The six-vertex model

Rob Klabbers

October 23, 2013

Contents

1 introduction

The model under consideration in these lecture notes is the six-vertex model. This model was invented by Linus Pauling in 1935 and was intended to model water ice. Water ice has several fascinating properties, such as that the density in water ice is lower than the density in water

itself. Additionally, it appears that water ice has a very degenerate ground state, resulting in a characteristically high zero-point entropy (entropy at zero temperature). Pauling used this model to explain this zero-point entropy of water ice with surprising accuracy. His prediction of 0.806 cal·mole^{-1}·deg^{-1} for the discrepancy in the zero-point entropy is in excellent agreement with the experimental value 0.82 cal·mole^{-1}·deg^{-1}. The surprise lies in the fact that his model is actually a two dimensional model (meaning a model with two spatial dimensions), whereas water ice is of course a three dimensional system. Apparently, the structure in two dimensions lies very close to the actual structure and studying a three dimensional version of his model will probably not give many additional insights. Pauling based his model on just a few observations about water ice:

1. the oxygen atoms in water ice occupy the vertices on a number 4 lattice (a lattice where each site has on average 4 nearest neighbours). This was derived from the observation that the actual ice lattice has a tetrahedron as its building block, with an oxygen atom in the center and hydrogen atoms close to the tips of the tetrahedron.

2. on every edge connecting two neighbouring oxygen atoms, exactly one hydrogen atom presides.

3. these hydrogen atoms can either be close to an oxygen atom (through a covalent bond) or far away (through a hydrogen bond) (as in figure 1).

4. Water ice is electrically neutral locally, implying that around each oxygen atom, two hydrogen atoms are close and two are far away.

The last observation is called 'the ice-rule' in the literature. In these lecture notes, we will introduce the model Pauling invented and show some properties of this model. We will then go on to solve the model exactly by introducing transfer matrices, recovering some familiar equations using the Bethe Ansatz. The references for this derivation can be found in the final section of these notes.

2 The model

The model is based on a two dimensional rectangular lattice with L vertical lines and L' horizontal lines, with periodic boundary conditions in both conditions. This means one must view this lattice as being wound around a torus instead of lying on a plane. The (classic) variables of the model live on the $2LL'$ edges of this lattice and 'interact' at the vertices, hence the name vertex model. To model the positions of the hydrogen atom, it is sufficient to introduce a (classical) spin variable

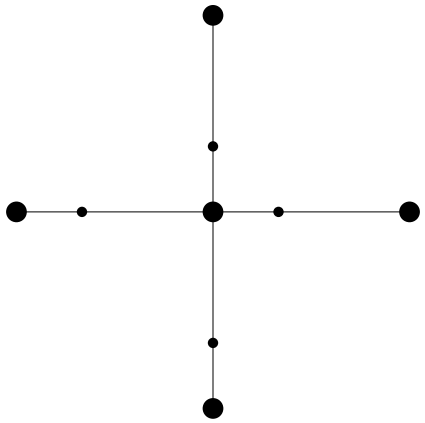


Figure 1: A typical situation in water ice: 5 oxygen atoms (the large dots) occupy the vertices of a lattice and the hydrogen atoms (the small dots) are either close or far away from the central oxygen atom.

$\epsilon_j \in \{+,-\}$, where j labels the edge. Whenever a hydrogen atom is situated in the right part of a horizontal edge or upper part of a vertical edge with label j, $\epsilon_j = +$ and when the atom is situated in the left part of a horizontal edge or the lower part a vertical edge, $\epsilon_j = -$ (compare with figure 1). We can now introduce the set $\{+,-\}^{2LL'}$ as the set of all configurations C of this lattice, even without specifying labels for each of the edges.

To actually create a statistical mechanical model from this set-up, we must introduce a weighing function $W : \{+,-\}^{2LL'} \to \mathbb{C}$, which assigns a Boltzmann weight to each of the possible configurations of the lattice. We will allow W to take complex values, since for the general model it is not necessary to restrict to real values. Also, it turns out that the hamiltonian of a one dimensional spin chain is simply an extension of the hamiltonian in our model to complex-valued Boltzmann weights[1]. Once we have defined a weighing function, we can already write down the partition function:

$$Z = \sum_C W(C), \tag{1}$$

where the summation is over all configurations $C \in \{+,-\}^{2LL'}$. Of course, the partition function is of great interest to us, because it allows us to find the correlation functions of our system, but also thermodynamic quantities such as pressure and energy.

In principle, this W can be any function, but to match with Pauling's observations , W has to

[1]For more information, see section 2.5.2 of [GRS], as given in the section References.

obey several rules. To ensure local (nearest neighbours) interactions, we will write W as

$$W(C) = \prod_n R^{\mu\nu}_{\rho\sigma}(C, n), \tag{2}$$

where the product runs over all vertices in the lattice and the R-matrix is simply a function $R : \{+, -\}^{2LL'} \times I \to \mathbb{C}$, where I is an index set to label the edges. The sub- and superscripts of R are defined as follows (compare with figure 2): for a given configuration C and a vertex n, the four edges joining at this vertex each have a value $+$ or $-$. The four greek letters ρ, σ, μ, ν indicate the value of the spin variable associated to the edge above, right of, left of and below of the vertex respectively. Enforcing the ice rule on the R-matrix reduces the possible configurations around a

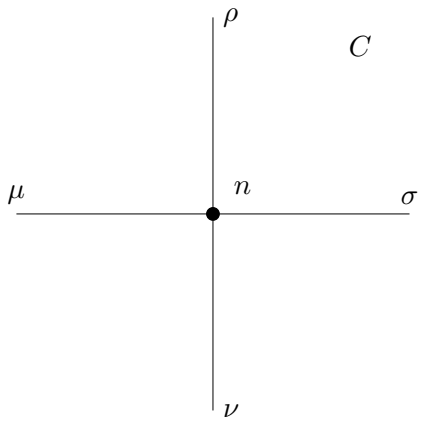
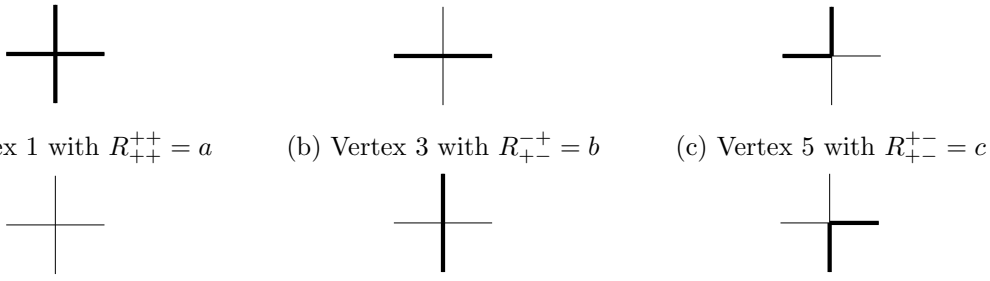


Figure 2: A vertex in the lattice with four values ρ, σ, μ, ν for the spin variables on the adjoining edges.

vertex from the possible $2^4 = 16$ down to 6, explaining the name six-vertex model. Notice that this implies that the value of the R-matrix on a vertex that is not allowed by the ice rule equals zero. The last restriction on R follows from *spin reversal symmetry*: in the absence of an external force, there is no preferred direction of spin, so one would expect that reversing all the spins in a state does not change the associated energy. This means that the six remaining vertex configurations fall apart in three groups of two, each with a fixed value (compare with table 3). In water ice, the values of the R-matrix for each of these three groups are equal, but we will consider a slightly more general case, which is summarized in table 3. In the line notation used in this table, an edge with $\epsilon_j = +$ is a filled line, while an edge with $\epsilon_j = -$ is an empty line. we will use this notation exclusively for the rest of these notes. In particular, we choose $a, b, c \in \mathbb{C}$ and they are all nonzero.

There is one more interesting thing to note here: spin reversal symmetry seemed to have restricted the weights of vertices 5 and 6 to be equal, but actually, this is not a new restriction. By looking at one horizontal line, one sees that vertex 5 allows a thick line to leave the line (to continue upwards), while vertex 6 introduces a new thick line to the line (coming from below). Since no other vertex has either of these properties (all other vertices either have a thick line on each horizontal edge or on no horizontal edge), it follows by periodicity that vertices 5 and 6 have to occur in equal numbers on each horizontal line. This means that they have to occur in equal numbers in every possible configuration C and thus their weights also have to occur in equal numbers in $W(C)$. We thus only have the combination $R^{+-}_{-+}R^{-+}_{+-}$ in $W(C)$, so assigning different values to each of them is meaningless; only the value of their product is meaningful. We can thus without loss of generality put them to be equal. This is the model as we will treat it in the following sections. Before we



(a) Vertex 1 with $R^{++}_{++} = a$ (b) Vertex 3 with $R^{-+}_{+-} = b$ (c) Vertex 5 with $R^{+-}_{+-} = c$

(d) Vertex 2 with $R^{--}_{--} = a$ (e) Vertex 4 with $R^{+-}_{-+} = b$ (f) Vertex 6 with $R^{-+}_{-+} = c$

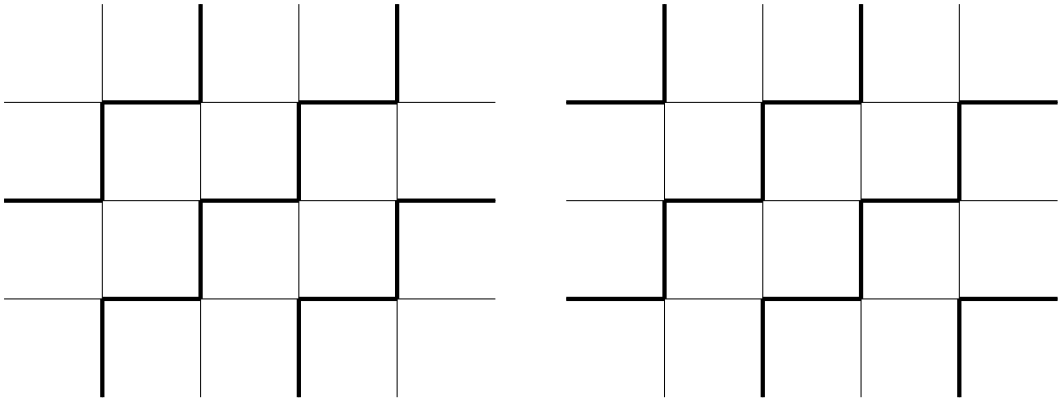
Figure 3: An overview of all the allowed vertex configurations and their respective weights

go into solving this model using the Bethe Ansatz and the transfer matrix method however, we will first show that one can already have some fun with this model and calculate some interesting features.

3 Low temperature expansion

The overall scaling of the partition function, as defined in equation 1, is irrelevant for the thermodynamic properties of the system, since all thermodynamic quantities are normalized with respect to the partition function. Additionally, under a rescaling of the three parameters, $a \to \frac{a}{k}$, $b \to \frac{b}{k}$, $c \to \frac{c}{k}$, for $k \in \mathbb{C}$ the partition function scales as

$$Z_{L,L'} \to \sum_C \frac{1}{k^{LL'}} W(C) = \frac{1}{k^{LL'}} Z_{L,L'}, \tag{3}$$



(a) Part of the ground state $C^{(+)}$ (b) Part of the ground state $C^{(-)}$

Figure 4: The two possible ground states

where now the subscripts on $Z_{L,L'}$ indicate that the size of our lattice. The factor $\frac{1}{k^{LL'}}$ arises because every nonzero $W(C)$ is a product of exactly LL' factors a, b and c. It follows that rescaling the parameters is not of physical significance and we can freely set $c = 1$ by a suitable rescaling. Consider the regime where a and b are real and write them as $a = e^{-\beta E_a}$, $b = e^{-\beta E_b}$, where $\beta = \frac{1}{k_B T}$ with k_B the Boltzmann constant and $E_a, E_b \in \mathbb{R}_{>0}$. Then we can regard E_a and E_b as energies associated with vertices of type 1 and 2 and type 3 and 4 respectively. In particular, the energy associated to $c = 1 = e^{\beta \cdot 0}$ has been set to zero by the rescaling. If we fix E_a and E_b, it is true that $\lim_{T\to 0} a(T) = 0$ and of course also $\lim_{T\to 0} b(T) = 0$. In this sense, the regime where $a, b \ll 1$ can be considered the low temperature regime of this model. For the rest of this section, we will assume $a, b \ll 1$.

Since a vertex of type 5 or 6 has zero energy, while all the other vertices have a positive energy, a configuration built out of only vertices 5 and 6 is a ground state of this model. In our model, there are two ground states, which are related by a global spin reversal. Part of these ground states are shown in table 4. Now, by choosing one particular edge on the lattice, we can classify these ground states as $C^{(+)}$ and $C^{(-)}$, according to the spin value of the chosen edge. Also, all other configurations can be classified in two sectors, by defining

$$d_i(C) = \text{the number of edges that are different from the ground state } C^{(i)}, \tag{4}$$

and dividing all configurations into two sectors: a sector S_+ consisting of all configurations C such that $d_+(C) < d_-(C)$ and a sector S_- consisting of all configurations C such that $d_+(C) > d_-(C)$.

Notice that since we will consider only the limit where $L, L' \to \infty$ the boundary between these sectors $d_+(C) = d_-(C)$ is irrelevant. Also, by spin reversal symmetry, for every state $C \in S_+$ there is a state $C' \in S_-$ such that C and C' are transformed into each other by a global reversal of spin and $W(C) = W(C')$. Therefore, we can write

$$Z_{L,L'} = 2 \sum_{C'} W(C'), \tag{5}$$

where now the summation is only over the sector S_+. In the large L, L'-limit, we can expand the partition function in powers of a and b. To find the first few orders, all one needs to do is find small alterations of the ground state $C^{(+)}$ using vertices with weight a and b. Finding the expansion up to eighth order in a and b has been an hand-in exercise; the result for the expansion is

$$2Z_{L,L'} = 1 + LL'a^2b^2 + LL'a^2b^2(a^2 + b^2) + \frac{LL'(LL'+1)}{2}a^4b^4 + LL'a^2b^2(a^2 + b^2). \tag{6}$$

A way to find this expansion is due to the following: as we will see in the next sections, the lines in the ground state pictures are a conserved quantity from row to row in the lattice. This means that any allowed configuration that deviates from the ground state must be the result of shifting the flow of one or some of the lines in the lattice (like building a dam in a river to change its flow). You can categorize these resulting states by the number of shifts you have performed on the ground state. By systematically going through all possible shifts, one finds the expansion as given in equation 6.

By trying to find the expansion for the partition function, it becomes clear very quickly that this is a pretty complicated model when viewed as a thermodynamical system. It is, however, possible to find the partition function exactly using a very neat trick. we will consider this in the following sections. From now on, we will again consider arbitrary values for a, b and c.

4 Transfer matrices

Up until this point, we have treated the six-vertex model as a two dimensional lattice in thermodynamic equilibrium. To solve this system, another approach will be very insightful. Consider one row of vertical edges on the lattice. If we define a time direction on the lattice, say upwards, each other row on the lattice can be viewed as intermediate stages in the time evolution of this first row. This is illustrated in figure 5: the lowest row represents the spin chain at $t = j$, and the rows above this one show intermediate stages at $t = j + 1$, $t = j + 2$ etc. Although this seems like a very innocent

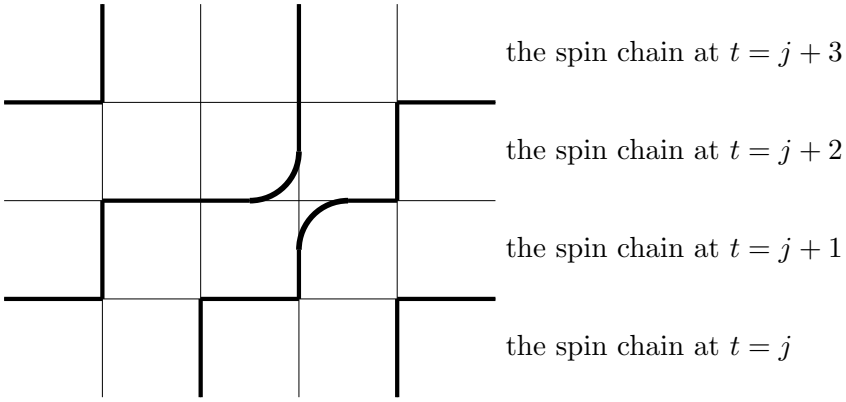


Figure 5: Part of the lattice in the interpretation as the time evolution of a spin chain.

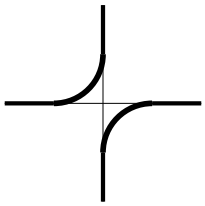


Figure 6: The new way to draw vertex 1, to make the flowing of magnon lines unambiguous.

change of view, its consequences are quite extensive, as we will show in due course. To make the new picture of the six-vertex model complete, we must choose some conventions. Additionally, we can give a physical interpretation to this picture: since we have chose the convention to let time flow upward, it is clear that also the lines in the lattice (cf. figure 5) flow upward. The lines can be viewed as the positions of some excitations of the classical spin, moving in discrete time; in analogy with what we know about spin chains we refer to these excitations as 'magnons'.

Now, the 'magnon' lines in this lattice have a very peculiar property: they never stop and are thus conserved from one row of vertical edges to the next. To see this, consider the six possible vertices again (in figure 3) and convince yourself that whenever a line flows toward a vertex, it also flows away from the vertex. Moreover, a line always approaches a vertex from the left or from below and always flows away from the vertex in the upper or right direction. Notice that in order to make the flow completely unambiguous, we must enforce on vertex 1 exactly the flow types as we stated in the previous: the line from the left continues on the line above the vertex and the line that comes from below the vertex continues to the right of the vertex. To make this also unambiguous in the pictures, from now on we will draw vertex 1 as pictured in figure 6.

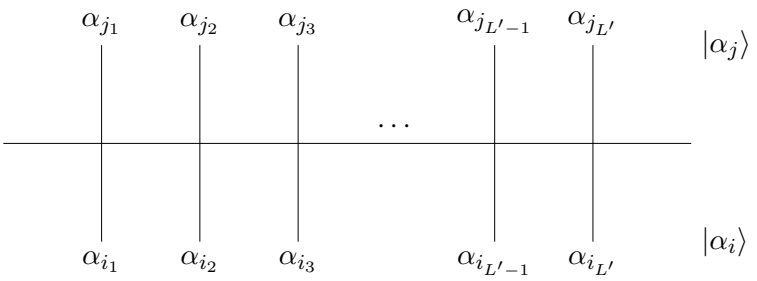


Figure 7: An in-state $|\alpha_i\rangle$ and an out-state $|\alpha_j\rangle$. The transfer from the in- to the out-state is governed by the transfer matrix T.

From this, it follows that lines can only move right or up and can thus never pass the same row by making a full turn. By periodic boundary conditions, it follows that the number of lines is conserved on each of the rows of vertical edges. This is a very special property of the six-vertex model and is for example not present in the eight-vertex model, an extension of the six-vertex model.

Until here, the transition from a two dimensional thermodynamic system to a one dimensional spin chain with time evolution has been fairly informal. Let us try now to formalize the above: Define a vector space V as $V = \mathbb{C}v_+ \oplus \mathbb{C}v_-$, the complex span of two vectors v_+ (which represents an edge with a magnon inhabiting it) and v_- (which represents an empty edge). We will define these to be orthonormal. This vector space V can be viewed as a vector space of one particular vertical edge, such that an arbitrary row state $|\alpha\rangle$ lives in the tensor product $\bigotimes^L V$ and the Hilbert space of this spin chain is simply $\mathcal{H} = \bigotimes^L V$. A basis for $\mathcal{H}$ is given by $\mathcal{B} = \left\{\otimes_{i=1}^L v_{p_i} | p_i \in \{+,-\}\right\}$, as the obvious extension of the basis of V. This is an orthonormal basis because we have defined v_+ and v_- to be orthonormal. These basis states can be written more conveniently as $|p_1, p_2, \ldots, p_L\rangle$. Additionally, we will need an operator to govern the time evolution of the spin chain: define $T : \mathcal{H} \to \mathcal{H}$ to be this operator and define its action on $\mathcal{H}$ by the action on the basis states of $\mathcal{H}$: let $|\alpha_i\rangle = |\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_L}\rangle$ and $|\alpha_j\rangle = |\alpha_{j_1}, \alpha_{j_2}, \ldots, \alpha_{j_L}\rangle$, then

$$\langle \alpha_i | \, T \, | \alpha_j \rangle = \sum_{\mu_1, \mu_2, \ldots, \mu_L} R^{\alpha_{j_1}\mu_1}_{\mu_L \alpha_{i_1}} R^{\alpha_{j_2}\mu_2}_{\mu_1 \alpha_{i_2}} \cdots R^{\alpha_{j_L}\mu_L}_{\mu_{L-1}\alpha_{i_L}}, \tag{7}$$

where the $\mu_i \in \{+,-\}$ for all $1 \leq i \leq L$.

This defines T as the operator that sums up the vertex weights of vertices between two rows of all possible ways to connect two row states using the line conventions from the previous paragraphs

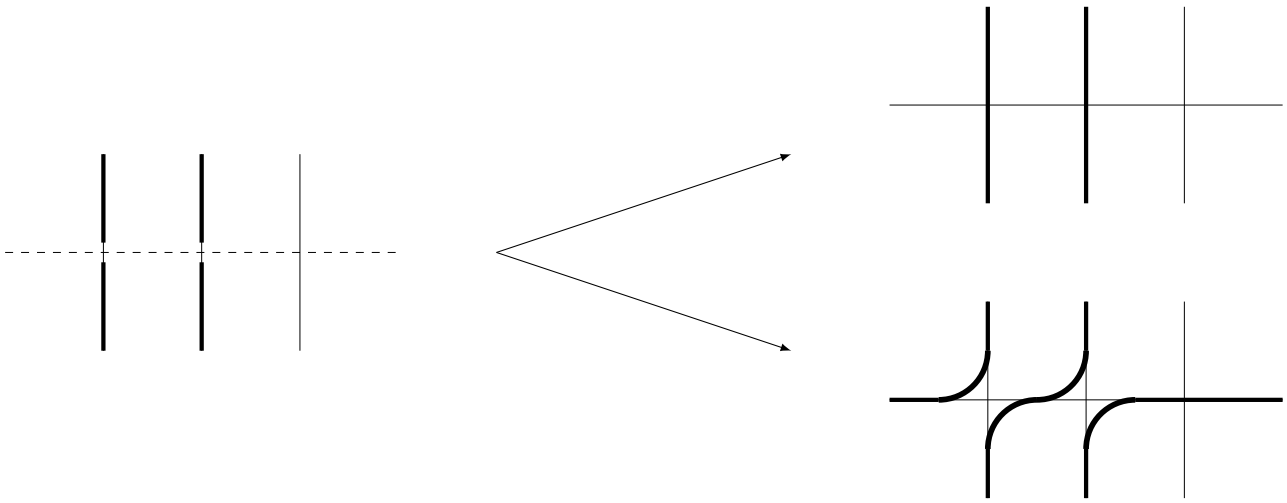


Figure 8: An example of an in- and out-state which can be connected using the six allowed vertices in two different ways. Here, $L = 3$.

(compare with figure 7). we will call it the *transfer matrix* for this system. Notice in particular that the in-state $|\alpha_i\rangle$ stands left of T in the above definition, contrary to the usual way of defining matrix elements of an operator. This will be useful in the following treatment. Notice also that it is not always the case that there is exactly one way to connect one row state to another row state; figure 8 shows an example where there are two ways of connecting the states. In general there may be more options. The fact that the transfer matrix allows for the transfer of a state to another state via a number of intermediate ones signals that this system is quantum mechanical in nature: it transfers a state into a linear combination of all states that can be created out of this first state using the six allowed vertices. One could see the quantum-mechanical nature also from the fact that by defining T as in equation 7, we have actually reinterpreted the weights from our original classical, statistical model as quantum-mechanical probabilities.

By line conservation, it follows that $\langle\alpha_i| T |\alpha_j\rangle = 0$ whenever $|\alpha_i\rangle$ and $|\alpha_j\rangle$ do not have an equal number of plusses (or particles). This means that T leaves subspaces of $\mathcal{H}$ spanned by states with a fixed number of plusses invariant, implying we can diagonalize T on these subspaces independently. We will use this fact greatly in the next section, so let us formalize this statement: to do so, we will introduce yet another way of writing the basis states: we can write the basis states as $|n_1, n_2, \ldots, n_m\rangle$, where the $n_i \in 1, 2, \ldots, L$ satisfy $n_1 < n_2 < \ldots < n_m$ and indicate for which i the p_i in the tensor product $\otimes_{i=1}^{L} v_{p_i}$ have $p_i = +$, so where a magnon resides on the row state. In

particular, we will write the basis state $\otimes_{i=1}^{L} v_-$ as $|0\rangle$. Now we can define the subspaces of fixed particle number as

$$S_m = \text{span}\left\{|n_1, n_2, \ldots, n_m\rangle \,|n_i \in \{1, 2, \ldots, L\} \text{ and } n_1 < n_2 < \ldots < n_m\right\}. \tag{8}$$

These subspaces have $\dim(S_m) = \binom{L}{m}$ and we can write $\mathcal{H} = S_0 \oplus S_1 \oplus \ldots \oplus S_L$ and $T(S_m) \subseteq S_m$ for each m.

Why are we interested in diagonalizing T? It follows from the definition of T that the partition function can be written as

$$Z_{L,L'} = \sum_{\alpha_1} \sum_{\alpha_2} \cdots \sum_{\alpha_{L'}} \langle\alpha_1| T |\alpha_2\rangle \langle\alpha_2| T \ldots |\alpha_{L'}\rangle \langle\alpha_{L'}| T |\alpha_1\rangle = \text{Tr}\left(T^{L'}\right), \tag{9}$$

by identifying that the sums run over a complete set of states of $\mathcal{H}$. Since the trace of $T^{L'}$ is completely determined by the eigenvalues of T, it follows that knowing the eigenvalues of T implies knowing the exact form of the partition function. Since all relevant quantities (such as energy and pressure) can be derived from the partition function, this means that knowing the eigenvalues of T essentially solves the system. This is exactly why in the following section, we will try to solve the eigenvalue equation for T using the Bethe Ansatz.

5 Bethe Ansatz

We want to find all the eigenvalues of T and already know that we can diagonalize T in each subspace S_m independently. Therefore, we must simply solve the equation $T|\psi\rangle = \Lambda|\psi\rangle$ in each subspace S_m. Notice that an arbitrary state in S_m can be written as

$$|\psi\rangle = \sum_{n_1, n_2, \ldots, n_m}^{*} f(n_1, n_2, \ldots, n_m) |n_1, n_2, \ldots, n_m\rangle, \tag{10}$$

where the star above the sum indicates that the summation variables satisfy $n_1 < n_2 < \ldots < n_m$. Furthermore, for each i, $n_i \in \{1, 2, \ldots, L\}$ and the function $f : \{1, 2, \ldots, L\}^m \to \mathbb{C}$ is arbitrary. Now the equation

$$\langle n_1, n_2, \ldots, n_m| T |\psi\rangle = \langle n_1, n_2, \ldots, n_m| \Lambda |\psi\rangle \tag{11}$$

implies that, by the orthonormality of the basis states $|n_1, n_2, \ldots, n_m\rangle$

$$\sum_{x_1, x_2, \ldots, x_m}^{*} \langle n_1, n_2, \ldots, n_m| T f(x_1, x_2, \ldots, x_m) |x_1, x_2, \ldots, x_m\rangle = \Lambda f(n_1, n_2, \ldots, n_m). \tag{12}$$

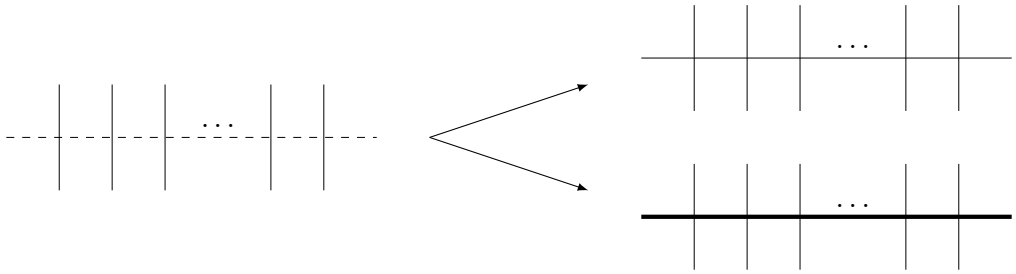


Figure 9: The two possible ways to connect the two rows of empty vertical lines.

For the function f we of course use the Bethe Ansatz, such that

$$f(n_1, n_2, \ldots, n_m) = \sum_{p \in S_m} A_p e^{i \sum_{i=1}^{m} k_{p(i)} n_i}, \tag{13}$$

where S_m is the symmetric group and $p(i)$ is an element $p \in S_m$ acting on the integer i. Notice that we choose $f = 1$ when $m = 0$. Finding the eigenvalue Λ now reduces to finding the matrix elements given on the left-hand side of equation 12. We will treat a few cases explicitly before tackling the case for general m.

5.1 S_0

The simplest case is of course the eigenvalue in the subspace S_0. S_0 is spanned by $|0\rangle$, thus the only matrix element left for computation is $\langle 0| T |0\rangle$. This is done by finding all the ways to connect a row of empty vertical lines to another row of empty vertical lines using the six allowed vertices.

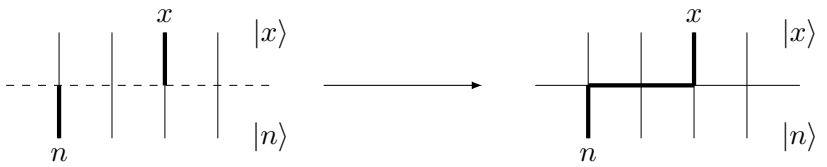
The possibilities are given in figure 9 and show that either all the vertices have weight a or all vertices have weight b. Thus the eigenvalue equation 12 reduces to

$$a^L + b^L = \Lambda, \tag{14}$$

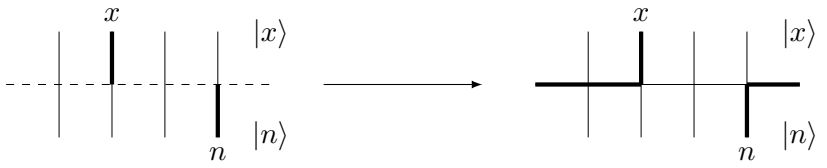
immediately giving the eigenvalue in this subspace with eigenvector $|0\rangle$.

5.2 S_1

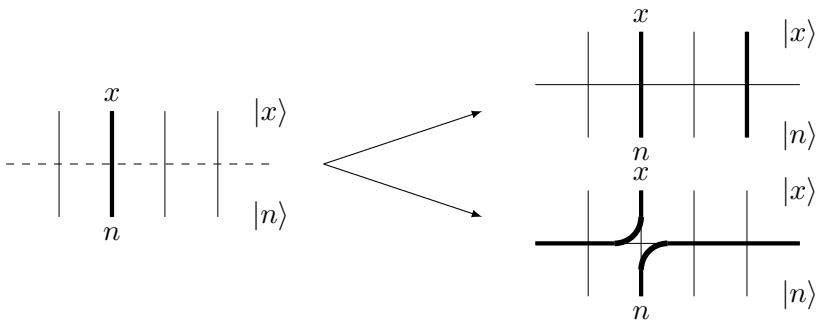
The subspace S_1 still is relatively easy – We will see that the real structure of calculating the matrix elements will emerge starting from $m = 2$ – but already shows the general approach for finding the matrix elements. S_1 is spanned by the L vectors $|n\rangle$, with $n \in \{1, 2, \ldots, L\}$. The Bethe Ansatz gives $f(n) = e^{ikn}$ for some $k \in \mathbb{C}$, but let us write $f(n) = z^n$ for brevity. For the matrix elements



(a) $n < x$



(b) $n > x$



(c) $n = x$

Figure 10: The relevant cases in the subspace S_1.

$\langle n| Tz^x |x\rangle$ from the eigenvalue equation 12, in general three cases can occur, which are displayed in table 10. From now on, to avoid overcomplicating the pictures, we will draw them without the "..." to indicate that there are actually other vertical lines in between the ones drawn; the reader will have to infer this from the context. Notice that due to periodic boundary conditions, cases (a) and (b) are really not different, but performing the relevant sums is easier when the cases are split into $n < x$ and $n > x$. From the pictures in table 10, we can immediately write down the complete

eigenvalue equation and compute the sums explicitly, recognizing geometric series as we go:

$$\begin{aligned}
\Lambda z^n &= \sum_x \langle n| T z^x |x\rangle \\
&= (ba^{L-1} + ab^{L-1})z^x + \sum_{x>n} c^2 b^{x-n-1} a^{L-2-(x-n-1)} z^x + \sum_{x<n} c^2 a^{x-n-1} b^{L-1-x+n} z^x \\
&= (ba^{L-1} + ab^{L-1})z^x + c^2 a^{L+n-1} b^{-n-1} \sum_{x=n+1}^{L} \left(\frac{bz}{a}\right)^x + c^2 a^{n-1} b^{L-n-1} \sum_{x=1}^{n-1} \left(\frac{bz}{a}\right)^x \\
&= (ba^{L-1} + ab^{L-1})z^x + c^2 a^{L+n-1} b^{-n-1} \sum_{x=0}^{L-n-1} \left(\frac{bz}{a}\right)^{x+n+1} + c^2 a^{n-1} b^{L-n-1} \left(\frac{bz}{a}\right) \sum_{x=0}^{n-2} \left(\frac{bz}{a}\right)^x \\
&= (ba^{L-1} + ab^{L-1})z^x + c^2 a^{L-2} z^{n+1} \sum_{x=0}^{L-n-1} \left(\frac{bz}{a}\right)^x + c^2 a^{n-2} b^{L-n} z \sum_{x=0}^{n-2} \left(\frac{bz}{a}\right)^x \\
&= (ba^{L-1} + ab^{L-1})z^x + c^2 a^{L-2} z^{n+1} \frac{1 - \left(\frac{bz}{a}\right)^{L-n}}{1 - \frac{bz}{a}} + c^2 a^{n-2} b^{L-n} z \frac{1 - \left(\frac{bz}{a}\right)^{n-1}}{1 - \frac{bz}{a}} \\
&= (ba^{L-1} + ab^{L-1})z^x + \frac{c^2 b^{L-1} a^{L-1} z^n + 1 - c^2 a^{n-1} b^{L-n} z^{L+1} + c^2 a^{n-1} b^{L-1} z - c^2 b^{L-1} z^n}{a - bz} \\
&= a^L L(z) z^n + b^L M(z) z^n + \frac{a^{n-1} b^{L-n} c^2 (z - z^{L+1})}{a - bz}, \qquad (15)
\end{aligned}$$

where we have defined the functions

$$L(z) = \frac{ab + (c^2 - b^2)z}{a^2 - abz} \qquad (16)$$

and

$$M(z) = \frac{a^2 - c^2 - abz}{ab - ab^2 z}. \qquad (17)$$

The right-hand side of the eigenvalue equation 15 has two terms which are nice and one that is not: the first two terms depend on n only through the factor z^n, which also occurs on the left-hand side. The third term, however, depends on n through various factors and will thus give an n-dependent contribution to Λ. This is very unwanted, because the translational invariance of the spin chain should be reflected in the eigenvalues and eigenvectors, meaning Λ cannot depend on n. Therefore, the third term must vanish. It is easy to see that this requires $z = z^{L+1}$ and thus we have $z^L = 1$. This has L solutions: $z = e^{ik}$ with $k \in \{\frac{2\pi l}{L} | 1 \leq l \leq L\}$. This defines L eigenvectors given by

$$|\psi_l\rangle = \sum_{n=1}^{L} e^{i\frac{2\pi l}{L}} |n\rangle , \qquad (18)$$

for $1 \leq l \leq L$, with eigenvalue $\Lambda = a^L L(z) + b^L M(z)$. Since $\dim(S_1) = \binom{L}{1} = L$, we have found a complete set of eigenvectors for this subspace.

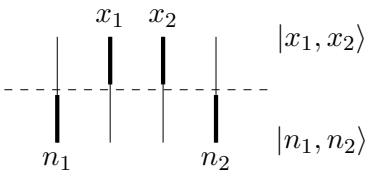


Figure 11: An example of an in-state and an out-state for which $\langle n_1, n_2| T |x_1, x_2\rangle$ vanishes

5.3 S_2: some structure appears

This case has all the features that occur at general m, but it is still simple enough to deal with it quite explicitly. We will not, however, list all possibilities in a table using the graphical representation as for the S_1-case: this is simply too cumbersome. We can however list the cases in another way: the eigenvalue equation 12 for this case becomes:

$$\sum_{x_1,x_2}^{*} \langle n_1, n_2| T f(x_1, x_2) |x_1, x_2\rangle = \Lambda f(n_1, n_2), \tag{19}$$

where the star indicates that $x_1 < x_2$. Notice that there are many particular cases for which the matrix element vanishes, see for example figure 11. This means that $\langle n_1, n_2| T |x_1, x_2\rangle$ is only non-zero when the integers n_1, n_2, x_1, x_2 satisfy either $n_1 \leq x_1 \leq n_2 \leq x_2$ or $x_1 \leq n_1 \leq x_2 \leq n_2$. This is a general feature which can also be extended to the general m case. To write down the left-hand side of 19, it is convenient and enlightening to first define some functions: let

$$D(x, n) = \begin{cases} \frac{a}{c} & \text{if } n = x \\ ca^{n-x-1} & \text{if } n > x \end{cases} \tag{20}$$

and

$$E(n, x) = \begin{cases} \frac{b}{c} & \text{if } n = x \\ cb^{x-n-1} & \text{if } n < x \end{cases}. \tag{21}$$

This defines D only for $n \geq x$ and E only for $n \leq x$. They can be interpreted as follows: for $D(x, n)$ we have (compare with the upper part of table 12): when $n > x$, it takes into account a line coming from the lower row onto the horizontal one (this has weight c) and going up (this has again weight c and is visible on the left of the left picture). It then counts all the vertices of type 2 between this line and the next incoming line from below. When $n = x$, it recognizes that there is no c weight coming from the vertex at the input line, no vertex of weight c is associated to the output line and there are no vertices of type 2 between the input and output line. This leaves just one vertex of type 1. We can follow the same line of reasoning for E: when $n < x$, E counts a c for

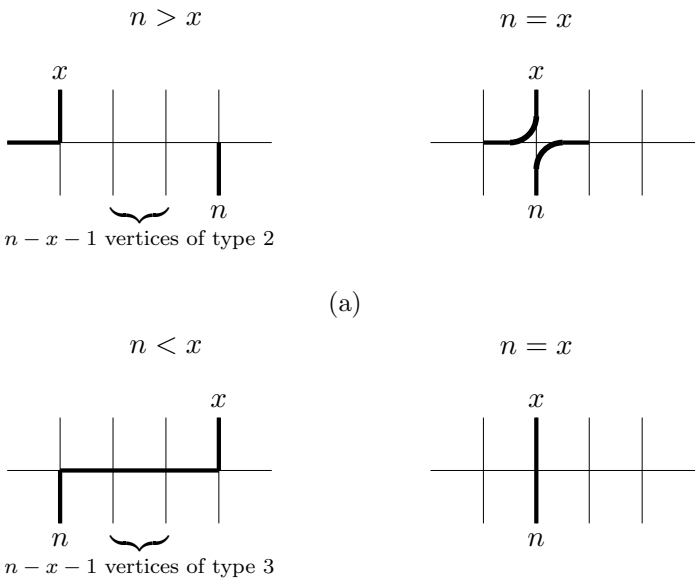


Figure 12: Pictures that show what $D(x, n)$ (in figure (a)) and $E(n, x)$ (in figure (b)) mean.

the vertex at the incoming line, then $x - n - 1$ vertices of type 3 with weight b and finally a vertex of type 6 for the vertex where the line leaves the horizontal line. Also, when $n = x$, it recognizes that there are no vertices with weight c and only one type-4 vertex with weight b. Notice that both these functions seem to overcorrect for the number of weight-c vertices (because in the cases where you expect two weight-c vertices, it counts only one and in the cases where you expect no weight-c vertex, it counts -1 vertices of this type), but we will see shortly that this is just convenient way of defining these two functions.

Using this definition, it is much easier to write down the eigenvalue equation 19, where for now we will leave the function f unspecified:

$$\begin{aligned}\Lambda f(n_1, n_2) = &\sum_{x_1=n_1}^{n_2} \sum_{\substack{x_2=n_2 \\ x_2 \neq x_1}}^{L} a^{n_1-1} E(n_1, x_1) D(x_1, n_2) E(n_2, x_2) c a^{L-x_2} f(x_1, x_2) \\ &+ \sum_{x_1=1}^{n_1} \sum_{\substack{x_2=n_1 \\ x_2 \neq x_1}}^{n_2} b^{x_1-1} D(x_1, n_1) E(n_1, x_2) D(x_2, n_2) c b^{L-n_2} f(x_1, x_2). \end{aligned} \tag{22}$$

The first summation corresponds to the case $n_1 \leq x_1 \leq n_2 \leq x_2$, while the second summation corresponds to $x_1 \leq n_1 \leq x_2 \leq n_2$. Notice that the restriction $x_1 = x_2$ under the second summation sign in each of the terms has an effect on just one term in each of them, namely the one for which

$x_1 = x_2 = n_1$ in the first summation and $x_1 = x_2 = n_2$ in the second summation. Let us now take as an ansatz for f that $f(n_1, n_2) = A_{12} z_1^{n_1} z_2^{n_2}$, which is only half of the usual Bethe Ansatz for the case where $m = 2$, but it saves us from writing a lot of different terms. Additionally, let us define the following:

$$L_i = L(z_i), \qquad M_i = M(z_i), \qquad \rho_i = \rho(z_i) = \frac{c^2 z_i}{a^2 - abz_i}. \tag{23}$$

Then, by computing the sums explicitly, equation 22 becomes,

$$\Lambda A_{12} z_1^{n_1} z_2^{n_2} = A_{12} \Bigg\{ a^{n_1} \big(\underbrace{L_1 a^{n_2-n_1} z_1^{n_1}}_{1} + \underbrace{M_1 b^{n_2-n_1} z_1^{n_2}}_{2} \big) \cdot \big(\underbrace{L_2 a^{L-n_2} z_2^{n_2}}_{1,2} - \rho_2 b^{L-n_2} z_2^L \big) - \underbrace{a^{L+n_1-n_2} b^{n_2-n_1} (z_1 z_2)^{n_2}}_{2} + (\rho_1 a^{n_1} + \underbrace{M_1 b^{n_1} z_1^{n_1}}_{1,2}) \cdot \big(\underbrace{L_2 a^{n_2-n_1} z_2^{n_1}}_{2} + \underbrace{M_2 b^{n_2-n_1} z_2^{n_2}}_{1} \big) b^{L-n_2} - \underbrace{a^{n_2-n_1} b^{L-n_2+n_1} (z_1 z_2)^{n_1}}_{2} \Bigg\}, \tag{24}$$

where we have categorized the terms into three classes (1, 2 and without a number) for future use. This equation is a pretty complicated mess, but using the same argumentation as in the case $m = 1$ we will be able to find the eigenvalue: translational invariance dictates again that the only terms on the right-hand side that can contribute are the ones where the only n_1- and n_2-dependence sits in the factor $z_1^{n_1} z_2^{n_2}$, such that it is divided out to leave us with a n_1- and n_2-independent Λ. This already gives us some idea about the eigenvalue: it contains the term $(a^L L_1 L_2 + b^L M_1 M_2)$ (which corresponds to the terms with a 1 underneath them), as one can verify by looking carefully at the terms in equation 24. Now, all other terms should vanish, but at the moment, they really don't. This is not surprising, since we know that the Ansatz we used for f is not the usual Bethe Ansatz. Let us first write out all other terms and see if we can conclude how we should modify f to let them vanish: The terms with a 2 under them give

$$A_{12}(M_1 L_2 - 1)((z_1 z_2)^{n_2} a^{L+n_1-n_2} b^{n_2-n_1} + (z_1 z_2)^{n_1} a^{n_2-n_1} b^{L-n_2+n_1}), \tag{25}$$

while all the other terms combine to give

$$A_{12} \left\{ a^{n_1} b^{L-n_2} (L_2 a^{n_2-n_1} + M_2 b^{n_2-n_1} z_2^{n_2}) \rho_1 \;\; - a^{n_1} b^{L-n_2} (L_1 a^{n_2-n_1} + M_1 b^{n_2-n_1} z_1^{n_2}) \rho_2 z_2^L \right\}.$$

By carefully looking at equation 25, one sees that the term

$$((z_1 z_2)^{n_2} a^{L+n_1-n_2} b^{n_2-n_1} + (z_1 z_2)^{n_1} a^{n_2-n_1} b^{L-n_2+n_1}) \tag{26}$$

is symmetric under exchange of $z_1 \leftrightarrow z_2$. If we therefore postulate a new Ansatz for f, namely $f(n_1, n_2) = A_{12} z_1^{n_1} z_2^{n_2} + A_{21} z_1^{n_2} z_2^{n_1}$, it follows that the whole term in equation 25 vanishes when

$$A_{12}(M_1 L_2 - 1) + A_{21}(M_2 L_1 - 1) = 0. \tag{27}$$

This implies the equation

$$\frac{A_{12}}{A_{21}} = -\frac{s_{21}}{s_{12}}, \tag{28}$$

where we have defined

$$M_1 L_2 - 1 = \frac{-c^2 s_{12}}{(a - bz_1)(a - bz_2)} \qquad s_{12} = 1 - 2\Delta z_2 + z_1 z_2 \qquad \Delta = \frac{a^2 + b^2 - c^2}{2ab}. \tag{29}$$

For those who have studied the Heisenberg XXZ-model, these equations should already look familiar. The functions s_{21} and s_{12} are exactly the scattering matrices in the XXZ-model, where Δ is the anisotropy parameter. We will go deeper into this similarity when we consider the case for general m.

Let us take a closer look at equation 26: if we include into this equation the terms that arise from extending f, we see it vanishes precisely when

$$\begin{aligned} A_{12} z_2^L &= A_{21}, \qquad z_2^L = -\frac{M_1 L_2 - 1}{M_2 L_1 - 1} = -\frac{s_{12}}{s_{21}}, \\ A_{21} z_1^L &= A_{12}, \qquad z_1^L = -\frac{M_2 L_1 - 1}{M_1 L_2 - 1} = -\frac{s_{21}}{s_{12}}. \end{aligned} \tag{30}$$

Following the Bethe Ansatz, we must impose $z_1 \neq z_2$, which implies here that if both these sets of equations have L solutions, in total we have $\binom{L}{2}$ possible combinations for the complex tuple (z_1, z_2). Since these z_j define a complex number of the form e^{ik_j}, we see that these equations give rise to $\binom{L}{2}$ choices for the momenta k_1 and k_2. Also, the momenta depend only on the weights a, b and c through the parameter Δ, which already hints that there is a larger symmetry hidden beneath this model[2].

In conclusion, we have found that the eigenvalues for the subspace S_2 are the same for all the eigenvectors and equal $\Lambda = (a^L L_1 L_2 + b^L M_1 M_2)$. The eigenvectors, which are specified by the momenta k_1 and k_2, follow as the solutions to the equations in 30, after plugging the value for $\frac{A_{12}}{A_{21}}$ as it follows from equation 28 into these equations. This can in principle be done and give rise to $\binom{L}{2}$ solutions for the eigenvectors, meaning we could construct a complete basis out of eigenvectors for S_2.

[2]This symmetry will be looked into in the lecture notes on the TQ-construction.

5.4 S_m: the case for general m

With the machinery and observations from the case $m = 2$, it is fairly straightforward to write down the eigenvalue equation when m is arbitrary. Firstly, it is very important to observe that the restriction that the nonzero matrix elements $\langle n_1, n_2, \ldots, n_m| T |x_1, x_2, \ldots, x_m\rangle$ must satisfy either $n_1 \leq x_1 \leq n_2 \leq x_2 \leq \ldots \leq n_m \leq x_m$ or $x_1 \leq n_1 \leq x_2 \leq n_2 \leq \ldots \leq x_m \leq n_m$. This greatly reduces the number of options we have to consider. Define for notational brevity $D(x_i, n_j) = D_{ij}$ and $E(n_i, x_j) = E_{ij}$. Now the eigenvalue equation follows as:

$$\Lambda f(n_1, n_2, \ldots, n_m) = \sum_{x_1=n_1}^{n_2} \sum_{x_2=n_2}^{n_3} \ldots \sum_{x_m=n_m}^{L} a^{n_1-1} \cdot E_{11} D_{12} E_{22} D_{23} \ldots E_{(m-1)(m-1)} D_{(m-1)m} E_{mm} \cdot \cdot c a^{L-x_m} f(x_1, x_2, \ldots, x_m) + \sum_{x_1=1}^{n_1} \sum_{x_2=n_1}^{n_2} \ldots \sum_{x_m=n_{m-1}}^{n_m} b^{x_1-1} \cdot D_{11} E_{12} D_{22} \ldots E_{(m-1)m} D_{mm} c b^{L-n_m} f(x_1, x_2, \ldots, x_m), \quad (31)$$

where in the summations we impose that $x_i \neq x_j$ whenever $i \neq j$. In principle, these summations can be carried out exactly once again. The general procedure is still the same as in the case where $m = 2$: we postulate a first Ansatz for f, namely $f(n_1, n_2, \ldots, n_m = A_{1,2,\ldots,m} z_1^{n_1} z_2^{n_2} \ldots z_m^{n_m}$. If we define $R_j(x, x') = L_j a^{x-x'} z_j^x + M_j b^{x-x'} z_j^{x'}$, we can write the first summation from the right-hand side of equation 31 as:

$$A_{1,2,\ldots,m} \left\{ a^{x_1} R_1 x_1, x_2 R_2(x_2, x_3) \ldots R_{m-1}(x_{m-1}, x_m) \cdot \left(L_m a^{L-x_m} z_m^{n_m} - \rho_m b^{L-x_m} z_m^L \right) - \text{ correction terms } \right\}, \quad (32)$$

where the correction terms arise to subtract terms in which $x_1 = x_2, x_2 = x_3, \ldots,$ or $x_{n-1} = x_m$. The second summation gives a similar term and is given by

$$A_{1,2,\ldots,m} \left\{ (\rho_1 a^{x_1} + M_1 b^{x_1} z_1^{x_1}) \cdot R_2(x_1, x_2) R_3(x_2, x_3) \ldots R_m(x_{m-1}, x_m) b^{L-x_m} - \text{ correction terms} \right\}. \quad (33)$$

Combining from these two expressions all the terms for which the n_j-dependence arises purely through the combination $z_1^{n_1} z_2^{n_2} \ldots z_m^{n_m}$ gives us again the eigenvalue, since all other terms are excluded by translational invariance. This gives

$$\Lambda = a^L L_1 L_2 \ldots L_m + b^L M_1 M_2 \ldots M_m, \quad (34)$$

which is a straightforward generalization of the eigenvalue in the case where $m = 2$. To make the other terms cancel, we can add other terms to f to end up with the actual Bethe Ansatz

$f(n_1, n_2, \ldots, n_m) = \sum_{p \in S_m} A_p e^{i \sum_{i=1}^{m} k_{p(i)} n_i}$. In the end, the terms containing factors $(z_j z_{j+1})_{j+1}^{x}$ or $(z_j z_{j+1})_{j}^{x}$ cancel provided that

$$s_{p(j)p(j+1)} A_{p(1),p(2)\ldots,p(m)} + s_{p(j+1)p(j)} A_{p(1),\ldots,p(j+1),p(j),\ldots,p(m)} = 0, \tag{35}$$

for all permutations $p \in S_m$ and all $j \in \{1, 2, \ldots, m-1\}$. Notice that $p(i)$ is just the permutation p acting on the integer i. The remaining terms, which (just as in the case where $m = 2$) all contain a ρ_j for some j, cancel only when

$$z_{p(1)}^{L} = \frac{A_{p(1),p(2),\ldots,p(m)}}{A_{p(2),p(3),\ldots,p(m),p(1)}}. \tag{36}$$

Of course, it is not obvious why these equations should have a solution, not in the least because the number of equations greatly outnumbers the number of variables. We can at least give a solution to equation 35, namely

$$A_{p(1),p(2)\ldots,p(m)} = \epsilon_p \prod_{1 \leq i \leq j \leq m} s_{p(i),p(j)}, \tag{37}$$

where $\epsilon_p = 1$ when p is an even permutation and $\epsilon_p = -1$ when p is an odd permutation. We can insert this solution into equation 36, to produce

$$z_{p(1)}^{L} = (-1)^{m-1} \prod_{j=2}^{m} \frac{s_{p(j),p(1)}}{s_{p(1)p(j)}}. \tag{38}$$

Notice that the right-hand side of this equation is actually symmetric under exchange of any two elements of the set $\{p(2), p(3), \ldots, p(m)\}$, thus we can just greatly restrict the number of independent equations here: it is only relevant where the element 1 gets mapped to by the permutation p. This has exactly m possibilities, thus we are left with

$$z_{j}^{L} = (-1)^{m-1} \prod_{\substack{i=1 \\ i \neq j}}^{m} \frac{s_{i,j}}{s_{j,i}}. \tag{39}$$

This gives us m equations for the z_i and can in principle be solved. From these, the coefficients A_p can be found, such that the complete solution is found.

6 Connection to the Heisenberg XXZ-model

Although the previous derivation of the solution of the six-vertex model is interesting in itself, it gets much more interesting when seen in another context. The Bethe Ansatz was originally used

to solve the Heisenberg XXX- and XXZ-models. The eigenvector equation for general m for the XXZ-model is given by

$$e^{ik_j L} = (-1)^{m-1} \prod_{\substack{i=1 \\ i \neq j}}^{m} \frac{s_{i,j}}{s_{j,i}}, \tag{40}$$

where k_j is the quasimomentum of the magnon excitations. By the identification $z_j = e^{ik_j}$, which is completely consistent with z_j being part of the Bethe Ansatz, we see that the eigenvector equations for the XXZ- an six-vertex model completely coincide! As seen before, the anisotropy parameter Δ from the XXZ-model also gets a new interpretation in the six-vertex model, as a fraction of the weights a, b and c. So we see that we can use all our knowledge of using the Bethe Ansatz on the XXZ-model to solve the six-vertex model, which is great. This was actually precisely the way that Lieb, who first solved the six-vertex model in 1967, found the eigenvectors!

Also, the fact that the eigenvector equation only depends on the weights through Δ is curious. We see that for every Δ, we get a set of eigenvectors from the XXZ-model. But this means that as long as $\Delta(a, b, c) = \Delta(a', b', c')$, the eigenvectors are the same. Thus the solutions of the six-vertex model obey an extra symmetry compared to the XXZ-model. This should give rise to conserved quantities, but we will not go into this here. This will be discussed later in the course.

7 Conclusion

We have seen that we can use the method of transfer matrices to solve the six-vertex model. This was done using the Bethe Ansatz. It turns out that the eigenvectors of this model can actually be found using the already known results of the XXZ-model, since the eigenvector equations for both models are identical. This is a curious fact, not in the least because there also seems to be a degeneracy in the solutions for the six-vertex model, because the solutions only depend on the weights (which really are free parameters) through the combination $\Delta = \frac{a^2+b^2-c^2}{2ab}$. This extra symmetry hints to much deeper connections, which we will discuss later in the course.

Using the transfer matrix method turned out very conveniently in this case, but this will not be so in a more general case: we could only use transfer matrics in this way because we could prove line conservation. This made it possible to treat the eigenvalue problem per subspace of fixed line number m, which greatly simplified the problem. Line conservation is unfortunately not a general property of vertex models: for example, the eight-vertex model, a simple extension of the six-vertex

model, does not have this property. Luckily, we can use a new method, called the TQ-construction, to solve that model. The TQ-construction is very powerful and we will look into it in the next set of notes.

8 References

To create these lecture notes and the accompanying presentation we have used the following references:

[Bax] Baxter: Exactly Solved Models in Statistical Mechanics. Academic Press (1982). Especially sections 8.1 to 8.5. In particular, the derivation in these notes for $m = 2$ and general m follow Baxter's derivation closely.

[GRS] Gmez, Ruiz-Altaba, Sierra: Quantum Groups in Two-Dimensional Physics. Cambridge University Press (1996). In particular section 1.3, but also section 2.5.2 for more general remarks.

[JM] Jimbo, Miwa: Algebraic Analysis of Solvable Lattice Models. American Mathematical Society (1995). Chapter 2

[Tol] Toledano Laredo: Exactly solved models of statistical mechanics. Lecture notes (2011). Chapter 8 contains a very detailed derivation of the eigenvalues and eigenvectors for general m.