6 Yang-Baxter equation

In the preceeding chapter we have seen how to solve the six-vertex model by the coordinate Bethe Ansatz. A natural next step is to identify the structures that make such a solution possible, and then to exploit those structures to solve more general classes of models. One would expect the structures in question to be of algebraic nature.

Obviously this programme will require to first take a slightly more abstract point of view. We shall expose the necessary ingredients in the first half of this chapter, and then connect them concretely to the six-vertex model in the second half.

6.1 R-matrix

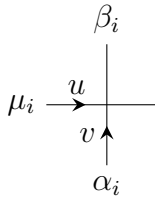
Let us start by highlighting some crucial results from our analysis of the six-vertex model:

1. The three Boltzmann weights of the six-vertex model define two independent ratios $a : b : c$. From these we have defined two parameters:
   - A parameter $\Delta = \frac{a^2+b^2-c^2}{2ab}$ which characterises the universality class of the model. Only this combination appears in the Bethe Ansatz equations, via the scattering amplitudes.
   - Another uniformising parameter—called w in (5.52)—on which the ratios $a : b : c$ depend, but which does not change the value of Δ.
2. The scattering amplitudes—encoded in the S-matrix—are such that that multi-particle scattering of several quasi-particles factorise into a pruduct of two-particle scattering amplitudes.

We now attempt to formalise these properties for statistical models of a particular type, defined on a so-called *Baxter lattice*. By this we mean any lattice that can be drawn in the plane as a collection of lines (one can think of them as straight lines, but this is not necessary) that undergo only pairwise intersections. The lattice does not at all need to be regular. The degrees of freedom live on the edges of the lattice, and interactions take place at the vertices.

Moreover, we suppose that to each line is associated a so-called *spectral parameter*, analogous to the uniformising parameter discussed above. The Boltzmann weights for a vertex where two lines with spectral parameters u and v intersect is supposed to have the *difference property*: it must depend only on the difference $u - v$. The weights of course also depends on Δ, and on the states of the degrees of freedom defined on the edges adjacent on the vertex.

The Boltzmann weights can be encoded as the matrix element of a linear operator, called the R-matrix:



$$= R^{\mu_{i+1}\beta_i}_{\mu_i\alpha_i}(u-v) = {}_{\mathrm{a}}\langle\mu_{i+1}| \otimes {}_i\langle\beta_i|R_{\mathrm{a}i}(u-v)|\mu_i\rangle_{\mathrm{a}} \otimes |\alpha_i\rangle_i \tag{6.1}$$

At the risk of appearing pedantic, let us explain very carefully this notation:

- The indices μ_i, α_i and μ_{i+1}, β_i label the statistical degrees of freedom defined on the lattice edges. The notation corresponds exactly to what we have seen in the definition of the transfer matrix for the dimer problem. The first pair of indices is the in-state, and the second pair is the out-state. The time evolution thus goes in the North-East direction.

- A spectral parameter u is attached to the horizontal line, and v to the vertical line. The R-matrix depends on the difference $u - v$. To avoid any confusion about the sign, the lines carry an orientation.[16] When looking along the direction of the time evolution, the argument of R is the spectral parameter seen on one's left (namely u) *minus* the spectral parameter seen on one's right (namely v), both counted with a sign that reflects the orientation of the lines along the direction of sight.

- We shall often refer to the spaces α and β as *quantum* and to the space μ as *auxiliary*.

- The R-matrix acting between the auxiliary space and the ith quantum is denoted by $R_{\mathrm{a}i}$. Its components are denoted $R^{\mu_{i+1}\beta_i}_{\mu_i\alpha_i}$. Note that the

[16] These arrows should of course not be confused with the (six-vertex) arrow degrees of freedom that live on lattice edges!

order of the out-indices has been permuted: this convention defines the R-matrix. One sometimes encounter the opposite, unpermuted convention: this defines what is called the $\check{R}$-matrix.

More formally, the R-matrix is a linear operator

$$R_{\mathrm{a}i} : V_{\mathrm{a}} \otimes V_i \mapsto V_{\mathrm{a}} \otimes V_i \,, \tag{6.2}$$

where the vector spaces V_{a} (auxiliary) and V_i (quantum) carry the edge degrees of freedom. For instance, in the six-vertex model they are both equal to the spin-$\frac{1}{2}$ representation space $\mathbb{C}^2$, since each arrow can be in two possible states: the R-matrix is then a 4×4 matrix.

The transfer matrix t is an endomorphism on the tensor product of all quantum spaces

$$t : V_1 \otimes V_2 \otimes \cdots \otimes V_L \mapsto V_1 \otimes V_2 \otimes \cdots \otimes V_L \,. \tag{6.3}$$

It can be written as

$$t = \mathrm{Tr}_{\mathrm{a}} \left(R_{\mathrm{a}L} R_{\mathrm{a}L-1} \cdots R_{\mathrm{a}2} R_{\mathrm{a}1} \right) , \tag{6.4}$$

where Tr_{a} denotes the trace over the auxiliary space V_{a}. For simplicity we have not written the dependence on the spectral parameters. Indeed, one has the possibility of taking *different* spectral parameters for each quantum space V_i, and also for V_{a}, which will correspond to a completely inhomogeneous lattice model. The matrix elements of t can be written very explicitly as

$$\langle \beta | t | \alpha \rangle = \sum_{\mu_1, \ldots, \mu_L} R^{\mu_1 \beta_L}_{\mu_L \alpha_L} R^{\mu_L \beta_{L-1}}_{\mu_{L-1} \alpha_{L-1}} \cdots R^{\mu_3 \beta_2}_{\mu_2 \alpha_2} R^{\mu_2 \beta_1}_{\mu_1 \alpha_1} \,. \tag{6.5}$$

Note that μ_1 appears both in the rightmost and the leftmost factor, so we indeed perform the operator Tr_{a}.

Before going on, it is legitimate to ask oneself whether this formalism is general enough to accommodate "all" statistical models of interest:

- A first question concerns the generality of the Baxter lattice. This of course encompasses all regular lattices for which all vertices are of degree four. What then about other lattices, such as the hexagonal and triangular lattices, which are widely used in statistical physics? One can usually find one's way out by making suitable transformations.

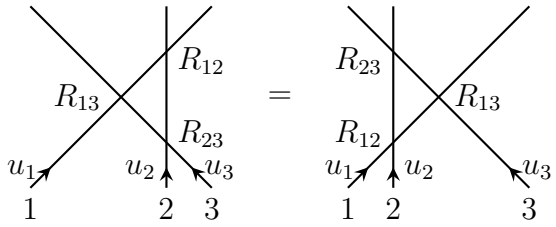
For instance, on the hexagonal lattice one can shrink all edges along one of the three principal direction to zero, thus regrouping pairs of 3-valent vertices to form 4-valent vertices:[17] This turns the hexagonal lattice into a square lattice. The triangular lattice can be turned into a hexagonal lattice by duality, or into a Kagomé lattice (which is a Baxter lattice) by going to the medial lattice. Other possible tricks include decimation procedures.

- A second question is how to deal with situations where the statistical degrees of freedom are not defined on the edges, but rather on the vertices. Going to suitably defined dual variables will usually enable us to transform such a model into one involving edge variables.

- Finally, how could one deal with long-range interactions, mediated by spatially extended objects, such as clusters or loops? One possibility is to define the R-matrix on appropriate representation spaces that take into account the non-locality of the interaction. Another option is to transform the model into one with local interactions. We shall see both possibilities at play in our subsequent treatment of the Potts model, which can be represented either as a Temperley-Lieb loop model with non-local interactions, or transformed into a six-vertex model with complex Boltzmann weights.

6.2 Commuting transfer matrices

A statistical model defined on a Baxter lattice is said to be integrable provided its R-matrix satisfies the Yang-Baxter equation and the inversion relation. The Yang-Baxter relation reads pictorially:



(6.6)

[17] This is a lattice version of the Hubbard-Stratonovich transformation used in field theory.

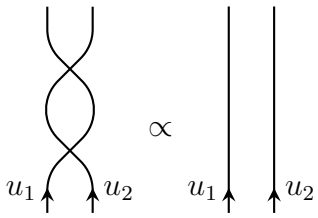
The algebraic transscription is

$$R_{12}(u)R_{13}(u+v)R_{23}(v) = R_{23}(v)R_{13}(u+v)R_{12}(u)\,, \tag{6.7}$$

where we have set $u = u_1 - u_2$ and $v = u_2 - u_3$, so that $u_1 - u_3 = u + v$. We stress again that the spectral parameters u_i always follow the lines of the Baxter lattice. The same is true for the labels of the representation spaces, that appear as subscripts for the R-matrix. It is sometimes convenient to have these labels stay well-ordered in space (i.e., with 1 on the left, 2 in the middle, and 3 on the right) at all times (in the diagram time flows upwards). In that case one uses instead the $\check{R}$-matrix, for which the Yang-Baxter equation reads

$$\check{R}_{23}(u)\check{R}_{12}(u+v)\check{R}_{23}(v) = \check{R}_{12}(v)\check{R}_{23}(u+v)\check{R}_{12}(u)\,. \tag{6.8}$$

The inversion relation can be represented pictorially as



(6.9)

and reads algebraically

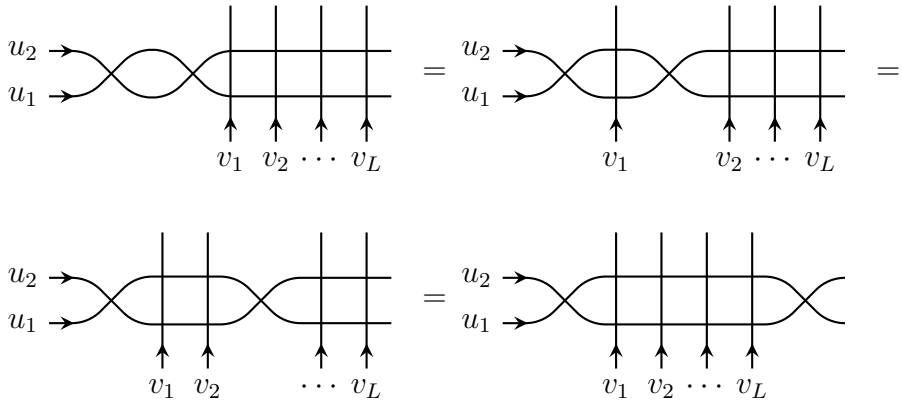
$$R_{12}(u)R_{12}(-u) \propto I\,. \tag{6.10}$$

The constant of proportionality could of course be set to unity by a suitable rescaling of R. Note also how the sign convention for spectral parameters comes into use when writing (6.10).

We shall show below that the R-matrix of the six-vertex model indeed satisfies (6.7) and (6.10).

The relations (6.7) and (6.10) imply the commutation of two transfer matrices corresponding to different choices of spectral parameters on the

auxiliary lines. This is best demonstrated graphically:



(6.11)

The first picture represents the product $t(u_2)t(u_1)$, since the two crossings to the left amount to the identity by (6.10).[18] In the second picture we have used (6.7) to push the v_1 line to the left. This is repeated in the third picture for the next v_2 line. Repeating this operation L times, we finally arrive at the last picture, which represents the product $t(u_1)t(u_2)$, apart from the crossings on the left and right. But the right crossing can be taken around the periodic boundary condition (more formally: we are using the cyclicity of the trace), and using once more (6.10) the two crossings annihilate. Summarising, we have shown that

$$t(u_2)t(u_1) = t(u_1)t(u_2)\,. \tag{6.12}$$

The existence of an infinite family of commuting transfer matrices has important consequences. Indeed the Bethe Ansatz technique permits us to diagonalise all these transfer matrices simultaneously.

Moreover, we can take derivatives of (6.12) with respect to u_2. All these derivatives commute with $t(u_1)$, hence are conserved by the time evolution process. In other words, an integrable system has an infinite number of conserved quantities. The first few derivatives can be identified with the Hamiltonian, the momentum operator, and so on. We shall present explicit examples below.

Note also that the various vector spaces in which the R-matrices act need not be isomorphic. In particular, one can have different representations on the quantum and auxiliary spaces. From a basic integrable model—such as

[18] The transfer matrices depend also on the spectral parameters $v_1, v_2, \ldots, v_L$ of the quantum spaces, but we omit this dependence for notational convenience.

the six-vertex model in the spin-$\frac{1}{2}$ representation—one can construct higher-spin solutions by appropriate fusions of representation spaces. One speaks in that case of descendent models.

6.3 Six-vertex model

Let us parametrise the weights of the six-vertex model (cf. Fig. 15) as follows:

$$\begin{aligned} \omega_1 = \omega_2 &= \sin(\gamma - u)\,, \\ \omega_3 = \omega_4 &= \sin u\,, \\ \omega_5 &= \mathrm{e}^{-i(u-\eta)} \sin\gamma\,, \\ \omega_6 &= \mathrm{e}^{i(u-\eta)} \sin\gamma\,. \end{aligned} \tag{6.13}$$

We have then $\Delta = -\cos\gamma$. The gauge parameter η can be chosen at will, since vertices of type 5 and 6 appear in pairs, and only the value of $\sqrt{\omega_5\omega_6}$ enters the computation of the partition function.

Denote by $\mathbb{C}^2 = 0, 1$ the occupation number of each edge in the particle picture defined by the lower half of Fig. 15. The $\check{R}$-matrix can then be written in the basis $(\mathbb{C}^2)^2 = \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ as

$$\check{R} = \begin{bmatrix} \omega_1 & 0 & 0 & 0 \\ 0 & \omega_5 & \omega_4 & 0 \\ 0 & \omega_3 & \omega_6 & 0 \\ 0 & 0 & 0 & \omega_2 \end{bmatrix}. \tag{6.14}$$

In the gauge $\eta = 0$ this becomes

$$\check{R}(u) = \begin{bmatrix} \sin(\gamma-u) & 0 & 0 & 0 \\ 0 & \mathrm{e}^{-iu}\sin\gamma & \sin u & 0 \\ 0 & \sin u & \mathrm{e}^{iu}\sin\gamma & 0 \\ 0 & 0 & 0 & \sin(\gamma-u) \end{bmatrix}. \tag{6.15}$$

We now claim that this $\check{R}$-matrix satisfies the Yang-Baxter equation (6.8), where the uniformising parameter u has been identified with the spectral parameter.

It is an instructive exercise to verify this. In tensor notation (6.8) reads in the space $(\mathbb{C}^2)^3$

$$(I\otimes\check{R}(u))(\check{R}(u+v)\otimes I)(I\otimes\check{R}(v)) = (\check{R}(v)\otimes I)(I\otimes\check{R}(u+v))(\check{R}(u)\otimes I)\,. \tag{6.16}$$

This identity between 8×8 matrices is greatly simplified by the symmetries of the problem. First, the number of particles is conserved. Second, the weights are invariant under a global negation ($0 \leftrightarrow 1$) of the occupation numbers.

The only equations to be verified thus concern a 1×1 matrix (in the 0-particle space $|000\rangle$) and a 3×3 matrix (in the 1-particle space $|100\rangle$, $|010\rangle$, $|001\rangle$). Only the latter gives rise to non-trivial equations.

The inversion relation (6.10) reads

$$\check{R}(u)\check{R}(-u) = \sin(\gamma - u)\sin(\gamma + u) I \,. \tag{6.17}$$

6.3.1 Temperley-Lieb algebra

The $\check{R}$-matrix (6.15) can be decomposed as

$$\check{R}(u) = \sin(\gamma - u) I + \sin(u) E \,, \tag{6.18}$$

where I is the identity operator in $V^2 = (\mathbb{C}^2)^2$ and

$$E = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \mathrm{e}^{-i\gamma} & 1 & 0 \\ 0 & 1 & \mathrm{e}^{i\gamma} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \tag{6.19}$$

The operator E satisfies the basis-independent relations

$$\begin{aligned} E^2 &= 2\cos\gamma\, E \,, \\ (E\otimes I)(I\otimes E)(E\otimes I) &= E\otimes I \,, \\ (I\otimes E)(E\otimes I)(I\otimes E) &= I\otimes E \,, \end{aligned} \tag{6.20}$$

where now I is the identity in V. For a system of width L one defines on $V_L = V^{\otimes L}$ a family of $L-1$ such operators:

$$E_m = I^{\otimes m-1} \otimes E \otimes I^{\otimes L-m-1} \,, \qquad \text{for } m = 1, 2, \ldots, L-1. \tag{6.21}$$

They verify the relations

$$\begin{aligned} (E_m)^2 &= 2\cos\gamma\, E_m \,, \\ E_m E_{m\pm 1} E_m &= E_m \,, \\ E_m E_{m'} &= E_{m'} E_m \text{ for } |m - m'| > 1 \end{aligned} \tag{6.22}$$

defining the so-called Temperley-Lieb (TL) algebra.

The TL algebra plays a major role in lattice models of statistical mechanics, the Potts model in particular. In addition to the above spin-$\frac{1}{2}$ arrow representation, the TL algebra can be represented in terms of Fortuin-Kasteleyn clusters, their surrounding loops, domain walls of Potts spins, two-row Young tableaux, and much more. We shall come back to some of those issues in a later chapter.

6.3.2 Pauli matrices

It is convenient to make manifest the spin-$\frac{1}{2}$ nature of the six-vertex model by reexpressing things in terms of the Pauli matrices

$$\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma^y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{6.23}$$

The arrow conservation then means that the transfer matrix $t(u)$ commutes with the total magnetisation

$$S^z = \frac{1}{2} \sum_{m=1}^{L} \sigma_m^z \,. \tag{6.24}$$

The Temperley-Lieb generator can be written as

$$E_m = \frac{1}{2} \left[\sigma_m^x \sigma_{m+1}^x + \sigma_m^y \sigma_{m+1}^y - \cos\gamma \, (\sigma_m^z \sigma_{m+1}^z - I) - i \sin\gamma \, (\sigma_m^z - \sigma_{m+1}^z) \right] . \tag{6.25}$$

6.3.3 Spectral parameter and anisotropy

The physical meaning of the spectral parameter u is that it controls the spatial anisotropy of the system. To see this qualitatively, note that in the $u \to 0$ limit, the $\check{R}$-matrix is proportional to the identity by (6.18). The transfer matrix $t(u)$ thus acts on a state just by shifting all spins one unit to the right (with periodic boundary conditions); note that this follows from the fact that time propagates in the North-East direction.

In a 1+1 dimensional quantum mechanical analogy, the $u \to 0$ limit thus means that interactions between spins happen very slowly. Equivalently, the time direction has been stretched with respect to the spatial direction. A homogeneous system can be retrieved by rescaling time by a certain anisotropy factor $\zeta(u)$.

It does not appear feasible to determine $\zeta(u)$ without invoking certain results of conformal field theory. Suffice it here to say that one finds

$$\zeta(u) = \sin\left(\frac{\pi u}{\gamma}\right) . \tag{6.26}$$

This predicts that the isotropic point $\zeta(u) = 1$ occurs for $u = \frac{\gamma}{2}$. One then has by (6.18) that $\check{R} \propto I + E$, a fact that can be accounted for geometrically within the loop representation of the Temperley-Lieb algebra.

6.3.4 Spin chain hamiltonian

Using (6.18) we thus see that in the completely anisotropic limit $u \to 0$ the transfer matrix becomes

$$t(0) = \sin^L(\gamma)\, \mathrm{e}^{-iP} , \tag{6.27}$$

where e^{-iP} is the shift operator that translate the lattice sites one unit to the right. Equivalently, P can be interpreted as the momentum operator.

We know from the path-integral formalism that the transfer matrix (the time evolution operator) is the exponential of the quantum hamiltonian. To make things completely precise, note that to first order in u, one may omit on of the factors $\sin(\gamma - u)I$ in (6.18) and take $\sin(u)E$ instead. The correct development in the limit $u \to 0$ therefore reads

$$t(u) \simeq t(0) \exp\left[-\frac{u}{\sin\gamma} H\right] , \tag{6.28}$$

where H is the Hamiltonian of the spin chain. Equivalently

$$H = -\sin\gamma \left.\frac{\partial}{\partial u} \log t(u)\right|_{u\to 0} = -\sin\gamma\, t(0)^{-1} t'(0) . \tag{6.29}$$

Here the inverse $t(0)^{-1} = (\sin\gamma)^{-L}\mathrm{e}^{iP}$ is just the shift in the opposite (left) direction. The derivative $t'(0)$ gives L terms, one for each of the factors in the product (6.5). Using (6.18) we have $\check{R}'(0) = -\cos\gamma\, I + E$. Therefore

$$H = L\cos\gamma\, I - \sum_{m=1}^{L} E_m . \tag{6.30}$$

Inserting the expression (6.25) for the TL generators in terms of Pauli matrices, the piece in $i \sin\gamma \, (\sigma_m^z - \sigma_{m+1}^z)$ simplifies by telescopy. With open boundary condition it would become a surface magnetic field acting on the first and last spins. We consider instead periodic boundary conditions, so this term vanishes alltogether. One is left with

$$H = -\frac{1}{2} \sum_{m=1}^{L} \left[\sigma_m^x \sigma_{m+1}^x + \sigma_m^y \sigma_{m+1}^y + \Delta(\sigma_m^z \sigma_{m+1}^z + I) \right] , \tag{6.31}$$

where we recall that $\Delta = -\cos\gamma$.

We thus arrive at the Hamiltonian of a Heisenberg-type spin chain, where however the interaction is anisotropic along the z-direction. For that reason, this is called the *XXZ spin chain* with anisotropy parameter Δ.[19]

Let us emphasize that due to the commutativity of transfer matrices, the eigenvectors of the six-vertex model transfer matrix and of the XXZ spin chain Hamiltonian are *identical*. It is thus equivalent to diagonalise one or the other, and in that sense the two models are equivalent.

[19] We are here referring to an anisotropy between the different components of the interaction in the space direction. This should not be confused with the space-time anisotropy linked with the spectral parameter u.