Algebraic Bethe Ansatz

Ben Werkhoven

universiteit Utrecht

16 November 2013

Contents

1 Introduction

Up untill now we have set up a nice way of describing the XXX and the XXY model for the Heisenberg spin-chain. One of the main results of the previous lectures was the Yang-Baxter equation:

$$\sum_{j_1,j_2,j_3} R^{k_1 k_2}_{j_1 j_2}(\lambda) R^{j_1 k_3}_{i_1 j_3}(\lambda+\mu) R^{j_2 j_3}_{i_2 i_3}(\mu) = \sum_{j_1,j_2,j_3} R^{j_2 j_3}_{i_2 i_3}(\mu) R^{j_1 k_3}_{i_1 j_3}(\lambda+\mu) R k_1 k_2{}_{j_1 j_2}(\lambda) \qquad (1)$$

This is one of the most important results and a lot of the computations are based on this equation. Realizing the importance of this equation has led people to set up a new approach to describe systems like the XXZ model, a new method to discover all the relevant properties: the Algebraic Bethe Ansatz. The idea of the Algebraic Bethe Ansatz is as follows: instead of taking the Yang-Baxter equation as a result, we take it as an assumption. The Ansatz part of the Algebraic Bethe Ansatz then consists of the Yang-Baxter equation and the mathematical structure underlying this equation.

This may seem strange at first, but especially for physicists this is not a new phenomenon. This approach is reminiscent of the Action Principle. The Action Principle was stated because people saw that in classical mechanics, where we assume Newton's equations, the action is always minimized. Instead of taking this as a result, they subsequently took it as an assumption. What they got was a whole new way of describing classical mechanics (even recovering Newton's equation) which could also be applied to other areas of physics besides just classical mechanics. This is analogous for the Algebraic Bethe Ansatz.

It turns out that this new method can be applied to other systems than just the XXX and XXY models. This is the most important property of the Algebraic Bethe Ansatz and the reason why it is worth studying. The purpose of this article is to first give a formal, general definition of a Yang-Baxter Algebra. This is the mathematical structure the whole technique is based on, so it is important to give a clear definition. Afterwards we will focus on more specific cases, six-vertex like models and after that the six-vertex model itself. We will see that this new method gives us the same results as before, most importantly the Bethe equations, and thus verifying our new method. However, while reading this article one must keep in mind that the basic structure outlined in the beginning of the article can be applied to other systems (the XYZ model for example), which is its main feature.

2 Recap of previous lectures

Before we define a our new method, it is useful to have a quick recap of the results and definitions of the previous chapters. Most importantly is the R-matrix, which tells us which Boltzmann weight to give to which vertex. To describe the six-vertex model we use vertices with 4 arms. Each arm has an index which can be either plus or minus. We can make a nice diagramatic representation of this: This gives us 16

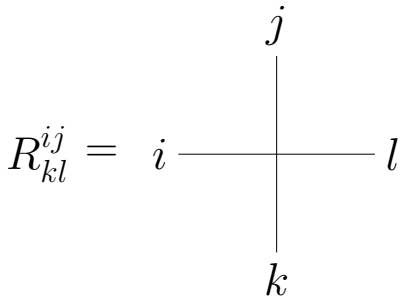

Figure 1: diagramatic representation of the R-matrix.

possible vertices, allowing us to write the R-matrix as an 4 by 4 matrix, each entry telling us what Boltzmann weight to assign to each vertex. For the representation of the R-matrix we choose the basis $e_i \otimes e_j$ where $i = \pm$ amd $e_+ = (1, 0)$ and $e_- = (0, 1)$. If we want to describe the XXZ model we assign non-zero Boltzmann weight to the following 6 vertices:

- $R^{++}_{--} = R^{--}_{--} = a$,
- $R^{+-}_{-+} = R^{-+}_{+-} = b$,
- $R^{+-}_{+-} = R^{-+}_{-+} = c$.

All the other R-matrix entries are 0.

The R-matrix obeys equation (1). With an R-matrix we can construct the monodromy matrix defined as follows:

$$T_a = R_{a,L} R_{a,L-1} ..... R_{a,1}, \tag{2}$$

where the subscripts are not indices, but the labels of the space they act on. $R_{a,L}$ acts on the auxiliairy space a and the local Hilbert space V_L associated to the lattice site L. The local Hilbert space is assigned to the vertical arms of the vertex and the auxiliairy space to the horizontal line. For the XXZ model the auxiliairy space and the Hilobert space are $\mathbb{C}^2$.

This has also a nice diagramtic representation:

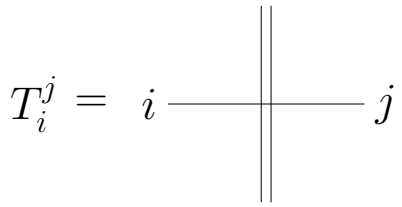

Figure 2: diagramatic representation of the monodromy matrix

The double line represents some physical Hilbert space, the global Hilbert space of the model we are considering. Using the Yang-Baxter equation (1), it is not difficult to show that T_a satisfies a similar equation (which has been derived in the previous chapter):

$$R_{ab}(\lambda - \mu)T_a(\lambda)T_b(\mu) = T_b(\mu)T_a(\lambda)R_{ab}(\lambda - \mu). \tag{3}$$

Where the subscripts denote the auxiliairy space the matrices work on.

Next we define the transfer matrix as the trace of the monodromy matrix:

$$t = Tr_a(T_a). \tag{4}$$

Here the trace is taken over the auxiliairy space V_a.

The goal of this method is to find the eigenstates and eigenvalues of the transfer matrix. As mentiuoned before, this is because the Hamiltonian and the transfer matrix commuute and therefore share the same eigenstates.

3 Yang-Baxter Algebra

3.1 General defenition

Before we begin searching for the eigenstates we must define the Yang-Bexter algebra $\mathscr{A}$. This algebra will be the underlying mathematical structure of our method. The Yang-Baxter algebra is a bialgebra. This means that it is, first of all, an associative algebra. An associative algebra is a vector space with an associative, bilinear operation:

$$m : \mathscr{A} \otimes \mathscr{A} \to \mathscr{A}. \tag{5}$$

With the associativity relation

$$m(m \otimes \mathbb{1}) = m(\mathbb{1} \otimes m). \tag{6}$$

The matrix multiplication will play the role of this operation, for which we know that it is associative and bilinear.

A bialgebra is an associative algebra with an addional map, called the comultiplication Δ. The comultiplication maps the algebra to the tensor product of the algebra with itself, whilst preserving the algebraic structure of the algebra:

$$\Delta : \mathscr{A} \to \mathscr{A} \otimes \mathscr{A}. \tag{7}$$

The co-product obeys its own associativity relation, called coassociativity:

$$(\Delta \otimes \mathbb{1})\Delta = (\mathbb{1} \otimes \Delta)\Delta. \tag{8}$$

This relation can also be put in words. After applying Δ once, we have got a tensor product of $\mathscr{A}$ with itself. If we then want to apply Δ again, we have a choice on which $\mathscr{A}$ we apply the second Δ. The co-associativity realtion then tells us that the result is the same if we apply Δ on the left $\mathscr{A}$ (of (7)) as when we apply Δ to the right $\mathscr{A}$ (of (7)).

The co-product may be new to the reader, but we already have encountered it before, namely when we defined the total spin operator $\sum_i S_z$. This operator is constructed by taking the properties of S_z, which acts on $\mathbb{C}^2$, and extending it so that it works on the tensorproduct of $\bigotimes_{i=1}^{L} \mathbb{C}^2$. This is exactly what the coproduct does, but then applied L times.

Now that we have a definition of a bi-algebra, we can define the Yang-Bxter algebra. The Yang-Baxter algebra is a bialgebra, where the monodromy matrices $T_i^j(\lambda)$ $(i, j \in \{1, ..., n\}; \lambda \in \mathbb{C})$ will play the role of the generator of the algebra, satisfiying the Yang-Baxter equation (3). The value of n will be determined by the dimension of the auxiliairy space. Equation (3) is a commutation realtion between different monodromy matrices, from which we can conclude that the R-matrix (actually the elements of the R-matrix) will serve as the structure constants (or rather structure functions since these entries will depend on λ). The R-matrix is an $n^2 \times n^2$ invertable matrix. The entries of the monodromy matrix are operators, which are a functions of λ, acting on some Hilbert sapce, which is two-dimensional in the six-vertex model. Constructing a representation of the Yang-Baxter algebra is equivalent to constructing an integrable vertex model.

The coproduct can take many forms, as long as it saitfies (8), but in the case of the Ynag-Baxter algebra Δ acts as follows (which is different from the way Δ acts for the total spin operator):

$$\Delta : T_i^j(\lambda) \to \sum_k T_i^k(\lambda) \otimes T_k^j(\lambda), \tag{9}$$

with the correspong diagramatic representation shown on the next page.

From the diagramatic representation we see that Δ links two monodromy matrixes in all possible ways and sums over all of them. It should be clear that (8) is satisfied by this because of the associativity of the tensorproduct.

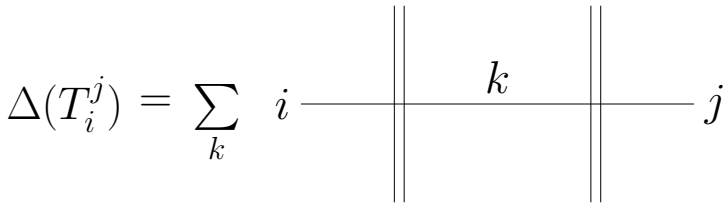

Figure 3: diagramatic representation of the action of Δ on the monodromy matrix

3.2 Focusing on six-vertex type R-matrices

Now that we have defined our algebra, we can focus on a more specific case. From now on we will assume the following form of our R-matrix:

$$\begin{pmatrix} a(\lambda) & 0 & 0 & 0 \\ 0 & b(\lambda) & c(\lambda) & 0 \\ 0 & c(\lambda) & b(\lambda) & 0 \\ 0 & 0 & 0 & a(\lambda) \end{pmatrix}. \tag{10}$$

Clearly the R-matrix of the 6-vertex model is of this form, but for now we have no reason to focus only on the six-vertex model. Basically, the computations of this section are relevant for any model which has an R-matrix of this form. A nice example is the spin chain with inhomogeneities.. In this model the translation symmetry is broken by translating each site by a different amount. However, with this model we can still find the Hamiltonian and its eigenvalues (although is may be cumbersome). This clearly shows why this model is worth studying. Note that he XYZ model doesn't fall in this category, since the lower left and upper right entries are non-zero for the XYZ model. The Yang-Baxter algebra can be used to describe the XYZ model, but it requires different calculations than the ones in this chapter.

By fixing the size of the R matrix we have fixed the value of $n\,(=2)$, which corresponds to fixing our auxuliairy space to $V_a = \mathbb{C}^2$. T_i^j is then a 2×2 matrix:

$$\begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix}. \tag{11}$$

Here we have defined $T_+^+ = A(\lambda)$, $T_-^+ = B(\lambda)$, $T_+^- = C(\lambda)$ and $T_-^- = D(\lambda)$. All four entries are operators acting on a Hilbert space (which we haven't defined yet).

Next we will choose a representation, which will correspond to choosing a physical Hilbert space. Eventually we want this Hilbert space to be the Hilbert space of our spin chain. A very logical first step is analogue to the adjoint represntation of a Lie-algebra, where the matrix elements of the generators are the structure constants:

$$(T_i^j)_l^k = R_{il}^{jk} \tag{12}$$

It is easy to verify this representation by plugging this into equation (3) and we immedeatly see that we retrieve equation (1). This choice fixes A, B, C and D as 2×2 matrices, which in turn fixes our Hilbert space to be $\mathbb{C}^2$. This becomes even more clear when we work out the explicit form of the generators using equation (12):

$$\begin{aligned} A(\lambda) = \begin{pmatrix} a(\lambda) & 0 \\ 0 & b(\lambda) \end{pmatrix} &= \tfrac{1}{2}(a+b)\mathbb{1} + \tfrac{1}{2}(a-b)\sigma_3, \\ B(\lambda) = \begin{pmatrix} 0 & 0 \\ c(\lambda) & 0 \end{pmatrix} &= c\sigma^-, \\ C(\lambda) = \begin{pmatrix} 0 & c(\lambda) \\ 0 & 0 \end{pmatrix} &= c\sigma^+, \\ D(\lambda) = \begin{pmatrix} b(\lambda) & 0 \\ 0 & a(\lambda) \end{pmatrix} &= \tfrac{1}{2}(a+b)\mathbb{1} - \tfrac{1}{2}(a-b)\sigma_3, \end{aligned} \tag{13}$$

with the diagramtic representation:

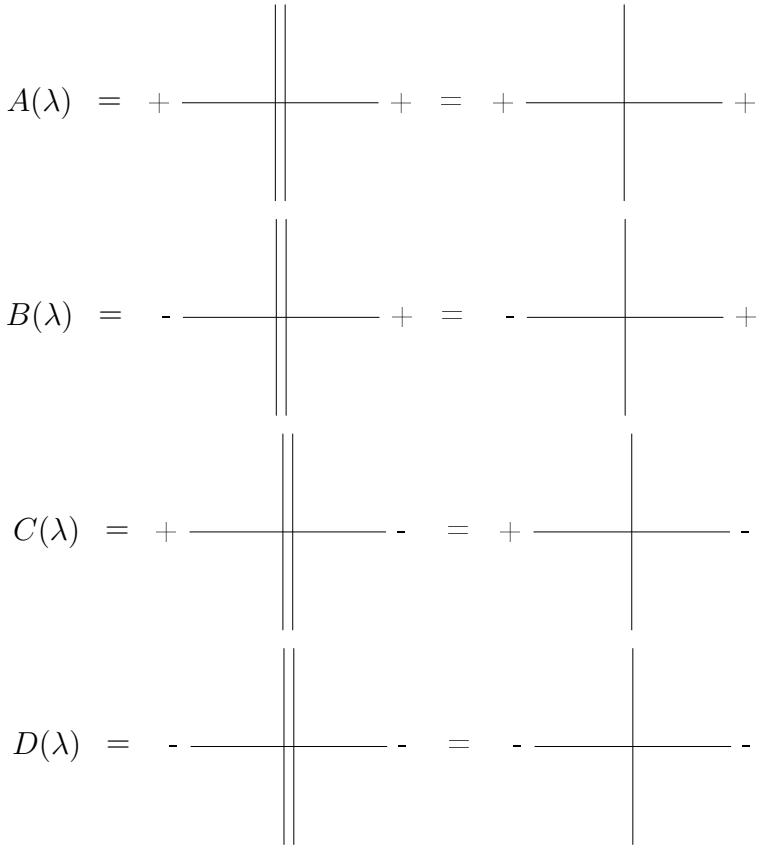

Figure 4: diagrams for the matrix elements of the monodromy matrix in the adjoint representation

We have just replaced our double line with a single one, since we have defined the R-matrix with a single line, for which we know that it represnts the Hilbert space $\mathbb{C}^2$.

3.3 The lowering operator

From this simple representation we get a clue what these generators A,B,C and D are physically: B acts as a lowering operator, C as a raising operatorand A and D span the Cartan subalgebra of $SU(2)$ in our Hilbert space. If we now act with the coproduct on these operators, we expect that these operators will keep their physical intergpretation, since Δ presserves the algebraic structure of the original algebra. If we want to extend these operators to the Hilbert space of the six-vertex model, $\bigotimes_{i=1}^{L} \mathbb{C}^2$, we just have to act Δ on them L-1 times.

But first let us consider the case $L = 2$ and see that the coproduct indeed preserves the physical function of the operators. For simplicity we will only work out $\Delta(B)$, but the computation for the other operators is very similar. Using equation (9), we can determine how Δ acts on $B(\lambda)$:

$$\Delta(B(\lambda)) = \sum_{k=\pm} T_-^k \otimes T_k^+ = B(\lambda) \otimes A(\lambda) + D(\lambda) \otimes A(\lambda). \tag{14}$$

Basically we can now work out how $\Delta(B)$ works on the 4 states $e_i \otimes e_j$, it is just a simple matrix multiplication. There is however also a nice diagramatic way to do this. The tensor product corresponds to connecting the diagrams right arm to left arm. Then we put the incoming state at the two lower arms of the vertices and impose conservation of $+$, using that the left and lower arm are incoming and that the right and upper arm are outgoing. We then figure out what the outgoing state will be by filling in the two upper arms, and the Boltzmann weights corresponding to the reulsting vertices will be the factor of the outgoing state. This method will make the computation a lot nicer and easier. The diagramatic result is on the following page.

In some diagrams the two upper arms of the vertex aren't filled in. This is because we can't fill them in if we want to obey + conservation. Therefore we must conclude that these diagrams are 0. We can use (2) to determine the factors is front of them.

$$\begin{aligned}
\Delta(B(\lambda))|++\rangle &= R_{-+}^{+-} R_{-+}^{-+} |+-\rangle + R_{-+}^{-+} R_{++}^{++} |-+\rangle \\
&= b(\lambda)c(\lambda)|+-\rangle + c(\lambda)a(\lambda)|-+\rangle \\
\Delta(B(\lambda))|+-\rangle &= R_{-+}^{+-} R_{+-}^{+-} |--\rangle = c(\lambda)b(\lambda)|--\rangle \\
\Delta(B(\lambda))|-+\rangle &= R_{--}^{--} R_{-+}^{-+} |--\rangle = a(\lambda)c(\lambda)|--\rangle \\
\Delta(B(\lambda))|--\rangle &= 0
\end{aligned} \tag{15}$$

From this computation we can indeed verify that $\Delta(B)$ acts like a lowering operator for the Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$, as we expected. In the same way one can show that $\Delta(C)$ is still a raising operator.

With the information that Δ preserves the function of B as a lowering operator we can extend this to a Hilbert space $\bigotimes_{i=1}^{L} \mathbb{C}^2$. Using the co-associativity relation (8) we know

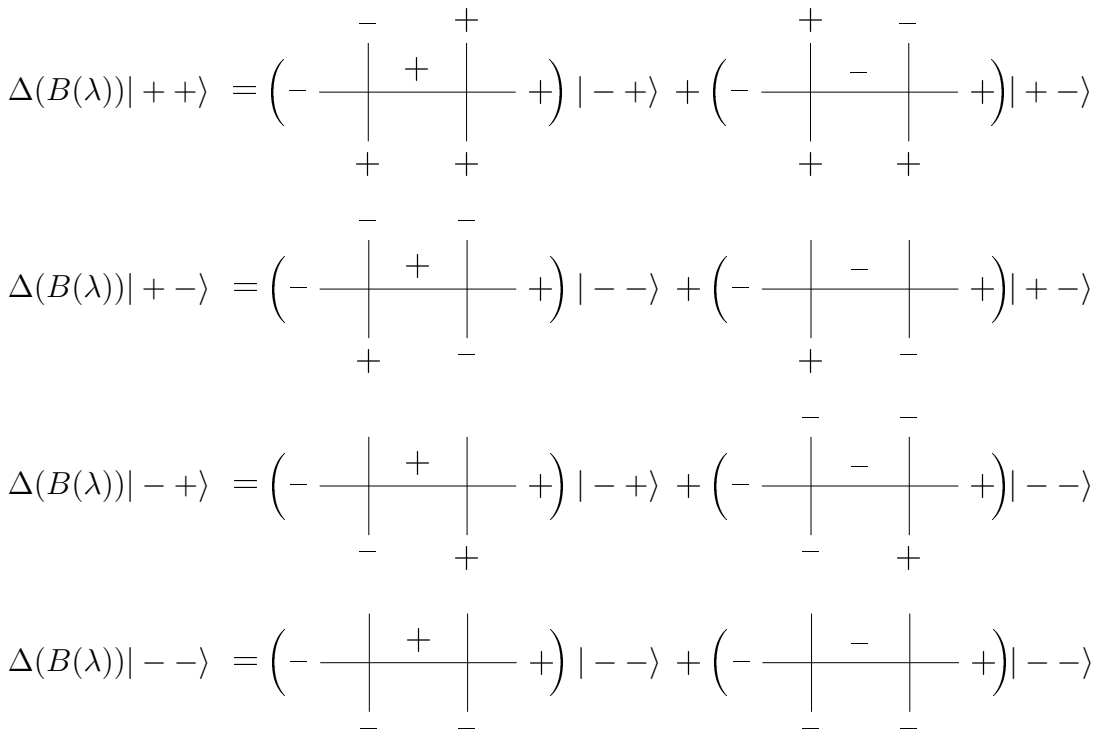

Figure 5: action of $\Delta(B)$ on the states of the spin chain of lenght 2.

that there is only one unique way to apply the Δ twice, so let us define $\Delta^2 := (\mathbb{1}\otimes\Delta)\Delta$. In the same way we can define Δ^{L-1} an Δ applied $L-1$ times, for which there is only one unique way, together with the generalized form of (8): $\Delta^{L-1} = (\mathbb{1}\otimes\Delta)\Delta^{L-2}$

4 Second Quantization

4.1 Analogy of Second Quantization

The main purpose of the previous discussion is that we have a nice definition of the raising and lowering operators for our spin chain which we are able to connect to the transfer matrix $t(\lambda) = A(\lambda) + D(\lambda)$ (namely via the Yang-Baxter equation for the monodromy matrix). We have also seen that the transfer matrix can be connected to the Hamiltonian. In other words, we have found the connection between the Hamiltonian and the lowering/raising operators of the spin chain. This sets us in a position to describe our spin chain analogous to the second quantization method.

The most important steps in setting up a second quantized system is to define the vacuum state $|0\rangle$. We define the vacuum state as the state with only spins up (we could of course also define it with only spins down but the convention is the state with only spin down). Since $\Delta^{L-1}(B(\lambda))$ (which we will call $B(\lambda)$ from now on) is the lowering

operator. Since it created a spin down (a magnon) this will now take the roll of the creation operator. With the same reasoning $C(\lambda)$ will take the roll of the annihilation operator, and lastly $A(\lambda)$ and $D(\lambda)$ will span the Cartan subalgebra.

With this new definition of the generators we are can define a state with M spins downs (magnons) as:

$$|\Psi_M\rangle = \prod_{i=1}^{M} B(\lambda_i)|0\rangle. \tag{16}$$

The question we now want to ask is: under what conditions is $|\Psi_M\rangle$ an eigenstate of t? We need to know what values of λ_i are allowed for this state to be an eigenstate. To find this out we must act on $|\Psi_M\rangle$ with $A(\lambda_i) + D(\lambda_i)$, but to know what this operator does on $|\Psi_M\rangle$ we need to know the commutator of A,B,C and D. To get these we use the Yang-Baxter equation (3). We need to rewrite this equation in a form with which we can do calculations, since the subscripts are not indices. To do this we must first realize that equation (3) is an equation on $V_a \otimes V_b$, so when we write down T_a we actually mean that we act with T on the space V_a and do nothing on V_b. Therefore we must replace T_a with $T_a \otimes \mathbb{1}$. Similarly we replace T_b with $\mathbb{1} \otimes T_b$. Then equation (3) becomes:

$$R_{ab}(\lambda - \mu)(T_a(\lambda) \otimes T_b(\mu)) = (T_b(\mu) \otimes T_a(\lambda))R_{ab}(\lambda - \mu). \tag{17}$$

We have:

$$T_a(\lambda) \otimes T_b(\mu) = \begin{pmatrix} A(\lambda)T(\mu) & B(\lambda)T(\mu) \\ C(\lambda)T(\mu) & D(\lambda)T(\mu) \end{pmatrix}. \tag{18}$$

To get $T_b(\mu) \otimes T_a(\lambda)$ we just swap λ and μ. Now using the R-matrix (10) we can calculate both sides of the Yang-Baxter equation. This is easy to do, but very tedious. In the end we get 16 equation, and 10 of these will give the commutation relations. Instead of doing this tedious calculation, we can just give the three most important 3, namely B with itself and with A and D:

$$\begin{aligned} B(\lambda)B(\mu) &= B(\mu)B(\lambda), \\ A(\lambda)B(\mu) &= \frac{a(\mu-\lambda)}{b(\mu-\lambda)}B(\mu)A(\lambda) - \frac{c(\mu-\lambda)}{b(\mu-\lambda)}B(\lambda)A(\mu), \\ D(\lambda)B(\mu) &= \frac{a(\lambda-\mu)}{b(\lambda-\mu)}B(\mu)D(\lambda) - \frac{c(\lambda-\mu)}{b((\lambda-\mu)}B(\lambda)D(\mu). \end{aligned} \tag{19}$$

These commutation relations are the ones we need to calculate the action of A and D on $|\Psi_M\rangle$. The first equation comes from the fourth column, first row, the second from third column first row and the third from the fourth column second row. Lastly we need the relation $A(\lambda)|0\rangle = a^L(\lambda)|0\rangle := \alpha(\lambda)|0\rangle$ and $D(\lambda)|0\rangle = b^L(\lambda)|0\rangle := \delta(\lambda)|0\rangle$. Now we have all the tools we need to diagonalize the transfer matrix. The fact that all B's commute with each other is an important result, since it states that it doesn't

matter in which order we create a spin down, completing the analogy with second quantization.

4.2 Diagonalizing the Transfer Matrix

As mentioned before, diagonalizing the transfer matrix comes down to asking the question: for what values of λ_i is $|\Psi_M\rangle$ an eigenvector of $t(\lambda)$, and what are its eigenvalues? We want to solve the following equation for λ_i:

$$(A(\lambda) + D(\lambda))|\Psi_M\rangle = \Lambda_M(\lambda, \{\lambda_i\})|\Psi_M\rangle. \tag{20}$$

Using our expression for $|\Psi_M\rangle$ (16) we want to use the commutation relations (19) between A and B and between D and B to commute A and D to the vacuum state, since we know how these act on this. Since the commutation relations for A and D are very similar, we will focus on one first. The other is easily obtainable from that.

Let us focus on how $A(\lambda)$ acts on $|\Psi_M\rangle$. First, for simplicity, define

- $f(\lambda) = \frac{a(\lambda)}{b(\lambda)}$,
- $g(\lambda) = -\frac{c(\lambda)}{b(\lambda)}$.

From the commutation relation we can see that, when commuting, the argument of A either stays with A, as in the first term, or switches to B, as in the second term. Then we can say that after commuting $A(\lambda)$ to the right, we can have that λ ends up with A, by only taking the first term when commuting, or we can have that λ ends up with a B. This allows us to write the action of A on $|\Psi_M\rangle$ as follows:

$$\begin{aligned} A(\lambda)|\Psi_M\rangle &= \prod_{i=1}^{M} f(\lambda_i - \lambda)B(\lambda_i)\alpha(\lambda)|0\rangle \\ &+ \sum_{k=1}^{M} M_k(\lambda, \{\lambda_i\})B(\lambda)B(\lambda_1)..\hat{B}(\lambda_k)..B(\lambda_M)|0\rangle. \end{aligned} \tag{21}$$

Here $\hat{B}(\lambda_i)$ means that that term is absent from the row. The coefficients M_k are in general very difficult, complicated expressions. They are a sum of many different contributions where the argument of A eventually ends up with $B(\lambda_k)$. The higher the value of k the more terms that will contribute to M_k. However, the coefficient M_1 is easy to calculate, since there is only one way the argument of $B(\lambda_1)$ can end up with A. First commute $A(\lambda$ and $B(\lambda_1)$:

$$A(\lambda)\prod_{i=1}^{M} B(\lambda_i) = f(\lambda_1 - \lambda)B(\lambda_1)A(\lambda)\prod_{i=2}^{M} B(\lambda_i) + g(\lambda_1 - \lambda)B(\lambda)A(\lambda_1)\prod_{i=2}^{M} B(\lambda_i). \tag{22}$$

Now that we have λ_1 as the argument of A, the only way to get the coefficient M_1 is to keep λ_1 with A, so that $B(\lambda_1)$ is absent from the resulting product of B's, which

is how we defined M_k. To keep λ_1 with A we can only take the first term of the commutatio relation when commuting A to the right. After doing this we get the following expression for M_1:

$$M_1 = g(\lambda_1 - \lambda)\alpha(\lambda_1)\prod_{i=2}^{M} f(\lambda_i - \lambda_1). \tag{23}$$

Instead of doing all the computations for the other M_k, there is a subtle argument we can use to get all M_k from M_1. Namely in determining M_1 we have first commuted A with the B on the far left. Determining M_1 was then easy because there was only one contribution. However, since all B's commute, we could just as well have put any other B on the far left, and then calculated another M_k just as easily. Therefore we have no other possaibility than to assume that all M_k's have the same form. If we want to get M_k, we just have to replace $1 \leftrightarrow k$ in the expression for M_1:

$$M_k = g(\lambda_k - \lambda)\alpha(\lambda_k)\prod_{i=1, i\neq k}^{M} f(\lambda_i - \lambda_k). \tag{24}$$

It might seem strange at first that all M_k take this relatively simple form, because if we would just calculate it in the naive way, by doing the full commutation, each M_k will be a sum of different contribtions (the number of contribtuions get higher as k grows). However, if we use this subtle but brilliant argument we can conclude that these sums, which consist of different products/quotients of the functions $f(\lambda)$ and $g(\lambda)$, must somehow combine to give the expression for M_k stated above. In principle we don't care about these sum rules. The only thing that is important is that they exist and, by the arguments explained above, we know they do.

Now we can do exactly the same calculations and argumentations for $D(\lambda)|\Psi_M\rangle$, and the result is the same as we got for A but with an extra minus sign in the argument of f and g, replacing $\alpha(\lambda)$ with $\delta(\lambda)$ and renaming the coefficients M_k to N_k:

$$\begin{aligned} D(\lambda)|\Psi_M\rangle &= \prod_{i=1}^{M} f(\lambda - \lambda_i)B(\lambda_i)\delta(\lambda)|0\rangle \\ &+ \sum_{k=1}^{M} N_k(\lambda, \{\lambda_i\})B(\lambda)B(\lambda_1)..\hat{B}(\lambda_k)..B(\lambda_M)|0\rangle, \end{aligned} \tag{25}$$

with:

$$N_k = g(\lambda - \lambda_k)\delta(\lambda_k)\prod_{i=1, i\neq k}^{M} f(\lambda_k - \lambda_i). \tag{26}$$

If we add (21) and (25), we see that the first terms of both equations are what we want, because they give us an expression with some factor times $|\Psi_M\rangle$. From these we can read of the eigenvalue of the transfer matrix:

$$\Lambda_M(\lambda, \{\lambda_i\}) = a^L(\lambda)\prod_{i=1}^{M} f(\lambda_i - \lambda) + b^L(\lambda)\prod_{i=1}^{M} f(\lambda - \lambda_i). \tag{27}$$

But this is of course only the eigenvalue if we demand that the coefficients M_k and N_k exactly cancel eachother:

$$M_k + N_k = 0. \tag{28}$$

This gives us the following equation:

$$g(\lambda - \lambda_k)\delta(\lambda_k) \prod_{i=1, i\neq k}^{M} f(\lambda_k - \lambda_i) = -g(\lambda_k - \lambda)\alpha(\lambda_k) \prod_{i=1, i\neq k}^{M} f(\lambda_i - \lambda_k). \tag{29}$$

Now we have all the equations we need to under which condition $|\Psi_M\rangle$ is an eigenvector of the transfer matrix.

5 The rediscovery of earlier results

Now that have diagonalized the transfer matrix, we can retrieve the familiar formulas for the six vertex model. First of all, remember that for the six-vertex model we had the following paramtrization:

- $a(\lambda) = \sinh(\lambda + \phi)$,
- $b(\lambda) = \sinh(\lambda)$,
- $c(\lambda) = \sinh(\phi)$,

with $\Delta = \cosh(\phi)$. Using that we defined as $f(\lambda)$ and $g(\lambda)$ as (4.2), we see that $g(\lambda)$ is an anti-symmetric function, so that it cancels while eliminating the minus sign in equation (29). After some rearranging we get:

$$\left(\frac{a(\lambda)}{b(\lambda)}\right)^L = \prod_{i=1, i\neq k}^{M} \frac{a(\lambda_k - \lambda_i)b(\lambda_i - \lambda_k)}{a(\lambda_i - \lambda_k)b(\lambda_k - \lambda_i)}. \tag{30}$$

And if we use the above parametrization we get:

$$\left(\frac{\sinh(\lambda + \phi)}{\sinh(\lambda)}\right)^L = \prod_{i-1, i\neq k}^{M} -\frac{\sinh(\lambda_k - \lambda_i + \phi)}{\sinh(\lambda_i - \lambda_k + \phi)}. \tag{31}$$

We can put it in a form which is often used in the literature by first changing $\phi \to i\phi$ and then shifting $\lambda_i \to \frac{1}{2}\phi(\lambda_i - i)$ to get it in the following familiar form:

$$\left(\frac{\sinh\frac{1}{2}\phi(\lambda + i)}{\sinh\frac{1}{2}\phi(\lambda - i)}\right)^L = \prod_{i-1, i\neq k}^{M} \frac{\sinh\frac{1}{2}\phi(\lambda_k - \lambda_i + 2i)}{\sinh\frac{1}{2}\phi(\lambda_k - \lambda_i + 2i)}. \tag{32}$$

This is of course the Bethe equation for the XXZ model. This parametrization is a nice one since it makes the limit $\phi \to 0$ easier, which is the same as the limit $\Delta \to 0$. In this way we obtain the Bethe equations for the XXX model:

$$\left(\frac{\lambda_i + i}{\lambda_i - i}\right)^L = \prod_{i-1, i\neq k}^{M} \frac{\lambda_k - \lambda_i + 2i}{\lambda_k - \lambda_i - 2i}. \tag{33}$$

Recall that from the lecture about the six-vertex model we got the following result for the eigenvalue Λ_M:

$$\Lambda_M = a^L \prod_{i=1}^{M} L(z_i) + b^L \prod_{i=1}^{M} M(z_i), \tag{34}$$

with

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

We know that z can be identitied with the quasi-momentum as $z = e^{k_i}$. If we compare this equation with 27 we see that we can identify $f(\lambda_i - \lambda)$ with $L(z)$ and $f(\lambda - \lambda_i)$ with $M(z_i)$. If we now evaluate the first equality at $\lambda = 0$ and with the parametrization of the six-vertex model we get:

$$f(\lambda_i) = \frac{a(\lambda_i)}{b(\lambda_i)} = \frac{\sinh(\lambda_i + \phi)}{\sinh \lambda_i} = \frac{a(0)b(0) + (c(0)^2 - b(0)^2)e^{ik_i}}{a(0)^2 - a(0)b(0)e^{ik_i}} = e^{ik_i}. \tag{35}$$

This gives us the parametrization between the spectral parameter λ_i and the corresponding quasi-momenta k_i. With this we can write the Bethe equation in terms of the quasi-momenta:

$$e^{ik_i L} = \prod_{i=1, i\neq k}^{M} \frac{2\Delta e^{ik_k} - e^{i(k_i + k_k)} - 1}{2\Delta e^{ik_i} - e^{i(k_i + k_k)} - 1}. \tag{36}$$

Now we are also in the position to calculate the energy and the momentum of the magnons. As shown in the previous chapter we can write the Hamiltonian and momentum operator in terms of the transfer matrix as follows:

$$H = i\frac{\partial}{\partial \lambda}\left(\log\left(\frac{t(\lambda)}{a^L(\lambda)}\right)\right)\Big|_{\lambda=0}, \quad P = i \log\left(\frac{t(0)}{a^L(0)}\right). \tag{37}$$

Since the Hamiltonian, momentum and transfer matrix operators all commute with each other, they share the same eigenstaes $|\Psi_M\rangle$. Therefore we can say:

$$E_M = i\frac{\partial}{\partial \lambda}\left(\log\left(\frac{\Lambda_M(\lambda, \{\lambda_i\})}{a^L(\lambda)}\right)\right)\Big|_{\lambda=0}, \quad p_M = i \log\left(\frac{\Lambda_M(0, \{\lambda_i\})}{a^L(0)}\right). \tag{38}$$

Let us first calculate the momentum of the state $|\Psi_M\rangle$ using equation 27 and the fact that $b(0) = 0$ and $a(0) = \sinh(\phi)$:

$$p_M = i \log \left(\prod_{i=1}^{M} \frac{\sinh(\lambda_i + \phi)}{\sinh \lambda_i} \right) = i \sum_{i=1}^{M} \log \left(\frac{\sinh(\lambda_i + \phi)}{\sinh \lambda_i} \right). \tag{39}$$

Before we calculate the energy, we first note that the second term in 27 doesn't contribute to the energy, since $b(0) = \frac{\partial b}{\partial \lambda}\big|_{\lambda=0} = 0$ for $L > 2$. Of course L takes much higher values than 2 since we are ultimatly looking for the thermodynamic limit where L becomes infinite. Therefore we only need to take the derivative of the logarithm of the first term in (27):

$$\begin{aligned} E_M &= i \, \frac{\partial}{\partial \lambda} \left(\log \left(\prod_{i=1}^{M} \frac{a(\lambda_i - \lambda)}{b(\lambda_i - \lambda)} \right) \right) \Bigg|_{\lambda=0}, \\ &= i \sum_{i=1}^{M} \left(-\frac{1}{a(\lambda_i - \lambda)} \frac{\partial a}{\partial \lambda}(\lambda_i - \lambda) + \frac{1}{b(\lambda_i - \lambda)} \frac{\partial b}{\partial \lambda}(\lambda_i - \lambda) \right) \Bigg|_{\lambda=0}, \\ &= -i \sum_{i=1}^{M} \left(\frac{\cosh(\lambda_i + \phi)}{\sinh(\lambda_i + \phi)} - \frac{\cosh \lambda_i}{\sinh \lambda_i} \right), \\ &= i \sum_{i=1}^{M} \frac{\sinh \phi}{\sinh(\lambda_i + \phi) \sinh \lambda_i}, \end{aligned} \tag{40}$$

where we have used the sum rule $\sinh\theta \cosh\phi - \cosh\theta \sinh\phi = \sinh(\theta - \phi)$. From here we see that the energy and momentum of the magnons are addative: the total energy/momentum is equal to the sum of the energy/momentum of the seperate magnons. If we define $E_M = \sum_{i=1}^{M} \epsilon_i$ and $p_M = \sum_{i=1}^{M} \pi_i$, we get for the cosine of π_i:

$$\begin{aligned} \cos \pi_i &= \cos \left(i \log \left(\frac{\sinh(\lambda_i + \phi)}{\sinh \lambda_i} \right) \right), \\ &= \frac{1}{2} \left(\frac{\sinh(\lambda_i + \phi)}{\sinh \lambda_i} + \frac{\sinh \lambda_i}{\sinh(\lambda_i + \phi)} \right), \\ &= \frac{\sinh^2(\lambda_i + \phi) + \sinh^2 \lambda_i}{2 \sinh(\lambda_i + \phi) \sinh \lambda_i}. \end{aligned} \tag{41}$$

With this we can calculate the following quantity:

$$\Delta - \cos \pi_i = \frac{2 \cosh \phi \sinh \lambda_i \sinh(\lambda_i + \phi) - \sinh^2(\lambda_i + \phi) - \sinh^2 \lambda_i}{2 \sinh(\lambda_i + \phi) \sinh \lambda_i}. \tag{42}$$

The easiest way to simplify this expression is to expand the cosh and sinh in exponents and then work out the brackets. This will result in many terms, but surprisingly many of them cancel and what we are left with is:

$$\Delta - cos\pi_i = \frac{\cosh^2 \phi}{2 \sinh(\lambda_i + \phi) \sinh \lambda_i}. \tag{43}$$

From this expression we recognize ϵ_i, giving us the dispersion relation for the magnons:

$$E_M = \sum_{i=1}^{M} \epsilon_i = \sum_{i=1}^{M} -i \frac{\cosh^2 \phi}{2 \sinh \phi} (\Delta - \cos \pi_i) = \sum_{i=1}^{M} J(\Delta - \cos \pi_i). \tag{44}$$

This momentum dependace of the energy is a familiar result of the XXZ model. Remember that ϕ can be imaginairy so J can (and should) be real.

Now we can reparametrize these again bu replacing $\phi \to i\phi$ and then $\lambda \to \frac{1}{2}\phi(\lambda - i)$ to get:

$$E_M = -\sum_{i=1}^{M} \frac{\sin \phi}{\sinh \frac{1}{2}\phi(\lambda_i + i) \sinh \frac{1}{2}\phi(\lambda_i - i)}, \quad p_M = i \sum_{i=1}^{M} \frac{\sinh \frac{1}{2}\phi(\lambda_i + i)}{\sinh \frac{1}{2}\phi(\lambda_i - i)}. \tag{45}$$

Lastly when we take the limit $\phi \to 0$ and get the energy and momentum of the XXX model:

$$E_M = -\sum_{i=1}^{M} \frac{4}{(\lambda_i + 1)^2}, \quad p_M = i \log \left(\frac{\lambda_i + i}{\lambda_i - i} \right) \tag{46}$$

6 Conclusion

In the last few sections I have shown that we can solve the six-vertex model from a whole new perspective. Just by assuming an algebra, the Yang-Baxter algebra, we found a new method of describing the spin chain, which is analogous to second quantization. Remember that the coordinate Bethe Ansatz was an assumptions on the wave-function of the magnons. Then we have to do calculations with superposition states. In the algebraic Bethe Ansatz we have made assumptions that give rise to the creation operators of the magnons. The calculations boil down to doing calculations with operators, namely commutation relations. This point makes a very nice analogy between Coordinate/Algebraic Ansatz and first/second quantization. The main purpose of this new method is similar, namely that we have a new way of describing spin chains. Eventually this method allows us not just to describe the six-vertex model, but also many other models. In the beginning all the definitons were very general, so the method described in this chapter can be applied to many other systems. The only thing that is important to know is the form of the R-matrix. Once you know this, the way to do the calculations are similar to what we've done in this chapter (of course, it can be much more difficult or tedious). A large part of this chapter was aimed at verifying this new approach by applying it to the model we know so well, the six-vertex model. After the analyses we indeed obtained the same equations as the longer, more cumbersome calculations from the six-vertex model. Now that we have verified this new approach, we have laid the foundations to describing other, more difficult models.