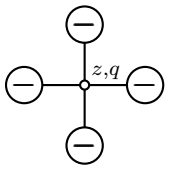
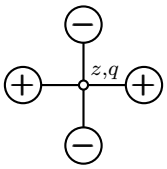
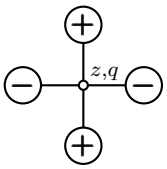
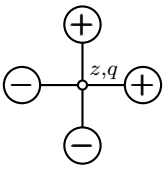
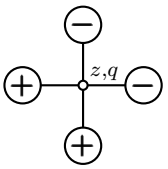
Remark 13.1. The quantity Δ here is different from Baxter's Δ, which is

$$(a_1a_2 + b_1b_2 - c_1c_2)/2a_1b_1.$$

Baxter's Δ is zero here, since all weights in this chapter are free-fermionic.

Let $\mathbf{i} = (i_m, i_{m-1}, \ldots)$ and $\mathbf{j} = (j_m, j_{m-1}, \ldots)$ be two sequences such that $i_m > i_{m-1} > \cdots$ and $j_m > j_{m-1} > \cdots$ and $i_i = j_k = k$ for k sufficiently negative. We define a simple system consisting of a single row, and either no states or a single state. We consider a grid with

only one row that is infinite in both directions. As boundary conditions, the spins of the vertical edges at the top are given by the Maya diagram for $\xi_{\mathbf{i}}$, and for the vertical edges at the bottom, by the Maya diagram for $\xi_{\mathbf{j}}$. There is also a "boundary condition" for the horizontal edges, that there are only finitely many + spins. We use the following Boltzmann weights:

	$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
Δ-ice						
	1	$-qz$	1	z	$(1-q)z$	1

Since as part of the boundary conditions there are only finitely many horizontal edges with + spins, all but finitely factors in the Boltzmann weight of a state are of type $\mathtt{b}_1$ (for vertices far to the left) or of type $\mathtt{a}_1$ (for vertices far to the right). Therefore the Boltzmann weight of a state is an infinite product with only finitely many terms not equal to 1, and so has a well-defined finite value.

Lemma 13.2. *The condition for the partition function to have a state (which is therefore unique) is that*

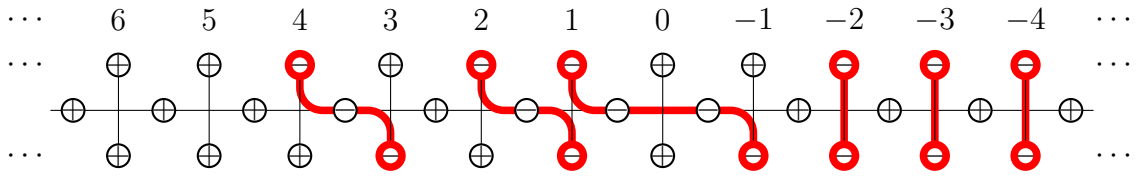
$$i_m \geqslant j_m \geqslant i_{m-1} \geqslant j_{m-1} \geqslant \cdots . \tag{13.2}$$

We express the conditions in (13.2) by saying that the sequences $\mathbf{i}$ and $\mathbf{j}$ *interleave*. We already encountered interleaving partitions in (3.4) in Chapter 3, Section 2.

Proof. This may be seen by consideration of the paths, which we recall from Chapter 3, Section 1 are obtained by joining edges with spin $-$. Because of our boundary condition, that there are only finitely many horizontal edges with spin $-$, each path must begin at the top and exit at the bottom for this system. For example, suppose that $m = 1$ and

$$\mathbf{i} = (4, 2, 1, -2, -3, -4, \ldots), \qquad \mathbf{j} = (3, 1, -1, -2, -3, -4, \ldots).$$

Then we have the following state:



Every path must start in the i_k column and end in the j_k column. Call this the k-th path. We must have $i_k \geqslant j_k$ since the paths move down and to the right. We also need $j_k \geqslant i_{k-1}$ since otherwise two paths will overlap between the i_{k-1} column and the j_k column. □

We quickly review Dirac notation for operators. Let $\mathcal{H}$ be a Hilbert space. A vector in $v \in \mathcal{H}$ is denoted alternatively as $|v\rangle$, and called a *ket*. On the other hand, a vector w gives rise to a linear functional $v \to (v, w)$ using the inner product on $\mathcal{H}$, and we denote this linear functional as $\langle w|$, also called *bra*. The notation works well in quantum mechanics due to the emphasis on Hermitian (self-adjoint) operators. If T is Hermitian, then $(Tv, w) = (v, Tw)$,

which we denote $\langle w|T|v\rangle$. We can either think of this as the linear functional $\langle w|$ applied to the vector $T|v\rangle$, or as the linear functional $\langle w|T$ applied to the vector v.

As a special case, the partition function of the monostatic system above will be denoted

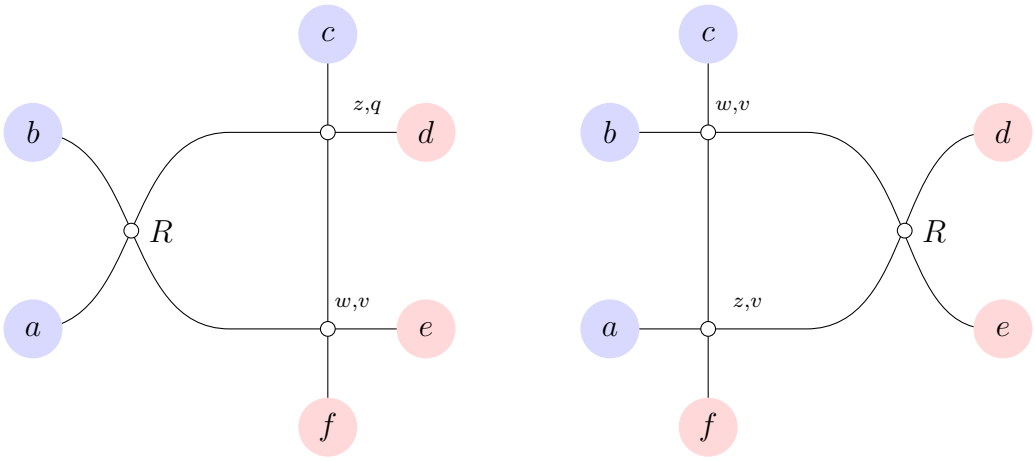
$$\langle \mathbf{j}|T_\Delta(z;q)|\mathbf{i}\rangle,$$

and we are now thinking of $T_\Delta(z;q)$ as being an operator on $\mathcal{H}$.

Theorem 13.3. *The operators $T_\Delta(z;q)$ all commute. That is, if w and v are other parameters, we have*

$$T_\Delta(z;q)T_\Delta(w,v) = T_\Delta(w;v)T_\Delta(z,q).$$

Proof. We make use of the general free-fermionic Yang–Baxter equation from Chapter 3. By Theorem 3.14 of Chapter 3, there exists an R-matrix R depending on z, q, w, v such that we have a Yang–Baxter equation in the form

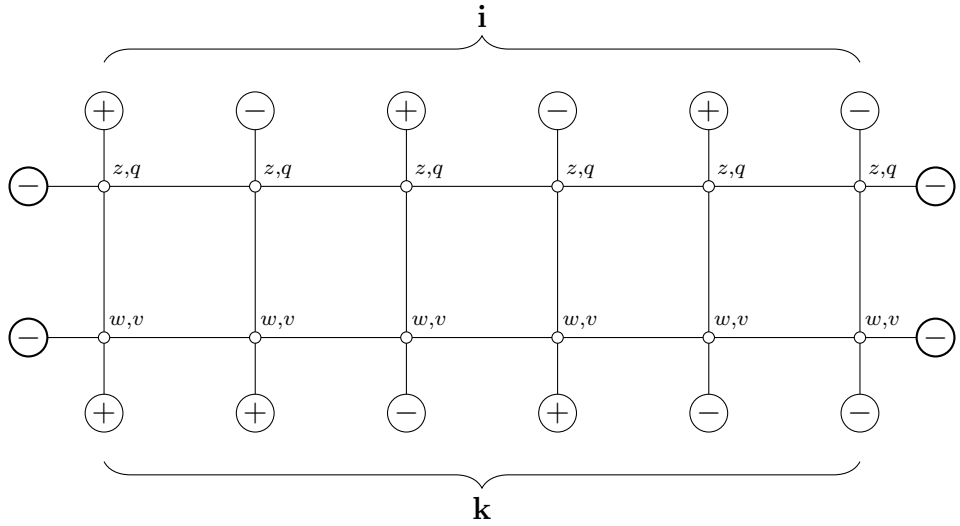


It is of course not hard to compute the Boltzmann weights but we do not need them for this proof. We only need that the $\mathtt{a}_2$ weight of R is nonzero. We fix $\mathbf{i}$ and $\mathbf{k}$ and will show that

$$\langle \mathbf{k}|T_\Delta(w;v)T_\Delta(z,q)|\mathbf{i}\rangle = \langle \mathbf{k}|T_\Delta(z;q)T_\Delta(w,v)|\mathbf{i}\rangle. \tag{13.3}$$

The left-hand side is the partition function of a 2-rowed infinite grid, but we may truncate this to a finite grid such that all sites of $|\mathbf{i}\rangle$ and $|\mathbf{k}\rangle$ to the right are occupied, and all sites

to the left are unoccupied. This partition function looks like this:



All vertices outside this finite grid have type $\mathtt{a}_1$ or $\mathtt{b}_1$, and Boltzmann weight 1, so discarding them does not change the partition function. So the partition function of this system is

$$\langle \mathbf{k}|T_\Delta(w;v)T_\Delta(z;q)|\mathbf{i}\rangle.$$

Now we attach the R-matrix, which multiplies the Boltzmann weight by $a_2(R)$. We apply the train argument, and discard the R-matrix on the right, which divides the Boltzmann weight by the same constant $a_2(R)$. The resulting system has the rows switched, proving (13.3). Since this is true for all $\mathbf{i}$ and $\mathbf{k}$, the row transfer matrices are proved to commute. □

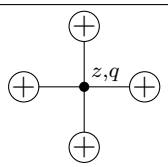
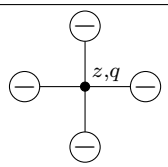
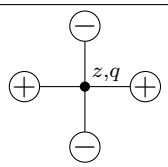
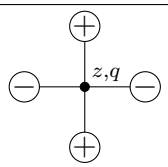
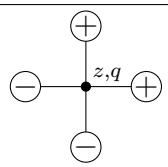
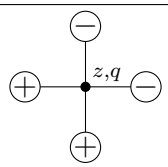
We can define $T_\Delta(z;q)$ as an operator on $\mathfrak{F}$ by

$$T_\Delta(z;q)|\mathbf{i}\rangle = \sum_{\mathbf{j}} \langle \mathbf{j}|T_\Delta(z;q)|\mathbf{i}\rangle\, |\mathbf{j}\rangle. \tag{13.4}$$

The sum on the right is finite, so this defines an element of $\mathfrak{F}$. However $T_\Delta(z;q)$ is not a bounded operator. That is, if we make $\mathfrak{F}$ into a Hilbert space where the semi-infinite monomials $|\mathbf{i}\rangle$ are an orthonormal basis, since the number of terms on the right side of (13.4) can be arbitrarily large, the map $T_\Delta(z;q)$ defined on basis elements does extend to an operator with bounded operator norm.

3. The row transfer matrix $T_\Gamma(z;q)$

There is another type of six-vertex model that is in a sense dual to the models in Section 2. For these we use the following Boltzmann weights:

	$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
Γ-ice						
	z^{-1}	1	$-qz^{-1}$	1	$1-q$	z^{-1}

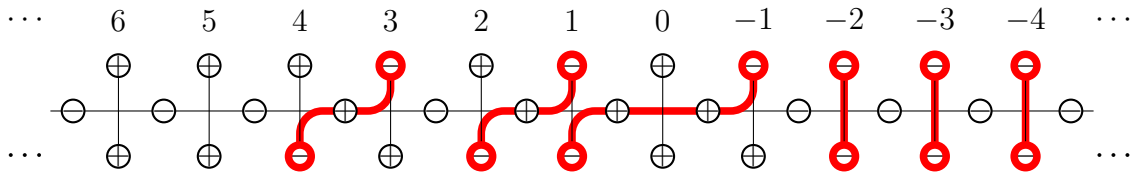
Remark 13.4. These are the same as the weights Tokuyama models introduced in Chapter 4, Section 2, divided by z. Since every weight is divided by the same constant, we could use these weights in the Tokuyama model, and the partition functions would be essentially unchanged, altered only be a constant monomial. However our boundary conditions will be different from the Tokuyama models.

Now we change the boundary conditions. We require all but finitely many horizontal spins to be $-$. This guarantees that the row transfer matrix is essentially a finite product, since all but finitely many spins will be of type $\mathtt{a}_2$ or $\mathtt{b}_2$.

We can define $\langle \mathbf{j}|T_\Delta(z;q)|\mathbf{i}\rangle$ as before, but now the condition for this to be nonzero is changed: we require

$$j_m \geqslant i_m \geqslant j_{m-1} \geqslant i_{m-1} \geqslant \cdots . \tag{13.5}$$

Here is a sample state with $\mathbf{i} = (3, 1, -1, -2, -3, \ldots)$ and $j = (4, 2, 1, -2, -3, \ldots)$. We modify the rule for describing the paths: now the paths follow the $-$ spins on vertical edges, and $+$ spins on the horizontal edges. This means that the paths move down and to the left, so the row transfer matrix is energy raising, in accordance with (13.5).



We can try to define $T_\Gamma(z;q)$ as an operator,

$$T_\Gamma(z;q)|\mathbf{i}\rangle = \sum_{\mathbf{j}} \langle \mathbf{j}|T_\Gamma(z;q)|\mathbf{i}\rangle\, |\mathbf{j}\rangle.$$

However (in contrast with Δ-ice) the sum on the right-hand side is no longer finite.

4. The Heisenberg Lie algebra

We now come to a representation of the *Heisenberg Lie algebra* $\mathfrak{s}$ with generators

$$\{j_k \mid k \in \mathbb{Z}\} \qquad \text{and} \qquad \mathbf{1},$$

with $\mathbf{1}$ central, and

$$[j_k, j_l] = \begin{cases} k & \text{if } k = -l, \\ 0 & \text{otherwise.} \end{cases}$$

The center of $\mathfrak{s}$ is spanned by $\mathbf{1}$ and j_0. This representation is at the heart of the *boson-fermion correspondence*. This is a relationship between the fermionic Fock space and the bosonic Fock space which originated in mathematical physics, and has important applications to representation theory and algebraic combinatorics (Frenkel 1981; Kac, Raina, and Rozhkovskaya 2013; Lam 2006).

We remind the reader that we have defined

$$u_{j_m} \wedge u_{j_{m-1}} \wedge \cdots$$

even if we do not have $j_m \geqslant j_{m-1} \geqslant \cdots$. It is only necessary that $j_k = k$ for k sufficiently negative. However, this monomial might be zero (if some index is repeated) or the negative of a basis element if putting the vectors in order produces an odd number of sign changes. If $j_m \geqslant j_{m-1} \geqslant \cdots$ we denote this vector as $|\mathbf{j}\rangle$. Otherwise we will avoid this notation.

Let $k \in \mathbb{Z}$. For the time being assume that $k \neq 0$. We define an operator J_k on V by $J_k(u_n) = u_{n-k}$. Then we transport J_k to acting on $\mathfrak{F}$ by the Leibnitz rule, so that

$$J_k|\mathbf{j}\rangle = (u_{j_m-k} \wedge u_{j_{m-1}} \wedge \cdots) + (u_{j_m} \wedge u_{j_{m-1}-k} \wedge \cdots) + \cdots .$$

In other words, to apply J_k, we pick one occupied location, and move the particle at that location k steps lower or higher (depending on the sign of k) to an unoccupied location. We also define J_0 to have eigenvalue m on $\mathfrak{F}_m$.

Theorem 13.5. *The operators J_k on $\mathfrak{F}_m$ satisfy*

$$[J_k, J_l] = \begin{cases} k \cdot 1_{\mathfrak{F}_m} & \text{if } k = -l, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $j_k \mapsto J_k$ defines a representation of the Heisenberg Lie algebra.

Proof. Let us first show that

$$J_k J_{-k}|\mathbf{j}\rangle - J_{-k}J_k|\mathbf{j}\rangle = k|\mathbf{j}\rangle. \tag{13.6}$$

We may assume $k > 0$ since the statements for k and $-k$ are trivially equivalent.

First suppose that $|\mathbf{j}\rangle = |\varnothing\rangle_m$ is the vacuum. Then $J_k|\varnothing\rangle_m = 0$. On the other hand, $J_{-k}|\varnothing\rangle$ is a sum of k terms, and applying J_k to each of these produces a copy of $|\varnothing\rangle_m$. Now we prove (13.6) for general $\mathbf{j}$. If $|\mathbf{j}\rangle = |\mathbf{j}\rangle_m$ is not the vacuum may write $|\mathbf{j}\rangle_m = u_j \wedge \eta$ where $j = j_m$ and

$$\eta = u_{j_{m-1}} \wedge u_{j_{m-1}} \wedge \cdots$$

has strictly smaller energy than $|\mathbf{j}\rangle_m$. By induction on enery we may assume that (13.6) is true for η. Now we have $J_{-k} = u_{j+k} \wedge \eta + u_j \wedge J_{-k}\eta$ and so

$$J_k J_{-k}|\mathbf{j}\rangle_m = u_j \wedge \eta + u_{j+k} \wedge J_k\eta + u_{j-k} \wedge J_{-k}\eta + u_j \wedge J_k J_{-k}\eta.$$

Similarly

$$J_{-k}J_k|\mathbf{j}\rangle_m = u_j \wedge \eta + u_{j-k} \wedge J_{-k}\eta + u_{j+k} \wedge J_k\eta + u_j \wedge J_{-k}J_k\eta.$$

Subtracting,

$$J_k J_{-k}|\mathbf{j}\rangle_m - J_{-k}J_k|\mathbf{j}\rangle_m = u_j \wedge (J_k J_{-k}\eta - J_{-k}J_k\eta) = u_j \wedge k\eta = k|\mathbf{j}\rangle_m,$$

where we have used our induction hypothesis.

We leave it to the reader to show that J_k and J_l commute unless $k = -l$. □

5. Row transfer matrices as vertex operators

We emphasize that the J_k with $k > 0$ all commute, and the J_{-k} with $-k < 0$ all commute, so we have two large commuting families of "operators" on $\mathfrak{F}$ or $\mathfrak{F}_m$. The J_{-k} are not operators in the usual sense, since each turns each basis vector into an infinite sum of basis vectors, which is not in $\mathfrak{F}$. Still, the two-point functions

$$\langle \mathbf{i}|J_k|\mathbf{j}\rangle$$

do make sense for all k, and as long as we couch our results in terms of these, there are no difficulties.

Now let us introduce two "Hamiltonians"

$$H_+(z;q) = \sum_{k=1}^{\infty} \frac{1}{k}(1-q^k)z^k J_k, \qquad H_-(z;q) = \sum_{k=1}^{\infty} \frac{1}{k}(1-q^k)z^{-k} J_{-k}.$$

Theorem 13.6 (Brubaker and Schultz 2018)**.** *We have*

$$e^{H_+(z;q)} = T_\Delta(z;q), \qquad e^{H_-(z;q)} = T_\Gamma(z;q). \tag{13.7}$$

The operator $H_+(z;q)$ commutes with $T_\Delta(w;v)$ for all w,v, and the operator $H_-(z;q)$ commutes with $T_\Gamma(w;v)$ for all w,v.

Proof. We will take this up in Section 6. For now we point out that the identities (13.7) imply the commutativity statements, since for example the operators $T_\Delta(w;v)$ and the operator $H_+(z;q)$ are all seen to be expressible in terms of the J_k with $k > 0$, which commute with each other. We also obtain a new proof of the commutativity statement in Theorem 13.3 from this observation in Section 7. □

"Operators" such as $e^{H_+(z;q)}$ and $e^{H_-(z;q)}$, particularly in combinations such as:

$$e^{H_-(z;q)}e^{H_+(z;q)} = \exp\left(\sum_{k=1}^{\infty} \frac{1}{k}(1-q^k)z^{-k}J_{-k}\right)\exp\left(\sum_{k=1}^{\infty} \frac{1}{k}(1-q^k)z^k J_k\right) \tag{13.8}$$

are called *vertex operators*. Here "operators" is in quotation marks since there is a nontrivial problem in making sense of this. Similar expressions appear in conformal field theory and in soliton theory. A purely algebraic and rigorous axiomatization of the underlying mathematics may be found in the theory of *vertex algebras*. In this context, expressions such as (13.8) appear in *lattice vertex algebras* (Frenkel and Ben-Zvi 2004 Chapter 5 or Kac 1998 Section 5.4). See also Kac and Leur 1987 and Jimbo and Miwa 1983.

As mentioned above, a proof of Theorem 13.3 follows by expressing the row transfer matrix $T_\Delta(z;q)$ as the exponential of the Hamiltonian

$$H_+(z;q) = \sum_{k=1}^{\infty} \frac{1}{k}(1-q^k)z^k J_k.$$

There is a corresponding result for T_Γ and H_+ but we will omit that. (It can be deduced from the T_Δ case by taking adjoints, as at the end of Section 4 in Brubaker, Buciumas, Bump, and Gustafsson 2020b.)

6. Fermionic operators

We introduce *fermionic creation operators* ψ_n^* ($n \in \mathbb{Z}$) on $\mathfrak{F}$ that create particles by

$$\psi_n^*(\eta) = u_n \wedge \eta.$$

If η is a basis vector of $\mathfrak{F}_m$, say

$$\eta = |\mathbf{j}\rangle = u_{j_m} \wedge u_{j_{m-1}} \wedge \cdots,$$

then $\psi^*(\eta) = 0$ if n is among the indices $j_m, j_{m-1}, \ldots$. Otherwise, $\psi_n^*(\eta)$ can be calculated by moving u_n to its proper place among the indices. This can involve interchanging some u_j, which can introduce sign changes and so $\psi_n^*(\eta)$ is either zero or $\pm|\mathbf{j}'\rangle$, where $\mathbf{j}'$ is obtained by sorting $\{n, j_m, j_{m-1}, \ldots\}$ into descending order. We see that $\psi_n^*\colon \mathfrak{F}_m \longrightarrow \mathfrak{F}_{m+1}$.

Dual to the creation operators ψ_n^* are their adjoints $\psi_n : \mathfrak{F}_{m+1} \longrightarrow \mathfrak{F}_m$. The operator ψ_n deletes u_n from the semi-infinite monomial if $n \in \{j_m, j_{m-1}, \ldots\}$, which can result in a sign change. If $n \notin \{j_m, j_{m-1}, \ldots\}$, then $\psi_n|\mathbf{j}\rangle = 0$.

Lemma 13.7. *We have*

$$[J_k, \psi_j^*] = \psi_{j-k}^*.$$

Proof. From the Leibnitz rule, if $\eta \in \mathfrak{F}$, then

$$J_k \psi_j^* \eta = J_k(u_j \wedge \eta) = J_k(u_j) \wedge \eta + u_j \wedge J_k(\eta) = u_{j-k} \wedge \eta + \psi_j^*(J_k \eta).$$

Rearranging,

$$[J_k, \psi_j^*]\eta = u_{j-k} \wedge \eta = \psi_{j-k}^*(\eta).$$

□

Now let us introduce the *fermion field*

$$\psi(x) = \sum_{j \in \mathbb{Z}} \psi_j^* x^j.$$

For our purposes this is just a formal expression that we can use to do a calculation. (The "field" terminology comes from quantum field theory.)

Proposition 13.8. *We have*

$$[H_+(z;q), \psi^*(x)] = \log\left(\frac{1-qxz}{1-xz}\right)\psi^*(x). \tag{13.9}$$

Proof. Note that by Lemma 13.7 we have

$$[J_k, \psi^*(x)] = \sum_j x^j [J_k, \psi_j^*] = \sum_j x^j [J_k, \psi_j^*] = \sum_j x^j \psi_{j-k}^* = x^k \psi^*(x).$$

Now the left-hand side of (13.9) equals

$$\sum_k \frac{1}{k}(1-q^k) z^k [J_k, \psi^*(x)] = \sum \frac{1}{k}(1-q^k)(xz)^k \psi^*(x) = -\log(1-xz) + \log(1-qxz)$$

from the identity

$$-\log(1-t) = \sum_{k=1}^{\infty} \frac{t^k}{k}.$$

□

Lemma 13.9. *Suppose that $xa - ax = ca$, where $c \in \mathbb{C}^\times$. Then*

$$e^x a e^{-x} = e^c a.$$

Proof. This is a special case of the Baker–Campbell–Hausdorff formula. We treat this as a formal identity, disregarding convergence. We need the following identity, for $k \geqslant 0$:

$$\sum_j \binom{k}{j} (-1)^j x^{k-j} a x^j = c^k a. \tag{13.10}$$

To avoid some bookkeeping we sum over all $j \in \mathbb{Z}$ but regard $\binom{k}{j}$ as zero unless $0 \leqslant j \leqslant k$, so most terms are zero. Assuming this true for $k-1$, we may establish (13.10) by induction, writing $\binom{k}{j} = \binom{k-1}{j-1} + \binom{k-1}{j}$. The left-hand side equals

$$x \cdot \left[\sum_j \binom{k-1}{j-1} (-1)^j x^{k-1-j} a x^j\right] - \left[\sum_j \binom{k-1}{j-1} (-1)^{j-1} x^{k-j} a x^{j-1}\right] \cdot x.$$

Both terms in brackets equal $c^{k-1}a$ by induction, so we obtain $c^{k-1}[x,a] = c^k a$. This proves (13.10).

Now expand the exponentials and collect terms of degree k to write

$$e^x a e^{-x} = \sum_k \frac{1}{k!} \sum_j \binom{k}{j} (-1)^j x^{k-j} a x^j = \sum_k \frac{1}{k!} c^k a = e^c a,$$

as required. □

Proposition 13.10. *Let $H = H_+(z;q)$. We have*

$$e^H \psi^*(x) e^{-H} = \frac{1-qxz}{1-xz} \psi^*(x). \tag{13.11}$$

Proof. This follows from our Proposition 13.8 by exponentiating (using Lemma 13.9). □

Now the key point is to show that the row transfer matrices $T_\Delta(z;q)$ satisfy the same identity. Let us introduce the operator $\rho_k(z)\colon \mathfrak{F}_m \longrightarrow \mathfrak{F}_{m+1}$ defined by

$$\rho_k(z) = \psi_k^* - z\psi_{k-1}^*.$$

Lemma 13.11. *Granted the invertibility of e^H, the identity (13.11) is equivalent to*

$$e^H \rho_k(z) = \rho_k(qz) e^H. \tag{13.12}$$

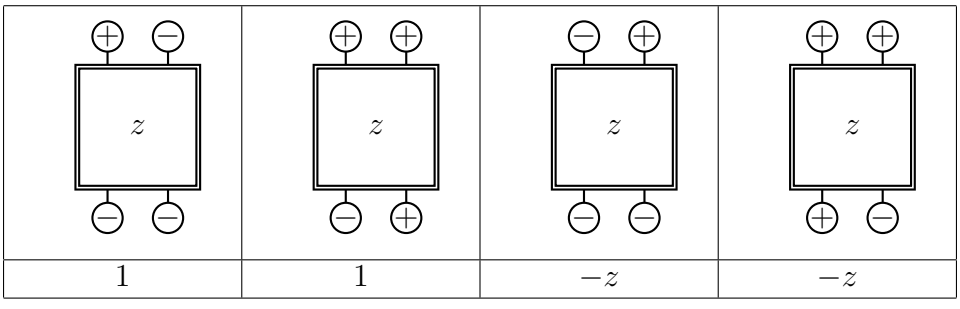
for $k \in \mathbb{Z}$.

Proof. We rewrite (13.11) in the form

$$(1-xz) e^H \psi^*(x) = (1-qxz) \psi^*(x) e^H.$$

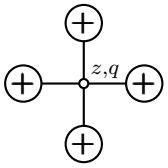
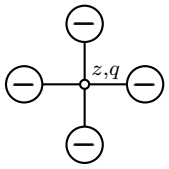
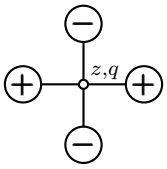
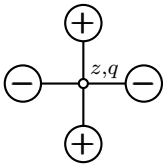
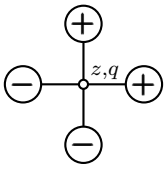
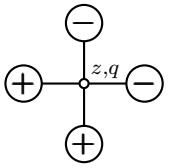
This is a formal identity that can be expanded in powers of x. Comparing the coefficient of x^k gives exactly the identity (13.12). □

Our goal is to show that the row $T = T_\Delta(z;q)$ satisfies the same identity $T\rho_k(z) = \rho_k(qz)T$ as e^H. Let us represent ρ_k graphically as a "gate" that can be attached to the lattice model. Remembering that ψ_k creates a particle in the k-th column, and that $+$ denotes the absence of a particle, $-$ its presence, we see that we have the following Boltzmann weights:



1	1	$-z$	$-z$

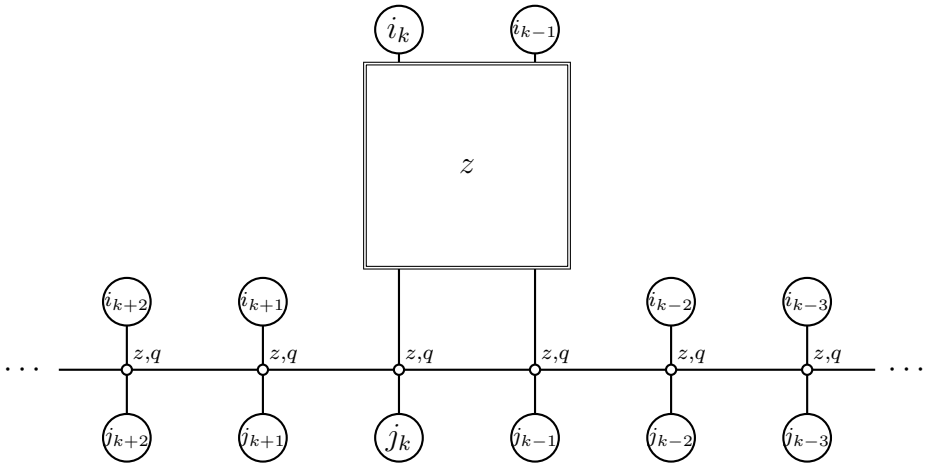
For reference, here are the Delta Boltzmann weights:

	$\mathtt{a}_1$	$\mathtt{a}_2$	$\mathtt{b}_1$	$\mathtt{b}_2$	$\mathtt{c}_1$	$\mathtt{c}_2$
Δ-ice						
	1	$-qz$	1	z	$(1-q)z$	1

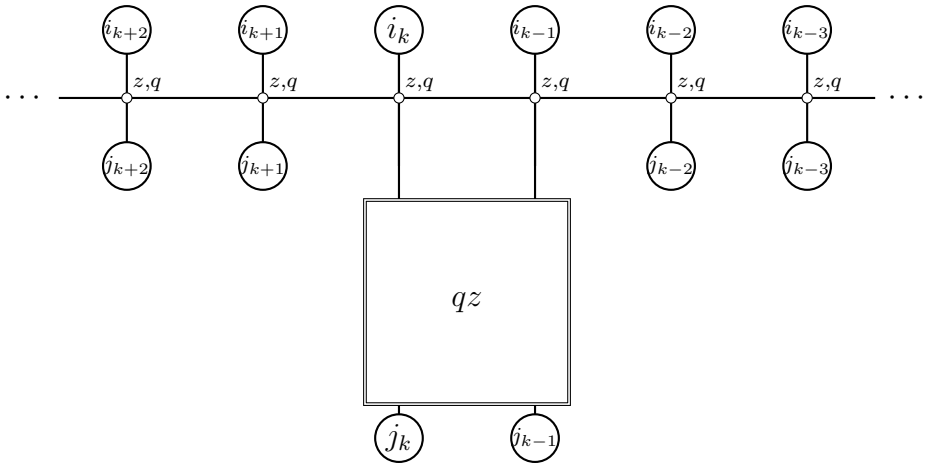
Proposition 13.12. *The row transfer matrix*

$$T\rho_k(z) = \rho_k(qz)T.$$

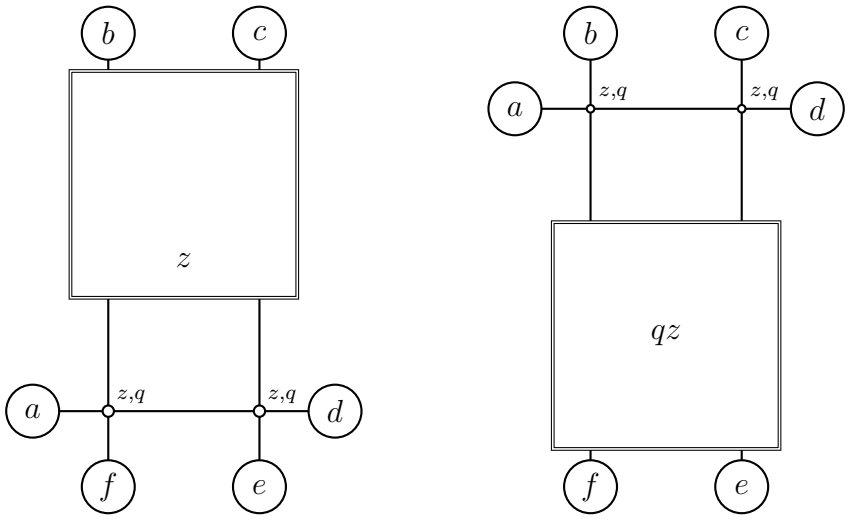
Proof. Graphically this means that we must show the equivalence of the two following partition functions:



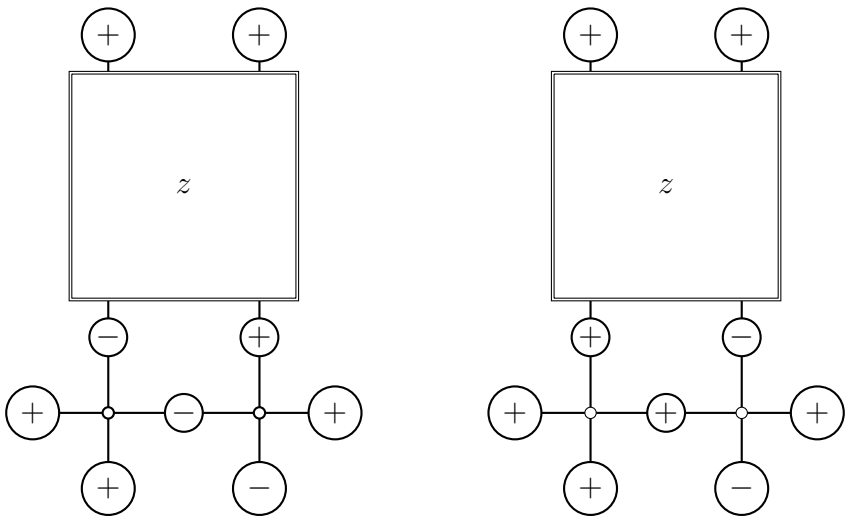
and



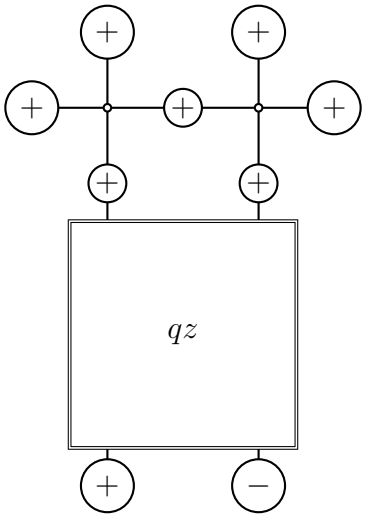
We can clip out the middle part and just prove the equivalence of these systems:



This can be thought of as a kind of a Yang–Baxter equation, but of the sort mentioned in Chapter 11 Section 7, where the R-matrix changes as it moves past the vertices. This verification is now subject to case by case verification. Let us check just one case. Suppose that the boundary values are $(a, b, c, d, e, f) = (+, +, +, +, -, +)$. On the left-hand side there are two admissible states:



Their Boltzmann weights are, respectively $(1-q)z$ and $-z$, for a total of $-qz$. On the right-hand side there is only one admissible state:



The Boltzmann weight is $-qz$. Since $(1-q)z+(-z) = -qz$, the required identity is satisfied in this case, and the remaining cases are similar. □

7. Proof of Theorem 13.3

We will only prove that $e^{H_+(z;q)} = T_\Delta(z;q)$. The identity $e^{H_-(z;q)} = T_\Gamma(z;q)$ can be deduced using adjointness considerations, as in Brubaker, Buciumas, Bump, and Gustafsson 2020b.

As in the last section, we abbreviate $H = H_+(z;q)$ and $T = T_\Delta(z;q)$. We have proved that both operators e^H and T both satisfy the same identities

$$e^H \rho_k(z) = \rho_k(qz) e^H, \qquad T\rho_k(z) = \rho_k(qz)T.$$

We need to show that there is enough information in this fact to deduce that $T|\mathbf{j}\rangle = e^H|\mathbf{j}\rangle$ for every semi-infinite monomial $|\mathbf{j}\rangle \in \mathfrak{F}$.

Recall that the *energy* of $|\mathbf{j}\rangle$, with $\mathbf{j} = (j_m, j_{m-1}, \dots) \in \mathfrak{F}_m$ is $\sum_k (j_k - k)$. This is actually a finite sum. The unique basis vector in $\mathfrak{F}_m$ of energy 0 is the vacuum

$$|\varnothing\rangle_m = u_m \wedge u_{m-1} \wedge \cdots .$$

The identity

$$e^{H_-(z;q)}|\varnothing\rangle_m = T_\Delta(z;q)|\varnothing\rangle_m$$

is clear since both sides are $|\varnothing\rangle_m$.

So assume that $|\mathbf{j}\rangle_m$ is not the vacuum. Then it has positive energy. This means $j_m > m$. We will show

$$e^{H_-(z;q)}|\mathbf{j}\rangle_m = T_\Delta(z;q)|\mathbf{j}\rangle_m. \tag{13.13}$$

We are assuming inductively that the identity is known for states of lower energy.

Let $|\mathbf{j}'\rangle = u_{j_{m-1}} \wedge u_{j_{m-2}} \wedge \cdots \in \mathfrak{F}_m$, so $|\mathbf{j}\rangle_m = \psi^*_{j_m}|\mathbf{j}'\rangle_{m-1}$. We have

$$|\mathbf{j}\rangle_m = \rho_{j_m}(z)|\mathbf{j}'\rangle_{m-1} + z\xi, \tag{13.14}$$

where

$$\xi = u_{j_m-1} \wedge |\mathbf{j}'\rangle.$$

Now *both* terms on the right-hand side of (13.14) have lower energy than $|\mathbf{j}\rangle_m$. It is possible that $\xi = 0$ (if $j_{m-1} = j_m - 1$) but if $\xi \neq 0$ it has lower energy than $|\mathbf{j}\rangle_m$. So by our induction hypothesis

$$e^H|\mathbf{j}'\rangle_{m-1} = T|\mathbf{j}'\rangle_{m-1}, \qquad e^H\xi = \xi. \tag{13.15}$$

Now we have

$$e^H|\mathbf{j}\rangle_m = e^H\rho_{j_m}(z)|\mathbf{j}'\rangle_{m-1} + ze^H\xi = \rho_{j_m}(qz)e^H|\mathbf{j}'\rangle_{m-1} + ze^H\xi,$$
$$T|\mathbf{j}\rangle_m = T\rho_{j_m}(z)|\mathbf{j}'\rangle_{m-1} + zT\xi = \rho_{j_m}(qz)T|\mathbf{j}'\rangle_{m-1} + zT\xi,$$

and using (13.15) we obtain (13.13). So the theorem is proved.

Exercises

Exercise 13.1. Prove the claim in Section 3 that for Γ-ice with finitely many horizontal spins equal to $-$, the row transfer matrix is a finite product.

Exercise 13.2. Show that J_k and J_l commute for $k \neq -l$ in the proof of Theorem 13.5.

Exercise 13.3. Argue as stated in Section 5 that

$$\langle\mathbf{i}|J_k|\mathbf{j}\rangle$$

is well-defined despite the fact that J_{-k} on a basis vector gives an infinite sum of basis vectors.

Exercise 13.4. Prove that

$$e^{H_-(z;q)} = T_\Gamma(z;q)$$

in Theorem 13.6 by using adjoints.

Exercise 13.5. Check the remaining cases in the proof of Proposition 13.12.

Bibliography

Aggarwal, Amol, Alexei Borodin, Leonid Petrov, and Michael Wheeler (2023). "Free fermion six vertex model: symmetric functions and random domino tilings". in: *Selecta Math. (N.S.)* 29.3, Paper No. 36, 138.

Aggarwal, Amol, Alexei Borodin, and Michael Wheeler (2023a). "Colored fermionic vertex models and symmetric functions". in: *Comm. Amer. Math. Soc.* 3, pp. 400–630.

Aggarwal, Amol, Alexei Borodin, and Michael Wheeler (2023b). "Colored fermionic vertex models and symmetric functions". in: *Comm. Amer. Math. Soc.* 3, pp. 400–630.

Assaf, Sami and Anne Schilling (2018). "A Demazure crystal construction for Schubert polynomials". in: *Algebr. Comb.* 1.2, pp. 225–247.

Baxter, Rodney J. (1972). "One-dimensional anisotropic Heisenberg chain". in: *Ann. Physics* 70, pp. 323–337.

Baxter, Rodney J. (1982). *Exactly solved models in statistical mechanics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, pp. xii+486.

Bergeron, Nantel and Sara Billey (1993). "RC-graphs and Schubert polynomials". in: *Experiment. Math.* 2.4, pp. 257–269.

Bernšteĭn, I. N., I. M. Gel'fand, and S. I. Gel'fand (1976). "A certain category of $\mathfrak{g}$-modules". in: *Funkcional. Anal. i Priložen.* 10.2, pp. 1–8.

Biedenharn, L. C. and J. D. Louck (1989). "A new class of symmetric polynomials defined in terms of tableaux". in: *Adv. in Appl. Math.* 10.4, pp. 396–438.

Billey, Sara C., William Jockusch, and Richard P. Stanley (1993). "Some combinatorial properties of Schubert polynomials". in: *J. Algebraic Combin.* 2.4, pp. 345–374.

Björner, Anders and Francesco Brenti (2005). *Combinatorics of Coxeter groups*. vol. 231. Graduate Texts in Mathematics. Springer, New York, pp. xiv+363.

Blum, Talia (2025). *Bicolored bosonic solvable lattice models*.

Boos, Hermann, Frank Göhmann, Andreas Klümper, Khazret S. Nirov, and Alexander V. Razumov (2014). "Quantum groups and functional relations for higher rank". in: *J. Phys. A* 47.27, pp. 275201, 47.

Borel, Armand (1953). "Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts". in: *Ann. of Math. (2)* 57, pp. 115–207.

Borodin, Alexei and Michael Wheeler (2022). "Colored stochastic vertex models and their spectral theory". in: *Astérisque* 437, pp. ix+225.

Bourbaki, Nicolas (2002). *Lie groups and Lie algebras. Chapters 4–6*. Elements of Mathematics (Berlin). Translated from the 1968 French original by Andrew Pressley. Berlin: Springer-Verlag, pp. xii+300.

Brion, Michel (2005). "Lectures on the geometry of flag varieties". in: *Topics in cohomological studies of algebraic varieties*. Trends Math.. Birkhäuser, Basel, pp. 33–85.

Brubaker, Ben, Valentin Buciumas, and Daniel Bump (2019). "A Yang-Baxter equation for metaplectic ice. Appendix joint with Nathan Gray". in: *Commun. Number Theory Phys.*

13.1, pp. 101–148.
Brubaker, Ben, Valentin Buciumas, Daniel Bump, and Solomon Friedberg (2017). "Hecke Modules from Metaplectic Ice". in: *Selecta Math. (N.S.)*.
Brubaker, Ben, Valentin Buciumas, Daniel Bump, and Henrik P. A. Gustafsson (2019). *Colored Vertex Models and Iwahori Whittaker Functions*.
Brubaker, Ben, Valentin Buciumas, Daniel Bump, and Henrik P. A. Gustafsson (2020a). *Metaplectic Iwahori Whittaker functions and supersymmetric lattice models*.
Brubaker, Ben, Valentin Buciumas, Daniel Bump, and Henrik P. A. Gustafsson (2020b). "Vertex operators, solvable lattice models and metaplectic Whittaker functions". in: *Comm. Math. Phys.* 380.2, pp. 535–579.
Brubaker, Ben, Valentin Buciumas, Daniel Bump, and Henrik P. A. Gustafsson (2021). "Colored five-vertex models and Demazure atoms". in: *J. Combin. Theory Ser. A* 178, pp. 105354, 48.
Brubaker, Ben, Daniel Bump, Gautam Chinta, and Paul E. Gunnells (2012). "Metaplectic Whittaker functions and crystals of type B". in: *Multiple Dirichlet series, L-functions and automorphic forms*. vol. 300. Progr. Math.. Birkhäuser/Springer, New York, pp. 93–118.
Brubaker, Ben, Daniel Bump, and Solomon Friedberg (2011a). "Schur polynomials and the Yang-Baxter equation". in: *Comm. Math. Phys.* 308.2, pp. 281–301.
Brubaker, Ben, Daniel Bump, and Solomon Friedberg (2011b). *Weyl group multiple Dirichlet series: type A combinatorial theory*. vol. 175. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ.
Brubaker, Ben, Daniel Bump, and Solomon Friedberg (2016). "Matrix coefficients and Iwahori-Hecke algebra modules". in: *Adv. Math.* 299, pp. 247–271.
Brubaker, Ben, Daniel Bump, and Anthony Licata (2015). "Whittaker functions and Demazure operators". in: *J. Number Theory* 146, pp. 41–68.
Brubaker, Ben, Claire Frechette, Andrew Hardt, Emily Tibor, and Katherine Weber (2023). "Frozen pipes: lattice models for Grothendieck polynomials". in: *Algebr. Comb.* 6.3, pp. 789–833.
Brubaker, Ben and Andrew Schultz (2015). "The six-vertex model and deformations of the Weyl character formula". in: *J. Algebraic Combin.* 42.4, pp. 917–958.
Brubaker, Ben and Andrew Schultz (2018). "On Hamiltonians for six-vertex models". in: *J. Combin. Theory Ser. A* 155, pp. 100–121.
Buciumas, Valentin and Travis Scrimshaw (2022a). "Double Grothendieck polynomials and colored lattice models". in: *Int. Math. Res. Not. IMRN* 10, pp. 7231–7258.
Buciumas, Valentin and Travis Scrimshaw (2022b). "Quasi-solvable lattice models for Sp_{2n} and SO_{2n+1} Demazure atoms and characters". in: *Forum Math. Sigma* 10, Paper No. e53, 34.
Bump, Daniel (2013). *Lie groups*. Second. vol. 225. Graduate Texts in Mathematics. Springer, New York.
Bump, Daniel, Peter J. McNamara, and Maki Nakasuji (2014). "Factorial Schur functions and the Yang-Baxter equation". in: *Comment. Math. Univ. St. Pauli* 63.1-2, pp. 23–45.
Bump, Daniel and Slava Naprienko (2022). *Colored Bosonic Models and Matrix Coefficients*.
Bump, Daniel and Slava Naprienko (2025). *The Six-Vertex Yang-Baxter Groupoid*.
Bump, Daniel and Anne Schilling (2017). *Crystal bases*. Representations and combinatorics. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, pp. xii+279.

Casselman, W. and J. Shalika (1980). "The unramified principal series of p-adic groups. II. The Whittaker function". in: *Compositio Math.* 41.2, pp. 207–231.

Cauchy, A. L. (1815). "Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opér'es entre les variables qu'elles renferment". in: *J. École Polyt.* 10, pp. 29–112.

Chari, Vyjayanthi and Andrew Pressley (1991). "Quantum affine algebras". in: *Comm. Math. Phys.* 142.2, pp. 261–283.

Chen, William Y. C. and James D. Louck (1993). "The factorial Schur function". in: *J. Math. Phys.* 34.9, pp. 4144–4160.

Cheng, Shun-Jen and Weiqiang Wang (2012). "Dualities for Lie superalgebras". in: *Lie theory and representation theory*. vol. 2. Surv. Mod. Math.. Int. Press, Somerville, MA, pp. 1–46.

Corwin, L., Y. Ne'eman, and S. Sternberg (1975). "Graded Lie algebras in mathematics and physics (Bose-Fermi symmetry)". in: *Rev. Modern Phys.* 47, pp. 573–603.

Demazure, Michel (1974). "Désingularisation des variétés de Schubert généralisées". in: *Ann. Sci. École Norm. Sup. (4)* 7. Collection of articles dedicated to Henri Cartan on the occasion of his 70th birthday, I, pp. 53–88.

Deodhar, Vinay V. (1977). "Some characterizations of Bruhat ordering on a Coxeter group and determination of the relative Möbius function". in: *Invent. Math.* 39.2, pp. 187–198.

Drinfeld, V. G. (1987). "Quantum groups". in: *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986)*. Amer. Math. Soc., Providence, RI, pp. 798–820.

Edelman, Paul and Curtis Greene (1987). "Balanced tableaux". in: *Adv. in Math.* 63.1, pp. 42–99.

Felder, Giovanni and Alexander Varchenko (1996). "Algebraic Bethe ansatz for the elliptic quantum group $E_{\tau,\eta}(\mathrm{sl}_2)$". in: *Nuclear Phys. B* 480.1-2, pp. 485–503.

Fomin, Sergey and Curtis Greene (1998). "Noncommutative Schur functions and their applications". in: vol. 193. 1-3. Selected papers in honor of Adriano Garsia (Taormina, 1994), pp. 179–200.

Fomin, Sergey and Anatol N. Kirillov (1996). "The Yang-Baxter equation, symmetric functions, and Schubert polynomials". in: *Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993)*. vol. 153. 1-3, pp. 123–143.

Fomin, Sergey and Richard P. Stanley (1994). "Schubert polynomials and the nil-Coxeter algebra". in: *Adv. Math.* 103.2, pp. 196–207.

Frenkel, Edward and David Ben-Zvi (2004). *Vertex algebras and algebraic curves*. Second. vol. 88. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, pp. xiv+400.

Frenkel, I. B. (1981). "Two constructions of affine Lie algebra representations and boson-fermion correspondence in quantum field theory". in: *J. Functional Analysis* 44.3, pp. 259–327.

Fulton, William (1997). *Young tableaux*. vol. 35. London Mathematical Society Student Texts. With applications to representation theory and geometry. Cambridge University Press, Cambridge, pp. x+260.

Gold, Sarah, Elizabeth Milićević, and Yuxuan Sun (2024a). "Crystal chute moves on pipe dreams". Preprint, arXiv:2403.07204.

Gold, Sarah, Elizabeth Milićević, and Yuxuan Sun (2024b). "Crystal chute moves on pipe dreams". in: *Sém. Lothar. Combin.* 91B, Art. 91, 12.

Gray, Nathan (2017). *Metaplectic Ice for Cartan Type C*.

Gustafsson, Henrik P. A. and Carl Westerlund (2025). *The Schützenberger involution and colored lattice models*.

Haines, Thomas J., Robert E. Kottwitz, and Amritanshu Prasad (2010). "Iwahori-Hecke algebras". in: *J. Ramanujan Math. Soc.* 25.2, pp. 113–145.

Hamel, A. M. and R. C. King (2007). "Bijective proofs of shifted tableau and alternating sign matrix identities". in: *J. Algebraic Combin.* 25.4, pp. 417–458.

Heisenberg, W. (1928). "Zur Theorie des Ferromagnetismus". in: *Z. Physik* 49, pp. 619–636.

Hong, Jin and Seok-Jin Kang (2002). *Introduction to quantum groups and crystal bases*. vol. 42. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xviii+307.

Howe, Roger (2002). "Affine-like Hecke algebras and p-adic representation theory". in: *Iwahori-Hecke algebras and their representation theory (Martina-Franca, 1999)*. vol. 1804. Lecture Notes in Math.. Springer, Berlin, pp. 27–69.

Humphreys, James E. (1978). *Introduction to Lie algebras and representation theory*. vol. 9. Graduate Texts in Mathematics. Second printing, revised. Springer-Verlag, New York-Berlin, pp. xii+171.

Humphreys, James E. (1990). *Reflection groups and Coxeter groups*. vol. 29. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, pp. xii+204.

Humphreys, James E. (2008). *Representations of semisimple Lie algebras in the BGG category $\mathcal{O}$*. vol. 94. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xvi+289.

Ivanov, Dmitriy (2012). "Symplectic ice". in: *Multiple Dirichlet series, L-functions and automorphic forms*. vol. 300. Progr. Math.. Birkhäuser/Springer, New York, pp. 205–222.

Jimbo, Michio (1985). "A q-difference analogue of $U(\mathfrak{g})$ and the Yang-Baxter equation". in: *Lett. Math. Phys.* 10.1, pp. 63–69.

Jimbo, Michio (1986). "Quantum R matrix related to the generalized Toda system: an algebraic approach". in: *Field theory, quantum gravity and strings (Meudon/Paris, 1984/1985)*. vol. 246. Lecture Notes in Phys.. Springer, Berlin, pp. 335–361.

Jimbo, Michio and Tetsuji Miwa (1983). "Solitons and infinite-dimensional Lie algebras". in: *Publ. Res. Inst. Math. Sci.* 19.3, pp. 943–1001.

Joyal, André and Ross Street (1993). "Braided tensor categories". in: *Adv. Math.* 102.1, pp. 20–78.

Kac, V. G. (1977). "Characters of typical representations of classical Lie superalgebras". in: *Comm. Algebra* 5.8, pp. 889–897.

Kac, V. G. and J. W. van de Leur (1987). "Super boson-fermion correspondence". in: *Ann. Inst. Fourier (Grenoble)* 37.4, pp. 99–137.

Kac, Victor (1998). *Vertex algebras for beginners*. Second. vol. 10. University Lecture Series. American Mathematical Society, Providence, RI, pp. vi+201.

Kac, Victor G. (1990). *Infinite-dimensional Lie algebras*. Third. Cambridge University Press, Cambridge, pp. xxii+400.

Kac, Victor G., Ashok K. Raina, and Natasha Rozhkovskaya (2013). *Bombay lectures on highest weight representations of infinite dimensional Lie algebras*. Second. vol. 29. Advanced Series in Mathematical Physics. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, pp. xii+237.

Kashiwara, M. (1991). "On crystal bases of the Q-analogue of universal enveloping algebras". in: *Duke Math. J.* 63.2, pp. 465–516.

Kashiwara, Masaki (1993). "The crystal base and Littelmann's refined Demazure character formula". in: *Duke Math. J.* 71.3, pp. 839–858.

Kashiwara, Masaki (1995). "On crystal bases". in: *Representations of groups (Banff, AB, 1994)*. vol. 16. CMS Conf. Proc.. Amer. Math. Soc., Providence, RI, pp. 155–197.

Kashiwara, Masaki (2002). *Bases cristallines des groupes quantiques*. vol. 9. Cours Spécialisés [Specialized Courses]. Edited by Charles Cochet. Société Mathématique de France, Paris, pp. viii+115.

Kashiwara, Masaki and Toshiki Nakashima (1994). "Crystal graphs for representations of the q-analogue of classical Lie algebras". in: *J. Algebra* 165.2, pp. 295–345.

Kassel, Christian (1995). *Quantum groups*. vol. 155. Graduate Texts in Mathematics. Springer-Verlag, New York, pp. xii+531.

Kim, Henry H. and Kyu-Hwan Lee (2011). "Representation theory of p-adic groups and canonical bases". in: *Adv. Math.* 227.2, pp. 945–961.

Kirillov, A. N. and A. D. Berenstein (1995). "Groups generated by involutions, Gelfand-Tsetlin patterns, and combinatorics of Young tableaux". in: *Algebra i Analiz* 7.1, pp. 92–152.

Knutson, Allen and Ezra Miller (2005). "Gröbner geometry of Schubert polynomials". in: *Ann. of Math. (2)* 161.3, pp. 1245–1318.

Knutson, Allen and Gabe Udell (2023). "Interpolating between classic and bumpless pipe dreams". in: *Sém. Lothar. Combin.* 89B, Art. 89, 12.

Kojima, Takeo (2013). "Diagonalization of transfer matrix of supersymmetry $U_q(\widehat{sl}(M+1|N+1))$ chain with a boundary". in: *J. Math. Phys.* 54.4.

Korepin, V. E., N. M. Bogoliubov, and A. G. Izergin (1993a). *Quantum inverse scattering method and correlation functions*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, pp. xx+555.

Korepin, V. E., N. M. Bogoliubov, and A. G. Izergin (1993b). *Quantum inverse scattering method and correlation functions*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, pp. xx+555.

Korff, Christian (2013). "Cylindric versions of specialised Macdonald functions and a deformed Verlinde algebra". in: *Comm. Math. Phys.* 318.1, pp. 173–246.

Kostant, Bertram and Shrawan Kumar (1986). "The nil Hecke ring and cohomology of G/P for a Kac-Moody group G". in: *Proc. Nat. Acad. Sci. U.S.A.* 83.6, pp. 1543–1545.

Kulish, P. P. (1991). "Contraction of quantum algebras, and q-oscillators". in: *Teoret. Mat. Fiz.* 86.1, pp. 157–160.

Kuperberg, Greg (1996). "Another proof of the alternating-sign matrix conjecture". in: *Internat. Math. Res. Notices* 3, pp. 139–150.

Kuperberg, Greg (2002). "Symmetry classes of alternating-sign matrices under one roof". in: *Ann. of Math. (2)* 156.3, pp. 835–866.

Kwon, Jae-Hoon (2014). "Crystal bases of q-deformed Kac modules over the quantum superalgebras $U_q(\mathfrak{gl}(m|n))$". in: *Int. Math. Res. Not. IMRN* 2, pp. 512–550.

Lam, Thomas (2006). "A combinatorial generalization of the boson-fermion correspondence". in: *Math. Res. Lett.* 13.2-3, pp. 377–392.

Lam, Thomas, Seung Jin Lee, and Mark Shimozono (2021). "Back stable Schubert calculus". in: *Compos. Math.* 157.5, pp. 883–962.

Lascoux, Alain and Marcel-Paul Schützenberger (1982). "Polynômes de Schubert". in: *C. R. Acad. Sci. Paris Sér. I Math.* 294.13, pp. 447–450.

Lascoux, Alain and Marcel-Paul Schützenberger (1985). "Schubert polynomials and the Littlewood-Richardson rule". in: *Lett. Math. Phys.* 10.2-3, pp. 111–124.

Lascoux, Alain and Marcel-Paul Schützenberger (1990). "Keys & standard bases". in: *Invariant theory and tableaux (Minneapolis, MN, 1988)*. vol. 19. IMA Vol. Math. Appl.. Springer, New York, pp. 125–144.

Lenart, Cristian (2004). "A unified approach to combinatorial formulas for Schubert polynomials". in: *J. Algebraic Combin.* 20.3, pp. 263–299.

Lieb, Elliott H. (June 1967a). "Exact Solution of the F Model of An Antiferroelectric". in: *Phys. Rev. Lett.* 18 (24), pp. 1046–1048.

Lieb, Elliott H. (July 1967b). "Exact Solution of the Two-Dimensional Slater KDP Model of a Ferroelectric". in: *Phys. Rev. Lett.* 19 (3), pp. 108–110.

Lieb, Elliott H. (Oct. 1967c). "Residual Entropy of Square Ice". in: *Phys. Rev.* 162 (1), pp. 162–172.

Littelmann, Peter (1995a). "Crystal graphs and Young tableaux". in: *J. Algebra* 175.1, pp. 65–87.

Littelmann, Peter (1995b). "Paths and root operators in representation theory". in: *Ann. of Math. (2)* 142.3, pp. 499–525.

Lusztig, G. (1990). "Canonical bases arising from quantized enveloping algebras". in: *J. Amer. Math. Soc.* 3.2, pp. 447–498.

Lusztig, George (1985). "Equivariant K-theory and representations of Hecke algebras". in: *Proc. Amer. Math. Soc.* 94.2, pp. 337–342.

Lusztig, George (1989). "Affine Hecke algebras and their graded version". in: *J. Amer. Math. Soc.* 2.3, pp. 599–635.

Macdonald, I. G. (1991a). *Notes on Schubert Polynomials*. vol. 6. Laboratoire de combinatoire et d'informatique mathématique (LaCIM), Université du Québec á Montréal: Publications du LaCIM.

Macdonald, I. G. (1991b). "Schubert polynomials". in: *Surveys in combinatorics, 1991 (Guildford, 1991)*. vol. 166. London Math. Soc. Lecture Note Ser.. Cambridge Univ. Press, Cambridge, pp. 73–99.

Macdonald, I. G. (1992). "Schur functions: theme and variations". in: *Séminaire Lotharingien de Combinatoire (Saint-Nabor, 1992)*. vol. 498. Publ. Inst. Rech. Math. Av.. Univ. Louis Pasteur, Strasbourg, pp. 5–39.

Macdonald, I. G. (1995). *Symmetric functions and Hall polynomials*. Second. Oxford Mathematical Monographs. With contributions by A. Zelevinsky, Oxford Science Publications. The Clarendon Press, Oxford University Press, New York.

Majid, Shahn (2002). *A quantum groups primer*. vol. 292. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, pp. x+169.

Mason, S. (2009). "An explicit construction of type A Demazure atoms". in: *J. Algebraic Combin.* 29.3, pp. 295–313.

Matsumoto, Hideya (1964). "Générateurs et relations des groupes de Weyl généralisés". in: *C. R. Acad. Sci. Paris* 258, pp. 3419–3422.

Molev, Alexander (1998). "Factorial supersymmetric Schur functions and super Capelli identities". in: *Kirillov's seminar on representation theory*. vol. 181. Amer. Math. Soc. Transl. Ser. 2. Amer. Math. Soc., Providence, RI, pp. 109–137.

Morse, Jennifer and Anne Schilling (2016). "Crystal approach to affine Schubert calculus". in: *Int. Math. Res. Not. IMRN* 8, pp. 2239–2294.

Motegi, Kohei (2017a). "Dual wavefunction of the Felderhof model". in: *Lett. Math. Phys.* 107.7, pp. 1235–1263.

Motegi, Kohei (2017b). "Dual wavefunction of the symplectic ice". in: *Rep. Math. Phys.* 80.3, pp. 391–414.

Musson, Ian M. (2012). *Lie superalgebras and enveloping algebras*. vol. 131. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, pp. xx+488.

Nagle, J. F. (Dec. 1966). "Lattice Statistics of Hydrogen Bonded Crystals. I. The Residual Entropy of Ice". in: *Journal of Mathematical Physics* 7.8, pp. 1484–1491.

Naprienko, Slava (2022). *Integrability of the Six-Vertex model and the Yang-Baxter Groupoid*.

Naprienko, Slava (2024). "Free fermionic Schur functions". in: *Adv. Math.* 436, Paper No. 109413, 44.

Nirov, Khazret S. and Alexander V. Razumov (2017). "Quantum groups, Verma modules and q-oscillators: general linear case". in: *J. Phys. A* 50.30, pp. 305201, 19.

Okounkov, Andrei (1998). "On Newton interpolation of symmetric functions: a characterization of interpolation Macdonald polynomials". in: *Adv. in Appl. Math.* 20.4, pp. 395–428.

Okounkov, Andreĭ and Grigori Olshanski (1998). "Shifted Schur functions. II. The binomial formula for characters of classical groups and its applications". in: *Kirillov's seminar on representation theory*. vol. 181. Amer. Math. Soc. Transl. Ser. 2. Amer. Math. Soc., Providence, RI, pp. 245–271.

Okunḱov, A. and G. Olśhanskiĭ (1997). "Shifted Schur functions". in: *Algebra i Analiz* 9.2, pp. 73–146.

Pauling, Linus (1935). "The Structure and Entropy of Ice and of Other Crystals with Some Randomness of Atomic Arrangement". in: *Journal of the American Chemical Society* 57.12, pp. 2680–2684.

Perk, Jacques H. H. and Cherie L. Schultz (1981). "New families of commuting transfer matrices in q-state vertex models". in: *Phys. Lett. A* 84.8, pp. 407–410.

Reshetikhin, N. Yu., L. A. Takhtadzhyan, and L. D. Faddeev (1989). "Quantization of Lie groups and Lie algebras". in: *Algebra i Analiz* 1.1, pp. 178–206.

Ross, Leonard E. (1965). "Representations of graded Lie algebras". in: *Trans. Amer. Math. Soc.* 120, pp. 17–23.

Stanley, Richard P. (1984). "On the number of reduced decompositions of elements of Coxeter groups". in: *European J. Combin.* 5.4, pp. 359–372.

Stanley, Richard P. (1999). *Enumerative combinatorics. Vol. 2*. vol. 62. Cambridge Studies in Advanced Mathematics. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge University Press, Cambridge.

Sutherland, Bill (July 1967). "Exact Solution of a Two-Dimensional Model for Hydrogen-Bonded Crystals". in: *Phys. Rev. Lett.* 19 (3), pp. 103–104.

Tabony, Sawyer James (2010). *Deformations of characters, metaplectic Whittaker functions, and the Yang-Baxter equation*. Thesis (Ph.D.)–Massachusetts Institute of Technology. ProQuest LLC, Ann Arbor, MI, (no paging).

Tokuyama, Takeshi (1988). "A generating function of strict Gelfand patterns and some formulas on characters of general linear groups". in: *J. Math. Soc. Japan* 40.4, pp. 671–685.

Yamane, Hiroyuki (1994). "Quantized enveloping algebras associated with simple Lie superalgebras and their universal R-matrices". in: *Publ. Res. Inst. Math. Sci.* 30.1, pp. 15–87.

Yang, Yingzi (2025). *Closed Colored Models and Demazure Crystals*.

Zhang, Huafeng (2017). "Fundamental representations of quantum affine superalgebras and R-matrices". in: *Transform. Groups* 22.2, pp. 559–590.

Zhong, Chenyang (2022). "Stochastic symplectic ice". in: *Lett. Math. Phys.* 112.3, Paper No. 55, 47.

Zinn-Justin, P. (2009). *Six-Vertex, Loop and Tiling models: Integrability and Combinatorics*.

Index