The six vertex model, R-matrices, and quantum groups

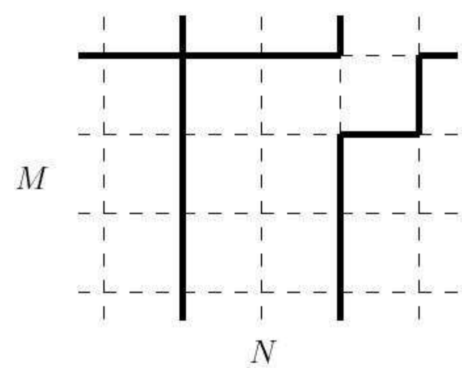
Jethro van Ekeren.

1 The six vertex model

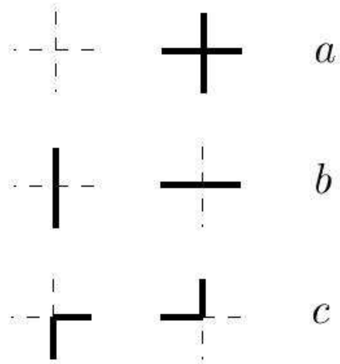
The six vertex model is an example of a lattice model in statistical mechanics. The data are

- A finite rectangular $M \times N$ grid, as shown below.
- An assignment of one of two states to each edge in the grid (subject to some restrictions). This is called a *configuration*.
- A number assigned to each vertex depending on the states of the adjacent edges. This is called the *weight* of the vertex.

We will distinguish the two states diagrammatically by making the edge thick or thin. Another common convention is to use arrows pointing in one direction or the other.



The restriction on the states is that at any vertex only one of the following arrangements may occur:



The complex parameters a, b and c are the weights of the vertices, the first two vertex configurations above have weight a, etc.

We also impose boundary conditions, the most commonly encountered types are:

- *Fixed boundary conditions*, in which we fix an arrangement of thick and thin lines around the edge of the grid and consider only interior configurations that are consistent with it.
- *Periodic boundary conditions*, in which we identify edges at opposite sides of the grid and require identified edges to be in the same state.

The configuration shown above satisfies periodic boundary conditions, and we will consider only periodic boundary conditions in this talk.

The weight of a configuration X is now defined to be the product of the weights of all the vertices in the grid:

$$w(X) = \sum_{\text{edges } e} w(e).$$

The *partition function* associated to the grid is the sum, over all possible configurations, of the weight of the configuration:

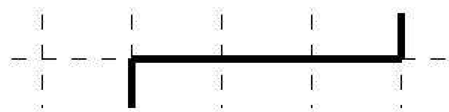
$$Z = \sum_{\text{config. } X} w(X).$$

The weight $w(X)$ of a configuration X is physically interpreted as the relative probability of the occurrence of X. The true probability of X is therefore $w(X)/Z$. Thus the partition function is a fundamental quantity in statistical mechanics, and the first step in the analysis of a statistical system is to calculate it.

In our case $Z = Z(M, N)$ may be quite complicated, but the asymptotics of Z as $M, N \to \infty$ turn out to be simpler. To see why we should expect such a simplification we introduce an object called the transfer matrix.

2 The transfer matrix

Let V_0 be the 2-dimensional complex vector space with the basis $\{\bullet, \circ\}$. Let $V = V_0^{\otimes N}$. A configuration of states on a row of N vertical edges may be identified with an element of the standard basis of V in the obvious way. In the diagram below for example, the lower and upper rows of edges are identified with $\circ \otimes \bullet \otimes \circ \otimes \circ \otimes \circ$ and $\circ \otimes \circ \otimes \circ \otimes \circ \otimes \bullet$ respectively.



We will construct an endomorphism $T : V \to V$ as follows: Let v and w denote lower and upper row configurations as shown. To each way of consistently filling in the intervening row of horizontal edges we assign its weight (ab^2c^2 in the case above). If α is the sum of the weights of all possible such configurations, then we define $T(v) = \alpha w$. This defines T as a $2^N \times 2^N$ matrix. It is called the *(row) transfer matrix.*

It is not hard to see that

$$Z = \mathrm{Tr}_V\, T^M.$$

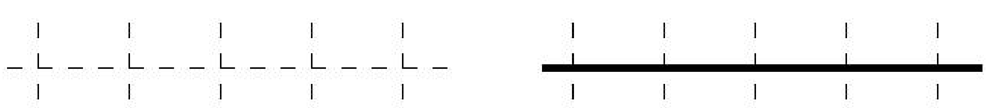
The trace of an operator is the sum of its eigenvalues. So if we fix N for the moment and let $M \to \infty$, we find

$$Z \sim \Lambda_{\max}^M \quad \text{as } M \to \infty.$$

where $\Lambda_{\max}$ is the largest eigenvalue of T. So our problem is to determine the largest eigenvalue of T; we would like to do this for arbitrary N, with the aim of letting $N \to \infty$ as well as M. Taking the simultaneous limit $M, N \to \infty$ is quite subtle; unless we let M grow faster than N, certain spurious effects can arise. We will not discuss these here.

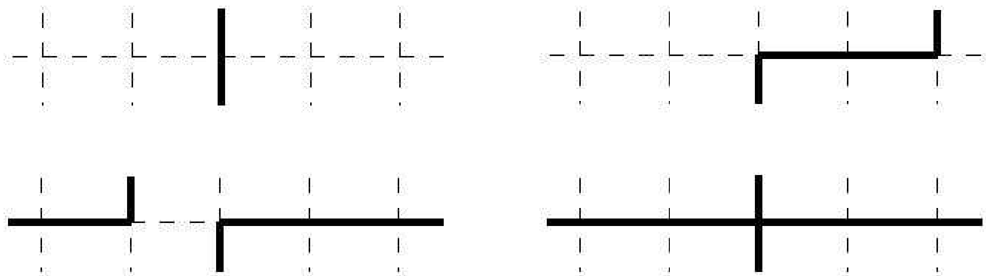
Let V_n be the subspace of V spanned by indecomposable tensors with n copies of $\bullet$, it has dimension $\binom{N}{n}$. By inspecting the allowed configurations at a vertex we see that the number of thick lines entering a row equals the number exiting, hence each space V_n is T-invariant. When we search for eigenvectors of T, it suffices to consider each V_n independently.

First let us tackle V_0. The possible configurations are



with weights a^N and b^N, so $T|_{V_0} = \Lambda = a^N + b^N$.

Now let us consider $n = 1$. There are 25 configurations, which come in the following four basic shapes



Let us identify a vector $g \in V_1$ with the function $g(x) : \{1, 2, \ldots N\} \to \mathbb{C}$ in the obvious way. The requirement that g be an eigenvector with eigenvalue Λ becomes

$$\begin{aligned}\Lambda g(x) = a^{N-1} b g(x) + &\sum_{y=x+1}^{N} a^{N+x-y-1} b^{y-x-1} c^2 g(y) \\ &+ \sum_{y=1}^{x-1} a^{x-y-1} b^{N+y-x-1} c^2 g(y) + a b^{N-1} g(x).\end{aligned}$$

The function $g(x)$ may be expanded as a linear combination of the functions $e^{2\pi i k x/N}$, where $k = 0, 1, \ldots N-1$. It turns out that these functions are precisely the N eigenvectors of T. To see this, we substitute $g(x) = z^x$, where $z^N = 1$. Using the formula for the summation of a geometric series we get

$$\begin{aligned}\Lambda z^x = a^N L(z) z^x &- a^{x-1} b^{N-x} c^2 z^{N+1}/(a - bz) \\ &+ b^N M(z) z^x + a^{x-1} b^{N-x} c^2 z/(a - bz),\end{aligned}$$

where

$$L(z) = \frac{ab + (c^2 - b^2) z}{a^2 - abz} \quad \text{and} \quad M(z) = \frac{a^2 - c^2 - abz}{ab - b^2 z}.$$

Since $z^N = 1$, the second and third terms cancel each other, leaving

$$\Lambda = a^N L(z) + b^N M(z).$$

The N roots of unity z give N eigenvalues Λ. To determine which eigenvalue is the largest we should do some more work, we could compute them all explicitly for example.

For $n \geq 2$ we may take the same approach, we write $g \in V_n$ as $g(S)$ mapping n-element subsets of $\{1, 2, \ldots N\}$ to $\mathbb{C}$. We may write down an equation for $g(S)$ directly. After some trial and error, it emerges that we may assume a particular form for $g(S)$ which yields an eigenvector of $T|_{V_n}$ in general. This general form is called the *Bethe ansatz*; we will not state it here, mentioning only that it leads to the following formula for the eigenvalues of $T|_{V_n}$:

$$\Lambda = a^N L_1 L_2 \cdots L_n + b^N M_1 M_2 \cdots M_n$$

where $L_i = L(z_i)$, $M_i = M(z_i)$, L and M as above, and $z_1, z_2, \ldots z_n$ are complex numbers which are to satisfy some further equations which we must determine.

The equations satisfied by the z_i are quite hard to determine directly. But we will now observe an interesting symmetry that occurs in the six vertex model that allows us to find the equations quite easily.

3 Commuting transfer matrices

To make the computations clearer we make a change of variables:

$$\begin{aligned} a &= \rho \sin(u+\eta), \\ b &= \rho \sin u, \\ c &= \rho \sin \eta. \end{aligned}$$

We also write

$$z = \frac{\sin v}{\sin(u+\eta)}$$

and similar equations with $z = z_i$, $v = v_i$ for $i = 1, 2, \ldots n$. We then have

$$L(z) = \frac{\sin(v-u+\eta)}{\sin(v-u)} \quad \text{and} \quad M(z) = \frac{\sin(v-u-\eta)}{\sin(v-u)}.$$

The transfer matrix depends on the parameters $T = T(a,b,c) = T(\rho,\eta,u)$. Suppose we fix ρ and η, so we have $T = T(u)$. We claim that

$$[T(u_1), T(u_2)] = 0 \quad \text{for all } u_1, u_2 \in \mathbb{C}.$$

We will prove this later. Now we wish to use this fact to derive the equations satisfied by the z_i of the last section.

Fix n as above. Since the $T(u)$ commute pairwise, they have a common set of eigenvectors, let P be the matrix whose columns are eigenvectors of $T(u)$. Note that P does not depend on u. If Λ is the eigenvalue corresponding to the r^{th} eigenvector, then

$$\Lambda = \Lambda(u) = (P^{-1}T(u)P)_{rr}.$$

The entries of T are polynomials in a, b and c, hence they are entire functions of $u \in \mathbb{C}$. Thus $\Lambda(u)$ is also an entire function of u.

But

$$\begin{aligned} \Lambda(u) &= a^N L_1 L_2 \cdots L_n + b^N M_1 M_2 \cdots M_n \\ &= \rho^N \sin^N(u+\eta) \prod_{i=1}^{n} \frac{\sin(v_i - u + \eta)}{\sin(v_i - u)} + \rho^N \sin^N u \prod_{i=1}^{n} \frac{\sin(v_i - u - \eta)}{\sin(v_i - u)}. \end{aligned}$$

The common denominator of the right hand side has zeros at $u = v_i$, so for $\Lambda(u)$ to be entire, the numerator must have zeros at $u = v_i$ too. This requirement yields the following n equations:

$$\left[\frac{\sin v_k}{\sin(v_k+\eta)}\right]^N = -\prod_{i=1}^{N} \frac{\sin(v_i - v_k + \eta)}{\sin(v_i - v_k - \eta)}.$$

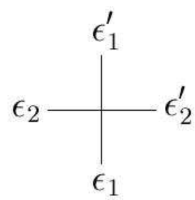
In principle we may solve these to find $z_1, z_2, \ldots z_n$, and then substitute to find the possible eigenvalues Λ.

4 The R-matrix

Our task now is to prove that $[T(u_1), T(u_2)] = 0$. To do this we introduce a new matrix $R : V_0 \otimes V_0 \to V_0 \otimes V_0$ defined as follows:

$$R(\epsilon_1 \otimes \epsilon_2) = \sum_{\epsilon_1', \epsilon_2'} w(\epsilon_1, \epsilon_2, \epsilon_1', \epsilon_2') \epsilon_1' \otimes \epsilon_2'$$

where $w(\epsilon_1, \epsilon_2, \epsilon_1', \epsilon_2')$ here denotes the weight of the following configuration:



With respect to the basis $\{\bullet \otimes \bullet, \bullet \otimes \circ, \circ \otimes \bullet, \circ \otimes \circ\}$, R becomes the matrix

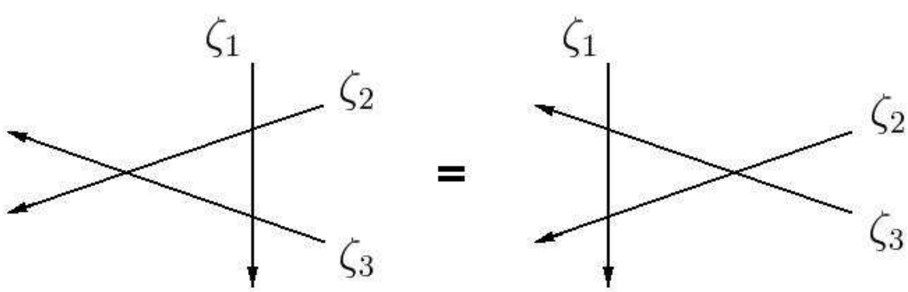
$$R = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ 0 & 0 & 0 & a \end{bmatrix} = \rho \begin{bmatrix} \sin(u+\eta) & 0 & 0 & 0 \\ 0 & \sin u & \sin \eta & 0 \\ 0 & \sin \eta & \sin u & 0 \\ 0 & 0 & 0 & \sin(u+\eta) \end{bmatrix}.$$

As in the last section we fix ρ and η, letting $R = R(u)$.

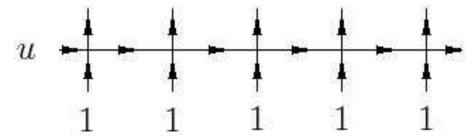
It may be verified by direct computation that for any nonzero complex numbers ζ_1, ζ_2 and ζ_3, $R(u)$ satisfies the *Yang-Baxter equation with parameter*:

$$R_{12}(\zeta_3/\zeta_2) R_{23}(\zeta_3/\zeta_1) R_{12}(\zeta_2/\zeta_1) = R_{23}(\zeta_2/\zeta_1) R_{12}(\zeta_3/\zeta_1) R_{23}(\zeta_3/\zeta_2),$$

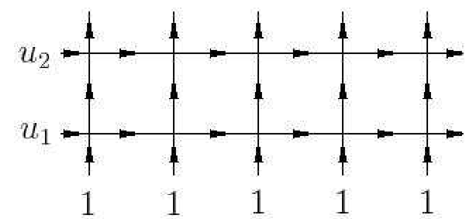
Here R_{12} means $R \otimes 1 : V_0 \otimes V_0 \otimes V_0 \to V_0 \otimes V_0 \otimes V_0$ and similarly for R_{23}. This equation has a nice geometric interpretation. We consider three crossing strands as shown below, and we associate a direction to each strand and also a complex number ζ_i. At the crossing of strands i and j suppose i comes before j in the clockwise ordering. Then we associate to their crossing, a copy of $R(u)$ where $u = \zeta_j/\zeta_i$. The Yang-Baxter equation then becomes the following equation, where a trace is taken over all internal edges:



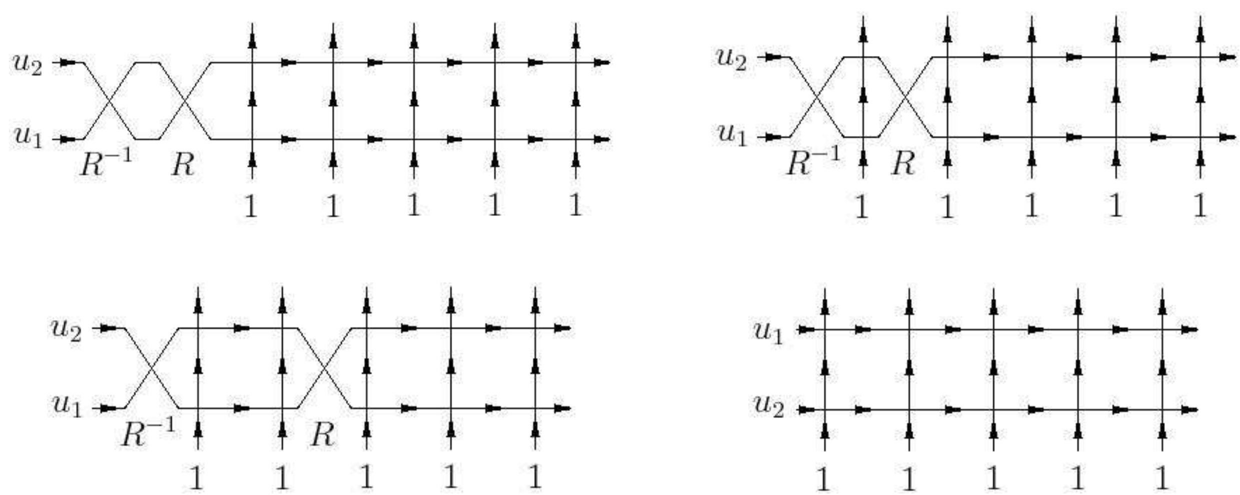
The transfer matrix $T(u)$ may be expressed in terms of the R-matrix as follows:



a copy of $R(u/1) = R(u)$ is placed at each crossing, and the trace is taken over all internal edges (including the edge that wraps around between the far left and right). As before, the arrows specify which lines are the inputs to the R-matrix and which are the outputs. The product of transfer matrices $T(u_1)T(u_2)$ is now



Now we introduce another copy of R, as well as its inverse, into this last diagram. We then manipulate using the Yang-Baxter equation to obtain, successively:



Thus $T(u_1)T(u_2) = T(u_2)T(u_1)$ as required.

5 Affine quantum groups

Let $\mathfrak{g}$ be a finite dimensional simple Lie algebra, and $U_q(\mathfrak{g})$ the associated quantum group. We have seen in previous lectures that if V and W are two $U_q(\mathfrak{g})$-modules then $V \otimes W$ and

$W \otimes V$ are isomorphic $U_q(\mathfrak{g})$-modules, but the isomorphism is not the usual 'swap'. There is a nontrivial intertwining map $R_{V,W} : V \otimes W \to W \otimes V$. If we fix a $U_q(\mathfrak{g})$-module V and put $R = R_{V,V}$, then R satisfies the (ordinary) *Yang-Baxter equation.*

So quantum groups are a source of R-matrices. However, the arguments above required an R-matrix that depended on a parameter (in addition to the usual q), and the R-matrices we obtain from representation of quantum groups do not naturally contain such parameters. To obtain R-matrices with parameter we need to consider *affine quantum groups.*

The affine quantum group associated to $\mathfrak{g}$ is just $U_q(\hat{\mathfrak{g}})$, i.e., it is constructed in the same way as $U_q(\mathfrak{g})$, but with the affine Kac-Moody algebra

$$\hat{\mathfrak{g}} = \mathfrak{g} \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}K \oplus \mathbb{C}d$$

in place of $\mathfrak{g}$. We also use a variant affine quantum group $U_q'(\hat{\mathfrak{g}}) = U_q(\hat{\mathfrak{g}}')$ which is constructed again in the same way, but using

$$\hat{\mathfrak{g}}' = \mathfrak{g} \otimes \mathbb{C}[t, t^{-1}] \oplus \mathbb{C}K$$

instead of $\hat{\mathfrak{g}}$.

Given a finite-dimensional representation V of $U_q(\mathfrak{sl}_2)$ it is possible to form an associated induced representation V^{aff} of $U_q(\hat{\mathfrak{sl}_2})$, it is infinite dimensional. V^{aff} becomes a representation of $U_q'(\hat{\mathfrak{sl}_2})$ automatically. Fix $\zeta \in \mathbb{C}$, neither zero nor a root of unity. V^{aff} factors to a $U_q'(\hat{\mathfrak{sl}_2})$-module V_ζ obtained by setting $t = \zeta$. As vector spaces (and in fact as representations of $U_q(\mathfrak{sl}_2) \subset U_q(\hat{\mathfrak{sl}_2})$), $V_\zeta \cong V$.

The construction is more subtle than the preceding paragraph would imply. There are analogous constructions in other types leading to finite dimensional modules V_ζ, but in general V_ζ is not irreducible as a representation of the underlying finite type algebra $\mathfrak{g}$.

Returning to the $\mathfrak{sl}_2$ case now: It turns out that for any ζ_1, ζ_2 we have an R-matrix

$$R(\zeta_1, \zeta_2) = R(\zeta_2/\zeta_1) : V_{\zeta_1} \otimes V_{\zeta_2} \to V_{\zeta_2} \otimes V_{\zeta_1},$$

which satisfies the Yang-Baxter equation with parameter.

The R-matrix we used above in relation to the six vertex model is essentially the same as the R-matrix obtained in this way from the standard representation $V = \mathbb{C}^2$ of $\mathfrak{sl}_2$. Other representations of $\mathfrak{sl}_2$ (and, with qualifications, representations of other affine quantum groups) yield new examples of R-matrices with parameter. Indeed, a large part of the initial motivation to study quantum groups came from the utility of such R-matrices in solving lattice models.