Part 2

Algebraic Bethe Ansatz

I would rather discover one cause than gain the kingdom of Persia.

Democritus (c. 460 BC—c. 370 BC)

Hans Bethe's 1931 seminal paper expanded on the one-dimensional version of the model of interacting magnetic moments, suggested by Heisenberg (1928) only a few years earlier. In today's terminology, the method Bethe developed is called the coordinate Bethe ansatz. We develop this approach in part III.

Another exact solution of a model for magnetism, the two-dimensional Ising model, was discovered by Onsager (1944). The Ising model, though conceived by Lenz (1920) to describe magnetism, and solved in its one-dimensional form by Ising (1925) as a contribution to the theory of ferromagnetism, is really a classical model as the 'spins' in this model are described by numbers rather than quantum mechanical operators. However, this fact neither diminishes the usefulness of the Ising model nor does it belittle the formidable mathematical feat of finding its exact solution in two dimensions.

Over the years a number of further exact solutions of two-dimensional classical statistical mechanical models and one-dimensional quantum mechanical models were found. It became apparent that these models and their solutions bore strong connections that were then formalized in the intimately connected quantum inverse scattering and algebraic Bethe ansatz methods, the former used to construct the models and demonstrate their integrability, the latter to find equations determining their spectra.

It must be noted that quantum integrability, as opposed to classical integrabilty, is far from being a uniquely defined concept. It remains an open question what relation there is between competing concepts of quantum integrability and whether and in what sense they may be equivalent. For discussions of the ongoing quest for a consistent definition of quantum integrability, see Weigert (1992), Caux and Mossel (2011), Larson (2013), and Batchelor and Zhou (2015). Here, however, we only develop the quantum inverse scattering method and the algebraic Bethe ansatz and the associated concept of quantum integrability, most recently referred to as Yang–Baxter integrability. We will not explore the relation of this quantum integrability concept with other proposals.

It is also for the sake of this concept of quantum integrability that we chose an ahistorical sequence in which to present the algebraic and the coordinate Bethe ansatz. The latter clearly demonstrates exact solutions of the interacting one-dimensional quantum models for which it works but leaves open a deeper understanding of the reasons why these models admit exact solutions, and does also not shed light on whether and in what sense these models might be quantum integrable. Starting our exposition with the algebraic Bethe ansatz thus provides a much deeper grounding also for our later exposition of the coordinate Bethe ansatz method, which remains a valuable tool in the arsenal of Bethe ansatz methods, especially since it may be considered to be more intuitive.

Our general strategy to introduce the algebraic Bethe ansatz is via the relation between two-dimensional classical statistical mechanical models whose prototypes are the Ising and ice-type models and one-dimensional quantum mechanical models whose prototype is the nearest–neighbour spin-1/2 Heisenberg quantum spin chain. In chapter 6, we reviewed the general connection based on the analogy between thermal and quantum fluctuations between statistical mechanics on the one side and quantum mechanics and quantum field theory on the other side. Here we discuss model systems in two and one dimensions, respectively, where this connection can be implemented in an exact way so that exact solutions of the two-dimensional model translate into exact solutions of the one-dimensional model.

The original physical motivation to study vertex models came from a peculiar observation of a thermodynamic property of (three-dimensional) ice at very low temperatures. Therefore, chapter 9 considers this property of ice and draw motivation from it to study a two-dimensional version of 'ice'.

Chapter 10, looks at general vertex models and introduce the machinery to compute the partition function and, hence, the thermodynamic properties of these models. As the key mathematical construct, we will again discuss the transfer matrix introduced in earlier chapters.

Chapter 11 develops the concept of quantum integrability for one-dimensional quantum systems will emerge. We develop this concept by specializing to the two-dimensional six-vertex model. The connection between the two- and one-dimensional models is brought about by the transfer matrix. On the one hand, its diagonalization is achieved in an algebraic manner and leads to a set of algebraic equations that encode the solution of the two-dimensional model. On the other hand, an infinite number of commuting quantum Hamiltonians can be constructed from the transfer matrix, among them the Hamiltonian of a Heisenberg quantum spin chain. Thus, with the transfer matrix, we have simultaneously diagonalized these quantum Hamiltonians and demonstrated that this Heisenberg quantum spin chain is a quantum integrable model.

Chapters 10 and 11 are partly inspired by Leon Takhtajan's (1985) review and, in particular, a special lecture series by Tuong Truong (1987)—Ferenc Woynarovich and the author were the only students.

Finally, chapter 12 demonstrates how the algebraic version of the Bethe ansatz can be used for models when there is no underlying two-dimensional model, or if we do not know of one. In such cases, it is sometimes possible to conjecture the building blocks

of the transfer matrix. The model we consider for this purpose is from quantum optics describing strong light-matter interactions, the Tavis–Cummings model introduced in section 8.10 as one of the descendent models of the Rabi model.

For vertex models and their exact solutions, the classical text by Baxter (1982) is invaluable. A similarly classical text on the quantum inverse scattering method and the algebraic Bethe ansatz is Korepin *et al.* (1993), while Slavnov (2018) leads up to recent developments.

9

Ice Model

There must be some definite cause why, whenever snow begins to fall, its initial formation invariably displays the shape of a six-cornered starlet. For if it happens by chance, why do they not fall just as well with five corners or with seven?

Johannes Kepler (1571–1630)

The crystal structure of several of the phases of ice shows a peculiarity associated with a special type of disorder (Ziman, 1979). In the words of Linus Pauling (1960): 'there is one proton along each oxygen-oxygen axis, closer to one or the other of the two oxygen atoms.' This peculiarity, arising from a double well potential along the oxygen-oxygen axis, so that the much smaller proton can sit in one or the other potential minimum, depends only on the configuration and is, hence, independent of temperature. It gives rise to a finite entropy of ice, even at zero temperature: the *residual entropy*.

Ice is a particular example of a class of hydrogen-bonded crystalline materials. There are, for instance, compounds like potassium dihydrogen phosphate (KH_2PO_4), commonly known as KDP where the electric dipoles of the crystal exhibit parallel order and which is, thus, a ferroelectric (cp. section 11.2.1). In ammonium dihydrogen phosphate ($NH_4H_2PO_4$, ADP), the electric dipoles have a preferred antiparallel orientation.

We use the observation of a residual entropy in these systems in section 9.1 as a physical motivation to study, in chapters 10 and 11, a certain type of *two-dimensional* statistical mechanical models, the so-called *vertex* models. A precise definition of the exemplary vertex model, the ice model, appears in section 9.2.

9.1 Physical motivation for the square lattice ice model

Careful experiments by Giauque and Stout (1936) on the heat capacity of ice down to very low temperatures led to the extrapolation that there will remain a residual entropy of ice crystals at zero temperature. Experimentally they obtained the value[1]

$$\frac{S}{N} \approx 0.41 \approx \ln(1.5) \tag{9.1}$$

for the entropy per oxygen atom.

[1] Boltzmann's constant has been set to unity.

Models of Quantum Matter: A First Course on Integrability and the Bethe Ansatz. Hans-Peter Eckle, Oxford University Press (2019). © Hans-Peter Eckle. DOI: 10.1093/oso/9780199678839.001.0001

Concurrent with these experiments, Pauling (1935) suggested an explanation for the residual entropy of ice, based on the observation discussed above.

Many different microscopic configurations of hydrogen atoms with respect to the oxygen atoms are possible which lead to the same macroscopic state. Then already the simplest assumption leads to a finite value for the residual entropy of ice: Assume a lattice of N oxygen atoms as vertices with one proton per bond (i.e. oxygen-oxygen axis), hence $2N$ hydrogen atoms in total. The proton can be in one of two positions: near to or far from one of the two oxygen atoms, and, consequently, in the opposite position with respect to the other oxygen atom. Therefore we have, ignoring boundaries and, more importantly, possible restrictions on them, 2^{2N} possible configurations. The number of possible microscopic configurations is therefore

$$Z = 2^{2N} = 4^N, \tag{9.2}$$

which leads to the residual entropy per lattice site

$$\frac{S}{N} = \ln(4). \tag{9.3}$$

This value is much too large compared to the experimental result. A simple consideration, going back to Bernal and Fowler (1933) and applied to the problem of the residual entropy of ice by Pauling (1935, 1960), improves this result considerably.

Since the oxygen atoms in ice are fourfold coordinated (cf. figure 9.1), each oxygen atom can be surrounded by either zero, one, two, three, or four hydrogen atoms in the adjacent position. This accounts for a total of 16 configurations. Pauling argued that only six out of these 16 configurations are, in fact, allowed, namely, those where there are two hydrogen atoms surrounding the oxygen atom. Therefore, we get, instead of (9.2),

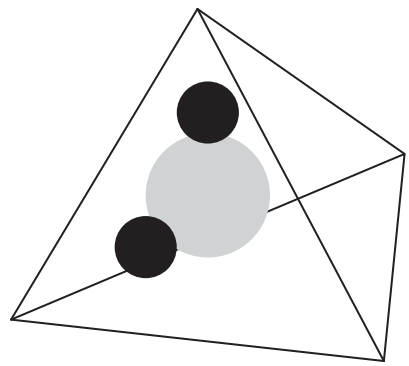

Figure 9.1 *Schematic representation of the coordination of water in an ice crystal.*

$$Z = 2^{2N}\left(\frac{6}{16}\right)^{N} = \left(\frac{3}{2}\right)^{N}, \tag{9.4}$$

and, thus,

$$\frac{S}{N} = \ln\left(\frac{3}{2}\right). \tag{9.5}$$

This value is in good agreement with the experimental result of Giauque and Stout (1936).

Introducing the number W such that $S = N\ln(W)$, numerical calculations by Nagle (1966) confirmed the value of $W \approx 1.5$.

Moreover, numerical calculations, also by Nagle (1966), for a caricature model of ice in two spatial dimensions yield

$$W \approx 1.540 \pm 0.001. \tag{9.6}$$

The latter result prompted mathematical physicist Elliott Lieb to investigate whether an exact solution of the problem of two-dimensional square ice was within reach (Lieb, 1967*a*, 1967*b*): 'It seemed worthwhile to try and find the value of W exactly for the following reasons: (a) It will serve as a check of Nagle's calculations. (b) It is an interesting graph theoretic problem. (c) It is the first step toward the solution of much more interesting problems having phase transitions ...' Lieb (1967*b*).

Let us give a definition of the ice model before we return to Lieb's result for W.

Square ice, also called the *ice model*, is defined on a two-dimensional square lattice. Put N 'oxygen' atoms at the vertices of the two-dimensional square lattice and 'decorate' the edges with one of the $2N$ 'hydrogen' atoms. Each of the $2N$ 'hydrogen' atoms can be in one of the two positions, close or far with respect to a given oxygen atom, as depicted in figure 9.2.

Lieb (1967*a*, 1967*b*) solved the ice model using transfer matrices. He obtained the exact value

$$W = \left(\frac{4}{3}\right)^{\frac{3}{2}} = 1.5396007\ldots \tag{9.7}$$

Although we shall, in the following, present the transfer matrix method in some detail, our main interest is not in two-dimensional statistical mechanical models such as the ice model. These we rather use as a convenient entry point to study one-dimensional quantum mechanical models. In particular, the approach via two-dimensional statistical mechanical models will enable us to understand the integrability of the corresponding one-dimensional quantum models. For the same reason—our primary interest in one-dimensional quantum models—we shall also not present the details of Lieb's calculation of W for the ice model.

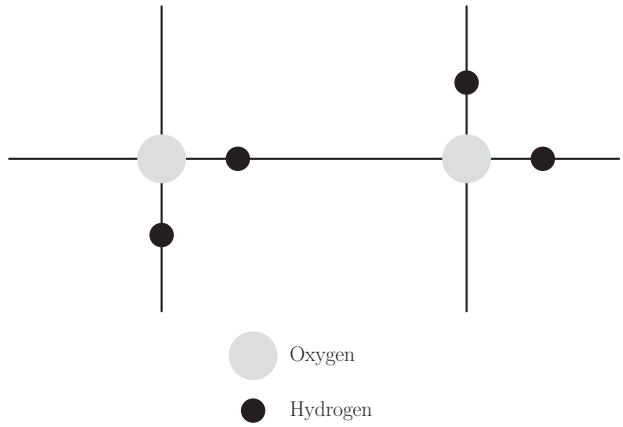

Figure 9.2 *Square ice.*

Let us reiterate the definition of the square lattice ice model and introduce some useful alternative graphical representation in section 9.2.

9.2 Definition of the ice model

The ice model is a two-dimensional model for square ice, i.e. it is a statistical mechanical model, as just described. However, we also need the (figure 9.3):

- *Ice rule*: there are exactly two hydrogen atoms near each oxygen atom.

We translate the 'ice picture' into the *vertex picture*: a hydrogen atom close to the oxygen atom, the vertex, is represented by an arrow pointing toward the oxygen atom, and a hydrogen atom far from the oxygen atom by an arrow pointing outward. Another way of decorating the lattice consists of using thick lines for arrows pointing downward or to the left.

While we will not use thick and thin lines to describe the ice model, we use the arrow picture, and call the ice model from now on mostly the six-vertex model, the oxygen atoms forming the vertices to or from which the arrows on the bonds or edges point.

The following exercise encourages playing with a small lattice to become familiar with the notions just introduced. Playing with small lattices is always a good idea to understand what is going on[2].

[2] Legend has it that Lars Onsager started his investigation of the two-dimensional Ising model by playing with small lattices.

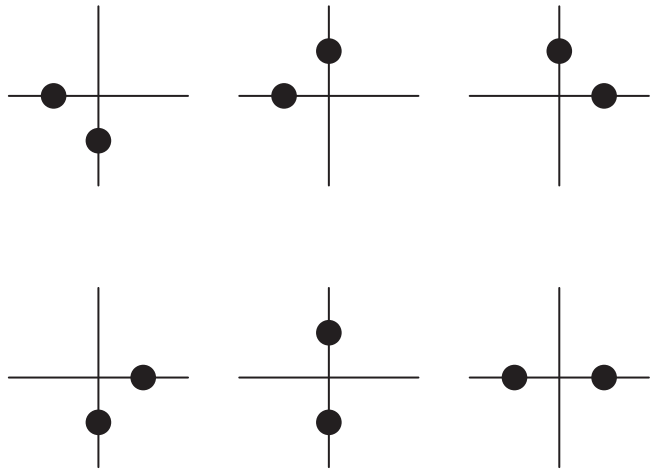

Figure 9.3 *The six local configurations for square ice allowed by the ice rule. The oxygen atoms at the vertices are not shown.*

EXERCISE 9.1 Various ways of decorating a small square lattice Draw a small square lattice, e.g. a 3×3 lattice. The vertices of the lattice correspond to oxygen atoms as described in the text. Decorate this lattice with hydrogen atoms such that the ice rule is satisfied. Now draw the lattice again and decorate it with the arrow configuration corresponding to the first lattice. Finally, draw the lattice a third time, and decorate it with thick and thin lines corresponding to the first two configurations.

- How did you treat the horizontal and vertical boundaries of the small finite lattice?
- What is the advantage of the last picture of thick/thin lines?

10

General Square Lattice Vertex Models

Whoever, in the pursuit of science, seeks after immediate practical utility, may generally rest assured that he will seek in vain.

Hermann von Helmholtz (1821–1894)

The class of models introduced in the previous chapter, the vertex models, is of physical interest in its own right. However, we study them for a particular reason: vertex models are a convenient path to introduce one-dimensional *integrable* quantum mechanical models and the Bethe ansatz. Quite apart from the questions of integrability, there is a correspondence between statistical mechanical lattice models, which are really classical models, and quantum mechanical models. As discussed in chapter 6, this correspondence relates the statistical mechanical models in $(d + 1)$ dimensions to quantum mechanical models in d dimensions (see, e.g. section 1 'Introduction' in Shenker (1982) for an elementary discussion).

Relaxing the ice rule, we can define vertex models on the square lattice with more than six possible vertices. Even more general vertex models are possible, as we shall briefly outline in section 10.1. The general two-dimensional vertex model on a square lattice has $n = 16$ different types of vertices, eight of which, shown in the upper two rows of figure 10.1, have an even number of outward and inward arrows. Especially vertex type 7 and 8 have all arrows outward and inward, respectively, and are, also respectively, sources and sinks of arrows. The other eight vertex types, in the lower two rows of figure 10.1, have either three inward and one outward arrow, or vice versa. Section 10.2 briefly summarizes the sixteen- and eight-vertex models before we examine them in the first step using concepts from chapter 4 on statistical mechanics, in particular, the partition function and the Boltzmann weight (section 10.3) for a vertex. These concepts are then tailored to the special case of the two-dimensional square lattice vertex models by the introduction of the R-matrix in section 10.4, which represents the possible vertex weights in matrix form. With the help of the R-matrix, another object which we encountered before, i.e. the transfer matrix, can be constructed in a systematic way. The transfer matrix also encodes the integrability of the two-dimensional vertex models, discussed in sections 10.5 and 10.6. In order to shed further light on the integrability concept, we introduce two more mathematical objects, the monodromy matrix and the L-matrix in section 10.7, the latter object being further investigated in section 10.8. Eventually, the concept of integrability can be subsumed in the Yang–Baxter relations discussed in section 10.9, expanded on in section 10.10, and then further exploited in section 10.11.

Models of Quantum Matter: A First Course on Integrability and the Bethe Ansatz. Hans-Peter Eckle, Oxford University Press (2019). © Hans-Peter Eckle. DOI: 10.1093/oso/9780199678839.001.0001

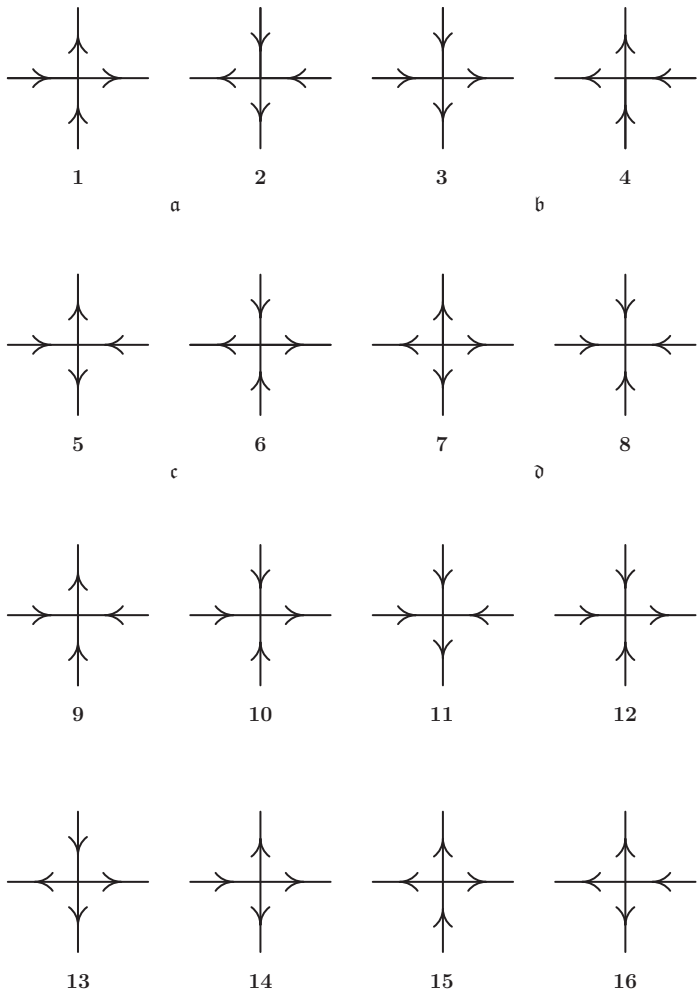

Figure 10.1 *The sixteen possible vertex configurations of the general square lattice vertex model (sixteen-vertex model). The energies and vertex weights (cf. main text) are* ϵ_j *and* $v_j, j = 1, \ldots, 16$, *respectively. The first eight vertex configurations are grouped into arrow reversal symmetric pairs labelled by* $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, *and* $\mathfrak{d}$. *Furthermore, these eight vertex configurations exhibit an even number of arrows going in or out, respectively, two arrows going in or out for configurations* $\mathfrak{a}$, $\mathfrak{b}$, *and* $\mathfrak{c}$, *four arrows in or out for configuration* $\mathfrak{d}$. *These eight vertices constitute the aptly named eight-vertex model. Considering only the first six vertices, those which obey the ice rule, leads to the six-vertex model.*

We learn that there are no vertex weights and corresponding R-matrix known for the general sixteen-vertex model that satisfy the Yang–Baxter relations. However, there are solutions in this sense for the eight- and six-vertex models. Chapter 11 pursues the solution of the six-vertex model, the eight-vertex model being considerably more involved technically.

10.1 Vertex models in two dimensions

If we allow also for bonds with no arrows, and/or other lattices, e.g. triangular or honeycomb lattices, other numbers of allowed vertices (e.g. seven and nineteen) can be considered. Such models are of interest, e.g. to study surface and interface phenomena, since bonds with no arrows can be interpreted as bonds with no interaction between the sites connected.

For example, Batchelor *et al.* (1989) discuss a nineteen-vertex model on a square lattice, Batchelor and Blöte (1989) consider a seven-vertex model on a honeycomb lattice, and Blum and Shapir (1990) explore a seven-vertex model on a square lattice.

Here, we concentrate our attention on the sixteen-vertex model on the square lattice and, in particular, on its special cases, the eight- and six-vertex models.

10.2 Sixteen- and eight-vertex models

For the time being, the number of allowed vertices will make no difference. Therefore, we can keep the discussion general for some time. Our focus will, for the present, be on the sixteen-vertex model in figure 10.1, which we shall eventually abandon for the eight-vertex model. The reason for this focus is that, while the eight-vertex model (Sutherland, 1970) has been solved in zero external field by Baxter (1971; see also Baxter, 1982), there is no solution known for the sixteen-vertex model. However, the sixteen-vertex model has been of some interest in the proof of the equivalence of boundary conditions of the zero field six-vertex model (Brascamp *et al.*, 1973),[1] where the sixteen-vertex model was used at an intermediate step of the proof, after which the limiting case to the six-vertex model was taken, i.e. $v_j \to 0$ for $j = 9, 10, \ldots, 16$.

10.3 Vertex Boltzmann weights and the partition function

In order to have some flexibility, let us call s the number of allowed vertices. We assign an energy

$$\epsilon_j, \qquad j = 1, \ldots s \tag{10.1}$$

to each of the s vertices. Equivalently, a vertex of type j can be assigned a Boltzmann or statistical weight, called vertex weight in this connection, by

$$v_j = e^{-\beta\epsilon_j}. \tag{10.2}$$

Then, a given arrangement or configuration of vertices has an energy

[1] I thank Tony Dorlas for drawing my attention to this point and the relevant literature.

$$\mathcal{E} = n_1\epsilon_1 + n_2\epsilon_2 + \ldots + n_s\epsilon_s = \sum_{j=1}^{s} n_j\epsilon_j, \tag{10.3}$$

where n_j is the number of vertices of type j in this configuration of vertices. The partition function becomes

$$Z = \sum_{\text{allowed configurations}} \exp(-\beta\mathcal{E}), \tag{10.4}$$

i.e. a vertex appearing n_j times in the configuration contributes with a statistical weight of $\exp(-\beta n_j\epsilon_j) = v_j^{n_j}$ to the partition function. Arrows are a good way to draw and visualize configurations of vertices, but they are a bit tedious for algebraic purposes. Therefore we assign the numbers $\sigma = \pm 1$ to the arrows according to the following correspondences

$$\uparrow \text{ and } \rightarrow \quad \Leftrightarrow \quad \sigma = +1, \tag{10.5}$$

$$\downarrow \text{ and } \leftarrow \quad \Leftrightarrow \quad \sigma = -1. \tag{10.6}$$

The numbers σ are called spins or spin variables in the sense of the classical spin variables of, e.g., the Ising model. They are *not* quantum mechanical spins.

The ice rule for the six-vertex model, written in spin variables is simply (cf. figure 10.2)

$$\alpha' + \gamma' = \alpha + \gamma. \tag{10.7}$$

The s energies ϵ_j or, equivalently, vertex weights v_j characterize the vertex model under consideration. They must be chosen according to the physics we want to describe with the

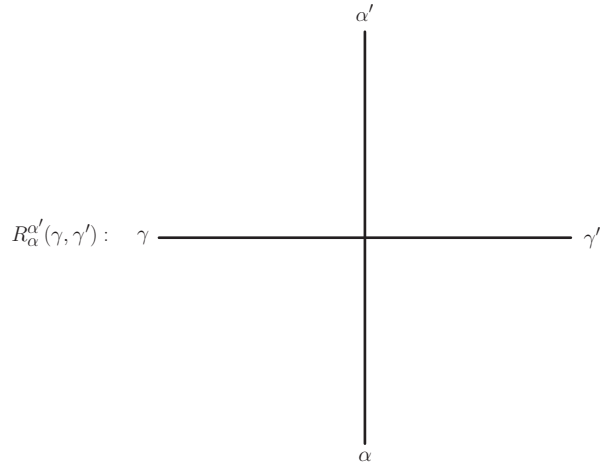

Figure 10.2 *The R-matrix element* $R_\alpha^{\alpha'}(\gamma, \gamma')$.

vertex model. However, these s values, energies or vertex weights, are fine to describe an individual vertex, but they are not yet very suitable to describe a whole two-dimensional lattice of vertices, or even just a single row of vertices. They have yet to be arranged in a proper way for that purpose.

This proper arrangement can be achieved by an object, called the R-matrix.

10.4 R-matrix: matrix of Boltzmann weights of a vertex

Each appearance of one of the $s = 16$ vertices contributes with a Boltzmann weight

$$v_j = \mathrm{e}^{-\beta\epsilon_j} \qquad j = 1, \ldots, 16 \tag{10.8}$$

to the partition function. It will soon prove important to arrange these 16 vertex weights in a 4×4 matrix, the so-called R-matrix. More precisely, a vertex configuration depends on four spin or arrow variables which we denote by $\alpha = \pm 1, \alpha' = \pm 1$ and $\gamma = \pm 1, \gamma' = \pm 1$, such that the Boltzmann weights of a single vertex can be represented by a 4×4 matrix, the R-matrix. The matrix elements

$$R_{\alpha}^{\alpha'}(\gamma, \gamma') \tag{10.9}$$

are defined by figure 10.2 and have entries

$$v_j = \exp(-\beta\epsilon_j) \qquad j = 1, \ldots, 16. \tag{10.10}$$

Figure 10.1 show, the sixteen possible arrow arrangements that may be translated into elements of the R-matrix with the help of the rules (10.5) and (10.6).

As the matrix elements in (10.9) initially may not look like proper matrix elements, let us spent a moment to clarify this point. Group together the unprimed and primed indices in (10.9) (cf. also figure 10.2) and make the following assignation

$$\{(\alpha, \gamma)\} = \{(+,+), (+,-), (-,+), (-,-)\} \quad \text{corresponding to} \quad \{1, 2, 3, 4\} \tag{10.11}$$

such that, with a similar assignation for $\{(\alpha', \gamma')\}$, the index pair $\{(\alpha, \gamma)\}$ corresponds to a row index and the index pair $\{(\alpha', \gamma')\}$ corresponds to a column index for the R-matrix elements.

We shall later often use a notation of the R-matrix where it is partitioned in blocks of 2×2 matrices according to

$$R = \left(R_{\alpha}^{\alpha'}(\gamma, \gamma')\right) = \begin{pmatrix} R_{+}^{+}(\gamma, \gamma') & R_{+}^{-}(\gamma, \gamma') \\ R_{-}^{+}(\gamma, \gamma') & R_{-}^{-}(\gamma, \gamma') \end{pmatrix} \equiv \begin{pmatrix} \hat{\alpha} & \hat{\beta} \\ \hat{\gamma} & \hat{\delta} \end{pmatrix}, \tag{10.12}$$

where $\{(\gamma, \gamma')\} = \{(+,+), (+,-), (-,+), (-,-)\}$.

From the vertices of figure 10.1 and the correspondences between arrows and spin variables, the arrangement of the weights v_j in the R-matrix, and hence the R-matrix itself, is now determined (once we have assigned the weights v_j).

The most important cases we discuss in some more detail further on and that are the integrable cases, i.e. the eight- and six-vertex models. Let us therefore give the R-matrices for these two cases explicitly introducing the convenient '$\mathfrak{abcd}$'-notation for the vertex weights (cf. the pairs of vertices labelled $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$, and $\mathfrak{d}$ in figure 10.1), where due to arrow reversal symmetry in the absence of external fields we have

$$R^+_+(+,+) = v_1 = R^-_-(-,-) = v_2 \equiv \mathfrak{a}, \tag{10.13}$$

$$R^-_-(+,+) = v_3 = R^+_+(-,-) = v_4 \equiv \mathfrak{b}, \tag{10.14}$$

$$R^+_-(+,-) = v_5 = R^-_+(-,+) = v_6 \equiv \mathfrak{c}, \tag{10.15}$$

$$R^+_-(-,+) = v_7 = R^-_+(+,-) = v_8 \equiv \mathfrak{d}, \tag{10.16}$$

or in matrix notation with the conventions (10.11) and (10.12), respectively,

$$R_{8\mathrm{v}} = \begin{pmatrix} \mathfrak{a} & 0 & 0 & \mathfrak{d} \\ 0 & \mathfrak{b} & \mathfrak{c} & 0 \\ 0 & \mathfrak{c} & \mathfrak{b} & 0 \\ \mathfrak{d} & 0 & 0 & \mathfrak{a} \end{pmatrix} = \begin{pmatrix} \hat{\alpha} & \hat{\beta} \\ \hat{\gamma} & \hat{\delta} \end{pmatrix}. \tag{10.17}$$

For the six-vertex model, the vertex weight $\mathfrak{d} = 0$. Later we shall evaluate the R-matrix in the '$\mathfrak{abc}$'-notation in order to find the exact solution of the six-vertex model.

For later use, it is also helpful to note that the R-matrix of the eight-vertex model can be written as a tensor product of Pauli matrices σ^j for $j = 1, 2, 3$ (or $j = x, y, z$) and the 2×2 unit matrix $\sigma^4 \equiv I_2$

$$R_{8\mathrm{v}} = \sum_{j=1}^{4} w_j \sigma^j \otimes \sigma^j. \tag{10.18}$$

For an introduction to and more on the tensor product, see section 10.7.1.

EXERCISE 10.1 Vertex weights of the eight-vertex model Determine the relations between the vertex weights $\{v_j\}$ and $\{w_j\}$ of the eight-vertex model.

But first, before we eventually focus on the integrable eight- and six-vertex cases, we further pursue the general properties of the R-matrix.

The R-matrix is the fundamental building block of the vertex model that we use to build up the two-dimensional lattice step by step, beginning with a single row, then attaching further rows.

If we put two vertices together, described by the R-matrix elements $R^{\alpha_1'}_{\alpha_1}(\gamma_1, \gamma_1')$ and $R^{\alpha_2'}_{\alpha_2}(\gamma_2, \gamma_2')$, respectively, the 'inner' spin variables γ_1' and γ_2 have to be equal or the arrows have to point in the same direction (i.e. both left or both right), i.e. we shall have to equate the 'inner' spin indices $\gamma_1' = \gamma_2$ (cf. figure 10.3).

Moreover, we can sum over this inner variable. We obtain the following object

$$\sum_{\gamma_2 = \pm} R^{\alpha_1'}_{\alpha_1}(\gamma_1, \gamma_2) R^{\alpha_2'}_{\alpha_2}(\gamma_2, \gamma_2'), \tag{10.19}$$

which represents the statistical weight of two vertices. There are six unspecified spin variable $\alpha'_{1,2}, \alpha_{1,2}$, and γ_1 and γ_2', and thus (10.19) represents $2^6 = 2^3 \cdot 2^3 = 64$ matrix elements. It is important to note that this construction, although it may superficially resemble one, is *not* an element of a matrix product. Below we describe how it can be expressed as a tensor product of the two R-matrices, each written as a 2×2 matrix of 2×2 blocks of matrix elements according to the convention (10.12).

In the same way, we can build a whole row of, say, N vertices and we obtain the matrix elements of a large $2^{N+1} \times 2^{N+1}$ matrix

$$T_{\{\alpha\},\{\alpha'\}}(\gamma_1, \gamma_N') = \sum_{\gamma_2} \cdots \sum_{\gamma_N} R^{\alpha_1'}_{\alpha_1}(\gamma_1, \gamma_2) R^{\alpha_2'}_{\alpha_2}(\gamma_2, \gamma_3) \ldots R^{\alpha_N'}_{\alpha_N}(\gamma_N, \gamma_N') \tag{10.20}$$

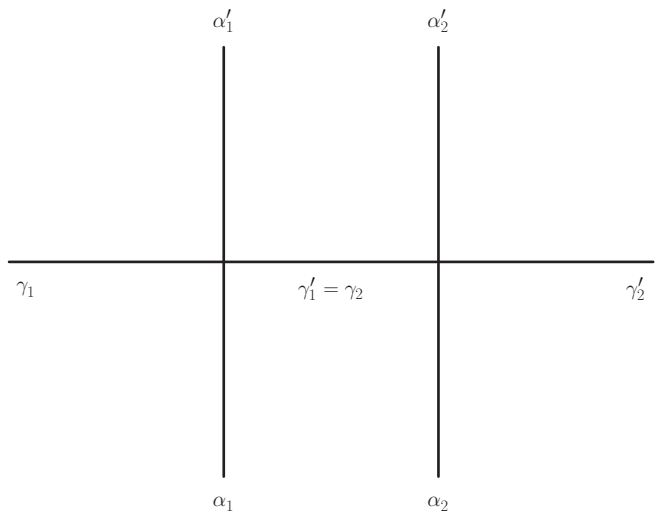

Figure 10.3 $\sum_{\gamma_2=\pm} R^{\alpha_1'}_{\alpha_1}(\gamma_1, \gamma_2) R^{\alpha_2'}_{\alpha_2}(\gamma_2, \gamma_2')$*: The R-matrices of two vertices combined together to obtain a matrix element of their statistical weight.*

with N unspecified upper spin variables $\{\alpha'\}$ and N unspecified lower spin variables $\{\alpha\}$. Also, the first and last horizontal spin variable γ_1 and γ'_N are not yet specified. The latter we specify through boundary conditions, the simplest of which (and the ones we shall adopt) are periodic ones, i.e. $\gamma_1 = \gamma'_N$.[2] Therefore, we can also sum over $\gamma_1 = \pm 1$ in (10.20) to obtain the $2^N \times 2^N$ matrix

$$T_{\{\alpha\},\{\alpha'\}} = \sum_{\gamma_1} \dots \sum_{\gamma_N} R^{\alpha'_1}_{\alpha_1}(\gamma_1, \gamma_2) R^{\alpha'_2}_{\alpha_2}(\gamma_2, \gamma_3) \dots R^{\alpha'_N}_{\alpha_N}(\gamma_N, \gamma_1). \tag{10.21}$$

This matrix will prove to be another central object, called the transfer matrix. It has still many 'dangling', i.e. unspecified or free, indices: $\{\alpha\}, \{\alpha'\}$.

The transfer matrix, in a sense a one-dimensional object since it describes a row of vertices, can now be used to build up the two-dimensional square lattice row by row. For example, the statistical weight of a lattice consisting of only two rows can now be written as the matrix product of two transfer matrices

$$T^2_{\{\alpha\},\{\alpha''\}} = \sum_{\{\alpha'\}} T_{\{\alpha\},\{\alpha'\}} T_{\{\alpha'\},\{\alpha''\}}. \tag{10.22}$$

Note that (10.22), in contradistinction to (10.19), (10.20), and (10.21), is indeed a matrix product.

We repeat this procedure M times. The resulting matrix elements will depend on the lower and upper row of outer horizontal spin variables $\{\alpha\} = \alpha_1, \dots, \alpha_N$ and $\{\alpha^{(M)}\} = \alpha_1^{(M)}, \dots, \alpha_N^{(M)}$, and, hence, will still be the matrix elements of a $2^N \times 2^N$ matrix. This matrix is the Mth power T^M of the transfer matrix T.

By imposing periodic boundary conditions also in the vertical direction, hence using the topology of a torus (toroidal boundary conditions), we can finally equate and then sum over the last remaining free spin variables $\{\alpha\} = \{\alpha^{(M)}\}$. With this last summation we have performed a trace in the space of the $2^N \times 2^N$ matrices. On the other hand, in total we have summed the statistical weights of all possible configurations of the two-dimensional square lattice with toroidal boundary conditions, which is nothing else than the partition function of the lattice model. We are therefore allowed to write

$$Z = \mathrm{Tr}(T^M). \tag{10.23}$$

The computation of the partition function Z has now been reduced to the task of calculating the eigenvalues of the transfer matrix T.

The transfer matrix, by construction from the non-negative vertex weights, is a matrix with non-negative matrix elements. The Perron–Frobenius theorem of matrix theory then tells us that, for such a matrix, there is a unique positive real eigenvalue λ_1 such that for all other eigenvalues λ_j

[2] It should be emphasized at this point that other boundary conditions *are* possible, the use of which may lead to quite interesting physics.

$$\lambda_1 > |\lambda_j| \quad j \neq 1 \tag{10.24}$$

holds.

Supposing that we had achieved this task of diagonalizing the transfer matrix T, then, according to the Perron–Frobenius theorem, the partition function becomes

$$Z = \sum_{j=1}^{2^N} \lambda_j^M = \lambda_1^M \left(1 + \left(\frac{\lambda_2}{\lambda_1}\right)^M + \left(\frac{\lambda_3}{\lambda_1}\right)^M + \dots\right), \tag{10.25}$$

which, in the limit $M \to \infty$ reduces to just

$$Z = \lambda_1^M. \tag{10.26}$$

This result looks marvelously simple, but we are, of course, not there yet. We first have to find a way to achieve the diagonalization of the transfer matrix.

10.5 Integrability and the transfer matrix

The transfer matrix allows us to address the question of integrability. We pose the following question: can we diagonalize a whole family of transfer matrices simultaneously? In other words, can we find arbitrary sets of Boltzmann weights v_j and v_j' and the corresponding R-matrices such that the corresponding transfer matrices $T = T(v_j)$ and $T' = T(v_j')$ commute

$$TT' - T'T \equiv [T, T'] = 0. \tag{10.27}$$

Before we search for an answer to this question, we shall pause for a moment and investigate the implications of (10.27).

10.6 Commuting transfer matrices

Recall the expression we found for the partition function of vertex models for the case $s = 8$, i.e. the eight-vertex model

$$Z = \sum_{\text{allowed configurations}} \exp(-\beta\mathcal{E}), \tag{10.28}$$

where

$$\mathcal{E} = n_1\epsilon_1 + n_2\epsilon_2 + \ldots + n_8\epsilon_8 = \sum_{j=1}^{8} n_j\epsilon_j. \tag{10.29}$$

We can also write the partition function in the following form

$$Z = \sum_{\text{allowed configurations}} \mathfrak{a}^{n_\mathfrak{a}}\mathfrak{b}^{n_\mathfrak{b}}\mathfrak{c}^{n_\mathfrak{c}}\mathfrak{d}^{n_\mathfrak{d}}, \tag{10.30}$$

using the $\mathfrak{abcd}$-notation for the vertex weights (cf. figure 10.1)

$$\mathfrak{a} = v_1 = v_2, \quad \mathfrak{b} = v_3 = v_4, \quad \mathfrak{c} = v_5 = v_6, \quad \mathfrak{d} = v_7 = v_8, \tag{10.31}$$

and the numbers of times the vertices appear in a configuration

$$n_\mathfrak{a} = n_1 + n_2, \quad n_\mathfrak{b} = n_3 + n_4, \quad n_\mathfrak{c} = n_5 + n_6, \quad n_\mathfrak{d} = n_7 + n_8. \tag{10.32}$$

In the rectangular lattice of $N \times M$ lattice sites we are considering, these numbers sum up such that $n_\mathfrak{a} + n_\mathfrak{b} + n_\mathfrak{c} + n_\mathfrak{d} = NM$. We assume that there is a spin or arrow inversion symmetry, i.e. flipping all spin variables, $\sigma \to -\sigma$ or, equivalently, reversing all arrows, of a vertex does not change the energy. This symmetry holds quite obviously as long as there are no external fields that might break the symmetry.

Writing the partition function in the form of (10.30) is quite instructive. First, we note that it can be written, e.g. as

$$Z = \mathfrak{d}^{NM} \sum_{\text{allowed configurations}} \left(\frac{\mathfrak{a}}{\mathfrak{d}}\right)^{n_\mathfrak{a}} \left(\frac{\mathfrak{b}}{\mathfrak{d}}\right)^{n_\mathfrak{b}} \left(\frac{\mathfrak{c}}{\mathfrak{d}}\right)^{n_\mathfrak{c}}. \tag{10.33}$$

The partition function thus depends only on the three ratios of vertex weights

$$\frac{\mathfrak{a}}{\mathfrak{d}}, \quad \frac{\mathfrak{b}}{\mathfrak{d}}, \quad \text{and} \quad \frac{\mathfrak{c}}{\mathfrak{d}}. \tag{10.34}$$

We can, thus, describe the eight-vertex model partition function in the three-dimensional space formed by these three vertex weight ratios.

The same conclusions hold also for the transfer matrix T. Generically, transfer matrices for different sets of vertex weights will not commute. However, we shall also see that we can find families of curves, parametrized by a parameter u, in the space of vertex weights $\mathbf{w} \equiv \left(\frac{\mathfrak{a}}{\mathfrak{d}}, \frac{\mathfrak{b}}{\mathfrak{d}}, \frac{\mathfrak{c}}{\mathfrak{d}}\right)^T$

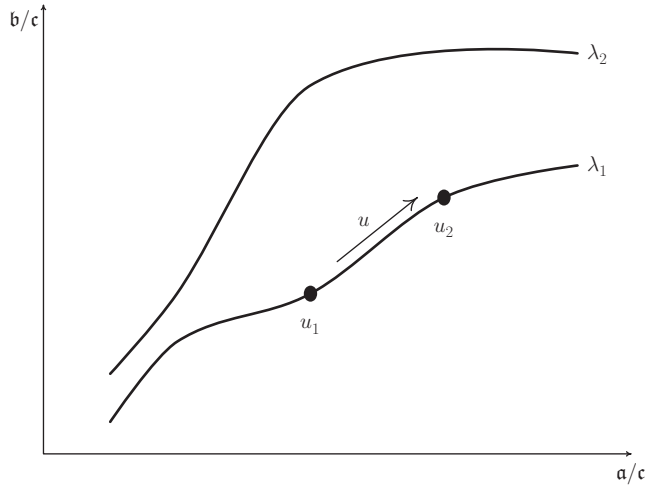

Figure 10.4 *Schematic representation of the families of curves* $\mathfrak{w}(u)$ *in the parameter space of the six-vertex model. The different families are specified by a parameter* λ*, the crossing parameter. For a fixed value of* λ*, pairs of transfer matrices for different positions along the curve, e.g.* u_1 *and* u_2*, commute.*

$$\mathbf{w} = \mathbf{w}(u) \equiv \begin{pmatrix} \mathfrak{a}(u) \\ \mathfrak{b}(u) \\ \mathfrak{c}(u) \end{pmatrix}, \tag{10.35}$$

such that the transfer matrices for different positions, i.e. different values of u, on those curves do commute. Different members of the families of curves are distinguished by a parameter λ, the crossing parameter. The parameter u is called the spectral parameter for reasons that will become clear a little later.

In figure 10.4, this is illustrated for the six-vertex model where there are only two parameters $\left(\frac{\mathfrak{a}}{\mathfrak{c}}, \frac{\mathfrak{b}}{\mathfrak{c}}\right)$ spanning the parameter space, since the $\mathfrak{d}$-vertex in the six-vertex model is not allowed by the ice rule.

Now we return to how we can actually perform this programme of finding commuting transfer matrices.

10.7 Monodromy matrix

In order to make progress toward an answer of the question posed in section 10.5, we recall an object encountered above in (10.20)

$$\mathcal{T} \equiv T_{\{\alpha\},\{\alpha'\}}(\gamma, \gamma') = \sum_{\gamma_2} \cdots \sum_{\gamma_N} R^{\alpha_1'}_{\alpha_1}(\gamma, \gamma_2) R^{\alpha_2'}_{\alpha_2}(\gamma_2, \gamma_3) \ldots R^{\alpha_N'}_{\alpha_N}(\gamma_N, \gamma'), \tag{10.36}$$

namely, the transfer matrix where no boundary conditions have been specified, i.e. where the first and last horizontal spin variables are free. This object bears the name *monodromy matrix*. If we focus on the dependence of the monodromy matrix of these first and last horizontal spin variables, it can be regarded as a 2×2 matrix

$$\mathcal{T} = \mathcal{T}(\gamma, \gamma') \tag{10.37}$$

whose four matrix elements are, however, themselves $2^N \times 2^N$ matrices. The transfer matrix can then be written as a trace in the two-dimensional space of these $2^N \times 2^N$ matrices

$$T = \mathrm{Tr}\,\mathcal{T} = \mathcal{T}(+,+) + \mathcal{T}(-,-). \tag{10.38}$$

The monodromy matrix $\mathcal{T}$, as well as the transfer matrix T, obviously are objects that describe an entire row of vertices. The derivation of $\mathcal{T}$ or T given above is built on a local object, the R-matrix. Next, we want to construct the monodromy matrix $\mathcal{T}$ from a different set of elementary matrices, which are, on the one hand also local objects in the sense that they describe a particular vertex. On the other hand, they are connected with a whole row as we shall see in a moment. The whole point is, similar to when we moved from vertex weights v_j to their neat arrangement as R-matrix, that we now write the R-matrix in such a way that the vertex to which it refers becomes easily localized in the row of vertices we consider. This is what the L-matrix does. The L-matrix carries a position label and is written more precisely as L_n, where n refers to the nth vertex of the row. The R-matrix does not carry such a position label.

First, let us write the R-matrix in a more explicit form as

$$R_{\alpha}^{\alpha'}(\gamma, \gamma')(w_{ij}) = \sum_{i,j=1}^{4} w_{ij} \sigma^{i}_{\gamma,\gamma'} \sigma^{j}_{\alpha,\alpha'} \tag{10.39}$$

where the matrices

$$\sigma^j = \begin{pmatrix} \sigma^j_{+,+} & \sigma^j_{+,-} \\ \sigma^j_{-,+} & \sigma^j_{-,-} \end{pmatrix} \qquad j = 1,2,3,4 \tag{10.40}$$

are the Pauli matrices and the 2×2 unit matrix

$$\begin{aligned} \sigma^1 = \sigma^x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 = \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma^3 = \sigma^z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma^4 = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{10.41}$$

Using the tensor products of the Pauli and unit matrices (10.41), the R-matrix can be written in the form

$$R = \sum_{i,j=1}^{4} w_{ij} \sigma^i \otimes \sigma^j \tag{10.42}$$

whose special case for the eight-vertex model we already encountered.

We won't yet discuss the detailed form of the Boltzmann weights w_{ij}, especially their connection to the previously introduced weights v_j, as they are not relevant yet for the still more formal development of our arguments. We just note now that any complex 4×4 matrix can be written in the form (10.42).

As we shall see, the drawback of the R-matrix is that, although it clearly is an object describing a vertex, it cannot be easily used to locate one particular vertex within the lattice of vertices, i.e. it does not carry a spatial label or coordinate. Motivated by the form (10.42) of the R-matrix, however, we can write down a matrix that does depend on the position within the lattice. More specifically, this matrix will depend on the position of a specific vertex in a row of vertices. This matrix based on the R-matrix in the form (10.42) will be a $2^{N+1} \times 2^{N+1}$ matrix, called the local L-matrix or, more commonly, the L-operator, which is given by

$$L_n(w_{ij}) = \sum_{i,j=1}^{4} w_{ij} \sigma^i \otimes \sigma_n^j \tag{10.43}$$

where

$$\sigma_n^j = \overbrace{I_2 \otimes I_2 \otimes \ldots \underbrace{\sigma^j}_{n} \otimes \ldots I_2}^{N} \tag{10.44}$$

is a $2^N \times 2^N$ matrix that acts non-trivially only on the position n along the row of vertices. In the notation I_2 for the 2×2 unit matrix we added the subscript 2 to distinguish from unit matrices of other dimensions, which we shall introduce shortly.

The L_n matrices are thus matrices located at site n in the row of vertices, as we anticipated above. They are a local version of the R-matrix.

Let us now pause for a brief introduction of the tensor product that has crept up in the foregoing considerations and, thus, merits some attention before we continue.

10.7.1 Interlude: direct or tensor product of matrices

In the following, the direct or tensor product of two square matrices of dimension N is important. Moreover we have just used it in writing down (10.43) and (10.44). Therefore, let us define the tensor product explicitly. However, we restrict ourselves to

the case $N = 2$ mainly to keep the equations neat and manageable. The generalization to arbitrary N and even to matrices partitioned into blocks, themselves square matrices, is straightforward.

For the 2×2 (complex) matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \tag{10.45}$$

the tensor product $A \otimes B$ is defined by[3]

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ \hline a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{array}\right). \tag{10.46}$$

There is an easy way to remember this

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix}. \tag{10.47}$$

The 2×2 matrices act on a two-dimensional complex vector space

$$V = \mathbb{C}^2, \tag{10.48}$$

whose elements are two-component vectors

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{10.49}$$

with $x_i \in \mathbb{C}$.

The tensor product $A \otimes B$ acts on the tensor product space $V \otimes V$ whose elements are of the form

$$x \otimes y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \left(\begin{array}{c} x_1y_1 \\ x_1y_2 \\ \hline x_2y_1 \\ x_2y_2 \end{array}\right). \tag{10.50}$$

[3] The horizontal and vertical lines help to organize the matrix elements; they have no other significance, and can be omitted. This practice follows Nepomechie (1999).

It is useful to recall the equation

$$(A \otimes B)(x \otimes y) = Ax \otimes By. \tag{10.51}$$

which generalizes to

$$(A_1 \otimes A_2 \otimes \ldots A_n)(x_1 \otimes x_2 \otimes \ldots x_n) = A_1 x_1 \otimes A_2 x_2 \otimes \ldots A_n x_n. \tag{10.52}$$

In other words, a 2×2 matrix in the ℓ's position ($1 \leq \ell \leq n$) of an n-fold tensor product acts non-trivially only on the two-dimensional vector space in the ℓ's position of the n-fold tensor product space

$$\overset{1}{\underset{\downarrow}{V}} \otimes \cdots \otimes \overset{\ell}{\underset{\downarrow}{V}} \otimes \cdots \otimes \overset{n}{\underset{\downarrow}{V}}, \tag{10.53}$$

i.e. on the corresponding vector in the ℓ's position of the n-fold tensor product of vectors.

There is an important connection between the ordinary matrix product and the tensor product that we shall be using frequently in the following. For matrices A, B, C, D, or, generally, A_ℓ and B_ℓ ($\ell = 1, \ldots, n$) whose matrix elements are complex numbers, we have

$$(A \otimes B)\,(C \otimes D) = AC \otimes BD \tag{10.54}$$

or

$$(A_1 \otimes A_2 \otimes \ldots \otimes A_n)\,(B_1 \otimes B_2 \otimes \ldots \otimes B_n) = A_1 B_1 \otimes A_2 B_2 \otimes \ldots \otimes A_n B_n. \tag{10.55}$$

What follows are a few simple observations about the tensor product we just defined.

10.7.1.1 Simple applications of the tensor product

Many of the tensor products we are interested in are related to the Pauli matrices and the corresponding spinors

$$|+\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle = |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{10.56}$$

Consider a lattice of two sites only to each of which we attach a quantum spin–$\frac{1}{2}$ object. Then the Pauli matrices acting on the first object only will be

$$\sigma_1^x = \sigma^x \otimes I_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right), \tag{10.57}$$

while for the second object, we have

$$\sigma_2^x = I_2 \otimes \sigma^x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right). \tag{10.58}$$

Similarly, we can obtain the 4×4 matrices σ_i^y and σ_i^z $(i = 1, 2)$.

The following little exercise is suggested to gain practice.

EXERCISE 10.2 Two spins
Show that for two spins we have

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = \sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z \tag{10.59}$$

$$= \sigma^x \otimes \sigma^x + \sigma^y \otimes \sigma^y + \sigma^z \otimes \sigma^z \tag{10.60}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{10.61}$$

The permutation operator (or permutation matrix)

$$\mathcal{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{10.62}$$

acts on the tensor product of two spinors of the tensor product space $V \otimes V$ as

$$\mathcal{P}x \otimes y = \begin{pmatrix} x_1 y_1 \\ x_2 y_1 \\ x_1 y_2 \\ x_2 y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \otimes \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = y \otimes x. \tag{10.63}$$

The permutation operator squares to the 4×4 identity matrix which can be written as the tensor product of the 2×2 identity matrix with itself

$$\mathcal{P}^2 = I_2 \otimes I_2 = I_4. \tag{10.64}$$

Furthermore, combining this with the exercise 10.2 we immediately observe

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 + I_4 = 2\mathcal{P}. \tag{10.65}$$

This is an interesting result to which we will return. For the moment, we note that the two-spin operator

$$h_{i,i+1} = \frac{1}{2}\left(\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} + I_4\right) = \mathcal{P}, \tag{10.66}$$

acting on spinors at neighbouring sites, is just the permutation operator $\mathcal{P}$.

We end this interlude section with an exercise which brings us back to R-matrices.

EXERCISE 10.3 Tensor product of two R-matrices Using the convention (10.12) for partitioning an R-matrix into 2×2 blocks, show that the vertex weight of two vertices (10.19) can be written as a tensor product of R-matrices in block form. Writing the elements of this composite object, we have

$$\sum_{\gamma_2=\pm} R^{\alpha_1'}_{\alpha_1}(\gamma_1,\gamma_2) R^{\alpha_2'}_{\alpha_2}(\gamma_2,\gamma_2') = (R_1 \otimes R_2)^{\alpha_1'\alpha_2'}_{\alpha_1\alpha_2}(\gamma_1,\gamma_1') \tag{10.67}$$

where

$$R_n = \left(R^{\alpha_n'}_{\alpha_n}(\gamma_n,\gamma_n)\right) = \begin{pmatrix} R^+_+(\gamma_n,\gamma_n') & R^-_+(\gamma_n,\gamma_n') \\ R^+_-(\gamma_n,\gamma_n') & R^-_-(\gamma_n,\gamma_n') \end{pmatrix} \tag{10.68}$$

are the R-matrices of the two vertices $n = 1, 2$.

Note that the vertex weights w_{ij} entering the R-matrices R_1 and R_2 are the same for both R-matrices.

10.8 Further to the L-operator

Let us now return to the operator L_n. Its definition (10.43), together with (10.44) and (10.55), implies that L_n-operators for different sites commute

$$[L_n, L_m] = 0 \qquad \text{for} \qquad n \neq m. \tag{10.69}$$

Furthermore, the L_n-operators can be formally decomposed and written as a 2×2 matrix with elements $\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n$ and $\hat{\delta}_n$, which are themselves matrices or operators of dimension $2^N \times 2^N$

$$L_n = \begin{pmatrix} \hat{\alpha}_n & \hat{\beta}_n \\ \hat{\gamma}_n & \hat{\delta}_n \end{pmatrix}. \tag{10.70}$$

EXERCISE 10.4 L_n-matrix Convince yourself the the L_n-matrix (10.43) can indeed be written in the form (10.70).

The monodromy matrix may thus be expressed in terms of a matrix product of the the local L_n-operators

$$\mathcal{T} = L_1 L_2 \dots L_N = \prod_{n=1}^{N} L_n \tag{10.71}$$

or, because of (10.70)

$$\mathcal{T} = \prod_{n=1}^{N} \begin{pmatrix} \hat{\alpha}_n & \hat{\beta}_n \\ \hat{\gamma}_n & \hat{\delta}_n \end{pmatrix} \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{10.72}$$

Note that the matrix product (10.71) has been performed in the two-dimensional space of the matrices $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\delta}$ as explicitly expressed in (10.72).

Now, we can formally perform the trace of (10.38) to obtain

$$T = A + D. \tag{10.73}$$

So far, of course, this was all purely formal and we still have not diagonalized the transfer matrix T. This will only be possible once we specify the model, i.e. once we define the energies ϵ_j and therewith the Boltzmann weights v_j. However, we can now give a *sufficient* condition (not a *necessary* one, though!) for the transfer matrices to commute.

10.9 Yang–Baxter relations

For the commutation relations (10.27) to hold for arbitrary mutually different sets of vertex Boltzmann weights $\{v_j\}$ and $\{v'_j\}$ (or, alternatively, $\{w_{ij}\}$ and $\{w'_{ij}\}$), requires a large number of conditions to be satisfied, which appears to be a remote possibility.

Historically, the question of whether and when these conditions may be met has, of course, initially not been answered in full generality, but rather by considering many special examples, starting with the ice model. For an early overview of these developments, see Lieb and Wu (1972).

The sufficient condition we now formulate is thus to be considered as the fruit of combining many research efforts across the years. It builds on concepts developed thus far, in particular, the monodromy matrix $\mathcal{T}$ in the form (10.72).

A vanishing commutator of the tensor product of two monodromy matrices $\mathcal{T}$ and $\mathcal{T}'$ for two different sets of vertex weights would imply commutation of the corresponding

transfer matrices, i.e. $\left[T, T'\right] = 0$. However, the tensor product of $\mathcal{T}$ and $\mathcal{T}'$ will in general also not commute. What could be inferred from the examples, however, was that it may be possible to find a third set of vertex weights $\{v''_j\}$ or $\{w''_{ij}\}$ and a corresponding R-matrix such that

$$\tilde{R}''(\mathcal{T} \otimes \mathcal{T}') = (\mathcal{T}' \otimes \mathcal{T})\tilde{R}'' \tag{10.74}$$

where the tensor product, using (10.72), is given by

$$\mathcal{T} \otimes \mathcal{T}' = \begin{pmatrix} A\mathcal{T}' & B\mathcal{T}' \\ C\mathcal{T}' & D\mathcal{T}' \end{pmatrix}. \tag{10.75}$$

This means we are trying to find an R-matrix depending on a third set of vertex weights that interchanges the vertex weights on the two rows of vertices. These are the sufficient conditions for commuting transfer matrices, called the *Yang–Baxter relations*, that could be distilled from the vertex models for which an exact solution could be found. In our context, we will exploit their further algebraic consequences for vertex and quantum spin models, including the connection of a notion of quantum integrability and the exact solutions via the Bethe ansatz.

We emphasize that the R-matrix and the two monodromy matrices are to be taken for different sets of vertex weights v_j or w_{ij}, which is indicated by no prime, prime, and double prime, i.e.

$$\mathcal{T} = \mathcal{T}(v_j), \qquad \mathcal{T}' = \mathcal{T}(v'_j), \qquad \text{and} \qquad \tilde{R}'' = \tilde{R}(v''_j). \tag{10.76}$$

The R-matrix entering the Yang–Baxter relations (10.74) is obtained from the R-matrix discussed so far by multiplication with the permutation matrix (10.62),

$$\tilde{R} = \mathcal{P}R. \tag{10.77}$$

This modification of the R-matrix is needed to properly connect it to the two rows of vertices whose statistical weight is represented by the two monodromy matrices $\mathcal{T}$ and $\mathcal{T}'$. This reason for the modification (10.77) will, however, become especially clear when we discuss the graphical proof of the Yang–Baxter relations in section 10.10.1.

The Yang–Baxter relations (10.74) implies that the transfer matrices for different sets of vertex weights commute, i.e that (10.27) holds. In order to see this, we rewrite the Yang–Baxter relations (10.74) as

$$(\mathcal{T} \otimes \mathcal{T}') = \tilde{R}''^{-1}(\mathcal{T}' \otimes \mathcal{T})\tilde{R}'' \tag{10.78}$$

and take the appropriate trace in the four-dimensional space of matrices (10.75), recalling that matrices can be permuted cyclicly under the trace operation.

EXERCISE 10.5 Commutation relation of the transfer matrix Convince yourself that taking the trace in an appropriate four-dimensional space of the previous equation, (10.78), is sufficient for the transfer matrices for different sets of vertex weights to commute, i.e. for (10.27) to hold.

Section 10.10.1 presents a useful graphical way to prove this result, after we discuss the Yang–Baxter relations for the local L-operators and also for the R-matrices.

Using the local L-operators we can also write down local Yang–Baxter relations

$$\tilde{R}''(L_n \otimes L'_n) = (L'_n \otimes L_n)\tilde{R}'' \tag{10.79}$$

or, using the weights w_{ij} explicitly

$$\tilde{R}(w''_{ij})\left(L_n(w_{ij}) \otimes L_n(w'_{ij})\right) = \left(L_n(w'_{ij}) \otimes L_n(w_{ij})\right)\tilde{R}(w''_{ij}). \tag{10.80}$$

A corresponding Yang–Baxter relation can be given involving only R-matrices of different sets of vertex weights

$$\tilde{R}(w''_{ij})\left(R(w_{ij}) \otimes R(w'_{ij})\right) = \left(R(w'_{ij}) \otimes R(w_{ij})\right)\tilde{R}(w''_{ij}), \tag{10.81}$$

which we shall discuss in more detail in section 10.10. The tensor products in this expression are formed when the R-matrices are interpreted as built up from 2×2 matrices as in the convention given by (10.12). Then, the remaining product is a product of 4×4 matrices where one factor acts as an ordinary 4×4 matrix, and the other as a 4×4 matrix where each element is a product of two 2×2 matrices. Note that only one R-matrix on each side of (10.81), the one acting as a 4×4 matrix, is modified by the permutation matrix $\mathcal{P}$. We shall encounter this type of Yang–Baxter relation in section 11.1 when we use it to solve the six-vertex model explicitly.

We pause for a moment to clarify what our results so far mean. We found the Yang–Baxter relations as a sufficient condition for the integrability of the transfer matrix of the vertex model. Integrability here means that we can diagonalize the transfer matrix of a given vertex model and hence, calculate the partition function.

10.10 More on Yang–Baxter relations

The Yang–Baxter relations appear in many different versions. For an overview of the various forms and their physical and mathematical backgrounds, see Perk and Au-Yang (2006). We have already encountered two versions, the first (10.74) involving, besides the R-matrix, the whole monodromy matrix $\mathcal{T}$ expressing the statistical weight of a row of vertices. In the second version (10.80), we have replaced the monodromy matrix $\mathcal{T}$ by

the *local* L_n matrices, which suggests that there is a version of the Yang–Baxter relations that employs only R-matrices. We have encountered this version already in (10.81) but now discuss it in more detail, mostly in the form of an exercise.

While the versions of the Yang–Baxter relations involving the monodromy matrix $\mathcal{T}$ or the L_n matrices are connected to the square lattice of the vertex model, the version involving only R-matrices consists of just three vertices attached to each other to form (the topology of) a triangle (cf. figure 10.5). This combination of vertices does, of course, not occur normally in a square lattice. We get from the left-hand side of figure 10.5 to the right-hand side, and vice versa, by a parallel shift of one of the sides of the triangle through the corner opposite that side. It is important to note that this shift implies that the vertex at the bottom of the right-hand side of the figure becomes the vertex at the top of the left-hand side, and vice versa, albeit with different spin indices on the bonds.

EXERCISE 10.6 Yang–Baxter relations for the R-matrices Convince yourself that the Yang–Baxter relations corresponding to figure 10.5 is explicitly given by

$$\sum_{\alpha''}\sum_{\beta''}\sum_{\gamma''} R^{\alpha'}_{\alpha''}(\gamma,\gamma'')(w)R^{\beta'}_{\beta''}(\gamma'',\gamma')(w')R^{\alpha''}_{\alpha}(\beta,\beta'')(w'') =$$
$$\sum_{\alpha''}\sum_{\beta''}\sum_{\gamma''} R^{\alpha'}_{\alpha''}(\beta'',\beta')(w'')R^{\beta''}_{\beta}(\gamma,\gamma'')(w')R^{\alpha''}_{\alpha}(\gamma'',\gamma')(w). \qquad (10.82)$$

We have suppressed the indices ij of the vertex weights w, w', and w''. These equations represent $2^6 = 64$ linear equations between the matrix elements of the R-matrices.

Note that, since we are dealing with matrix *elements*, their sequence in the expression above does not matter but is conventionally written in the order we use. In many other forms of the Yang–Baxter relation we deal with matrices or operators where the corresponding order *does* matter.

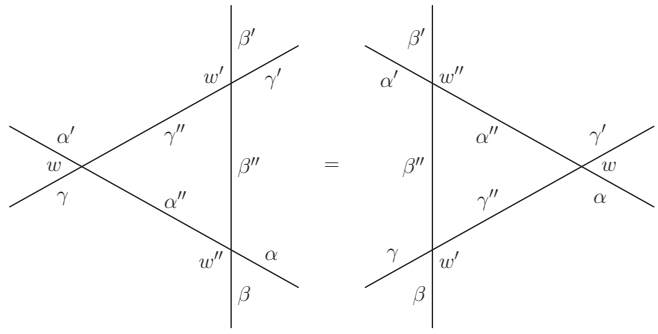

Figure 10.5 *Yang–Baxter relations for the R-matrices. The vertices are labelled by the weights w, w', and w'' (suppressing the indices ij) and four surrounding spin indices, cf. (10.39) or (10.84). Cf. section 10.10 and exercise 10.6.*

EXERCISE 10.7 Equivalence of Yang–Baxter relations Convince yourself that the Yang–Baxter relations in the forms (10.82) and (10.81) are indeed equivalent.

The local forms of the Yang–Baxter relations (10.82) and (10.81) for the R-matrices and their graphical representation in figure 10.5 are useful for a graphical demonstration of the commutation relation of the transfer matrix.

10.10.1 Graphical proof of the Yang–Baxter relation

First, we require a further relation to hold between the R-matrices. We assume that there are two sets of vertex weights w'''_{ij} and w''_{ij} such that the corresponding R-matrices satisfy

$$\sum_{\upsilon\nu} R^{\tau}_{\upsilon}(\sigma,\nu)(w'''_{ij})R^{\upsilon}_{\beta}(\upsilon,\alpha)(w''_{ij}) = \delta_{\alpha\tau}\delta_{\beta\sigma}. \tag{10.83}$$

In other words, the two R-matrices in (10.83) are matrix inverses of each other.

EXERCISE 10.8 Inverse of the R-matrix Represent (10.83) graphically in a way similar to, e.g. figure 10.5.

Using (10.82) and (10.83) the graphical proof of the commutativity of the transfer matrix can be carried out; see figure 10.6 for the details.

10.11 Exploiting Yang–Baxter integrability

We now exploit the local Yang–Baxter relations (10.80), i.e. the integrability condition. In order to do so, let us finally establish an explicit relation between the vertex weights v_j and the elements of the R-matrix (cf. 10.39):

$$R^{\alpha'}_{\alpha}(\gamma,\gamma')(w_{ij}) = \sum_{i,j=1}^{4} w_{ij}\sigma^{i}_{\gamma,\gamma'}\sigma^{j}_{\alpha,\alpha'}. \tag{10.84}$$

The logic is the following: we are searching for vertex weights w_{ij} such that the local Yang–Baxter relations are satisfied.

So far, we have not imposed any restrictions on the model. Thus the R-matrix (10.84) still corresponds to the most general case of a sixteen-vertex model. Unfortunately, there are no vertex weights known for a sixteen-vertex model such that the corresponding R-matrix satisfies the local Yang–Baxter relations (10.80) or (10.81).

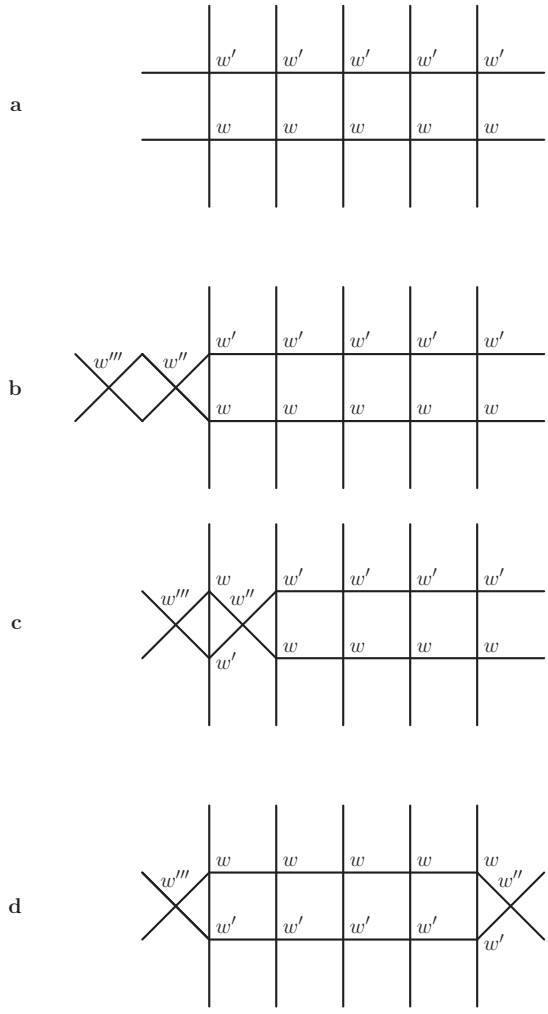

Figure 10.6 *Graphical proof of the Yang–Baxter relation (10.74). The figure in panel* ***a*** *represents the product of two transfer matrices as in (10.22) but with two different weights w and w′ (the subscripts ij are again suppressed). All inner spin or arrow variables are identified and summed over. The rightmost and leftmost spin variables are also identified and summed over, i.e. we assume periodic boundary conditions in the horizontal direction. For the figure of panel* ***b****, the matrix product (10.83) representing a unit matrix has been inserted at the left end, which does not cause any change. These weights w, w′, and w″ are assumed to satisfy the Yang–Baxter relations. The Yang–Baxter relations in the form for the R-matrices is now used to move the vertex with weight w″ to the right (panel* ***c****). Thereby, the upper vertex with weight w′ and the lower vertex with weight w are exchanged. We can keep moving the vertex with weight w″ to the right thereby exchanging the upper and lower vertices until the end of the two rows is reached (panel* ***d****). Due to the periodic boundary conditions, the vertices at the left and right ends (panel* ***d****) can now be recombined to give a unit matrix. The net effect of this procedure is that the transfer matrices represented by the two rows of vertices are commuted. This constitutes a graphical proof the Yang–Baxter relations (10.74).*

We shall accordingly restrict our attention to the eight-vertex model and eventually even the six-vertex model where (10.84) simplifies to

$$R_{\alpha}^{\alpha'}(\gamma,\gamma')(w_j)=\sum_{j=1}^{4} w_j \sigma^j_{\gamma,\gamma'}\sigma^j_{\alpha,\alpha'} \tag{10.85}$$

and the L-operators are

$$L_n(w_j)=\sum_{j=1}^{4} w_j \sigma^j \otimes \sigma_n^j = \begin{pmatrix} w_4 I_n + w_3\sigma_n^z & w_1\sigma_n^x - iw_2\sigma_n^y \\ w_1\sigma_n^x + iw_2\sigma_n^y & w_4 I_n - w_3\sigma_n^z \end{pmatrix} = \begin{pmatrix} \hat{\alpha}_n & \hat{\beta}_n \\ \hat{\gamma}_n & \hat{\delta}_n \end{pmatrix}. \tag{10.86}$$

The R-matrix elements are given by (10.13) through (10.16) or, using (10.85), in terms of the weights w_j as

$$\mathfrak{a}=w_3+w_4, \qquad \mathfrak{b}=w_4-w_3, \qquad \mathfrak{c}=w_1+w_2, \qquad \text{and} \qquad \mathfrak{d}=w_1-w_2. \tag{10.87}$$

For the six-vertex model, the ice rule requires

$$\mathfrak{d}\equiv v_7=v_8=0 \qquad \Leftrightarrow \qquad w_1=w_2. \tag{10.88}$$

The local Yang–Baxter relations (10.80) relate the different sets of vertex weights w_j'', w_j', and w_j. For the eight–vertex model, the Yang–Baxter relations describe an elliptic curve that can be parametrized by elliptic functions $\mathrm{sn}(u,k)$ (elliptic sine of modulus k). These remarks connect the present discussion with the remarks in section 10.6, especially figure 10.4.

Here, we finally focus on the six-vertex model (see chapter 11), and refer the interested reader to the literature (see Baxter (1982)) for the further detailed (and technically much more elaborate) treatment of the eight-vertex model.

11

Six-Vertex Model

Algebra is generous; she often gives more than is asked of her.

Jean-Baptiste le Rond d'Alembert (1717–1783)

We restrict our attention in this chapter to the six-vertex model, where the curves obtained from the local Yang–Baxter relations (10.80) or (10.81) and its equivalent relation (10.82) can be parametrized by trigonometric sine functions[1] instead of elliptic functions, as in the eight-vertex model. The trigonometric functions would also emerge from appropriate limits (i.e. $k \to 1$) of the elliptic functions.

In order to achieve this, we will exploit the Yang–Baxter relations for the example of the six-vertex model (see section 11.1) in the trigonometric parameterization (see section 11.2) and find Lieb's parameter Δ, which encapsulates integrability of the model, i.e. the commutation of the transfer matrices for different sets of vertex weights. We shall list the parameterizations and use these to interpret the six-vertex model as a model for two-dimensional ferroelectrics.

Section 11.3, given its parameterization, explores diagonalization of the transfer matrix of the six-vertex model, i.e. finding the Bethe ansatz equations as equations a set of parameters has to satisfy as conditions for this diagonalization to be valid. The algebraic procedure with which the diagonalization of the transfer matrix is achieved is called the algebraic Bethe ansatz.

Section 11.4 makes the correspondence explicit between a two-dimensional model of classical statistical mechanics and a corresponding one-dimensional quantum model, starting with the two-dimensional classical six-vertex model, and ending up with the one-dimensional anisotropic Heisenberg quantum spin chain, having established, along the way, integrability, and again in particular the role of the Yang–Baxter relations as the heart of integrability, for these models. From the transfer matrix, we can construct an infinite set of commuting operators; in particular, we find that the transfer matrix of the six-vertex model and the Hamiltonian of the quantum spin chain commute. Moreover, given the algebraic diagonalization of the transfer matrix, we have automatically also

[1] As we will see later, this is only one possible parameterization that corresponds to the disordered phase of the six-vertex model in 11.2.1, or to the anisotropic XXZ Heisenberg quantum spin chain. The Lieb parameter is $|\Delta| \leq 1$ and determines the position in the phase diagram in the six-vertex model while it describes the anisotropy of the quantum spin chain.

Models of Quantum Matter: A First Course on Integrability and the Bethe Ansatz. Hans-Peter Eckle, Oxford University Press (2019). © Hans-Peter Eckle. DOI: 10.1093/oso/9780199678839.001.0001

obtained the exact solution of the one-dimensional quantum spin chain Hamiltonian given by the same Bethe ansatz equations.

This is the constructive algebraic programme of the quantum inverse scattering method (QISM) and the *algebraic* Bethe ansatz. In the words of Evgeny Sklyanin (1992), 'The basic idea of QISM [...] is purely algebraic.'

Finally, section (11.5) demonstrates how the Yang–Baxter relations, which are still valid in spatially inhomogeneous situations, can be used to examine new classes of quantum integrable spin chains.

11.1 Yang–Baxter relation for the six-vertex model

Let us recall the R-matrix (10.17) for the six-vertex model for which the ice rule must be satisfied, which implies that we have to set $\mathfrak{d} = 0$

$$R_{6\mathrm{v}} = \left(\begin{array}{cc|cc} \mathfrak{a} & 0 & 0 & 0 \\ 0 & \mathfrak{b} & \mathfrak{c} & 0 \\ \hline 0 & \mathfrak{c} & \mathfrak{b} & 0 \\ 0 & 0 & 0 & \mathfrak{a} \end{array}\right) \equiv \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{11.1}$$

EXERCISE 11.1 Two vertices Show that for a system consisting of only $N = 2$ vertices, the matrices $\hat{\alpha}_n$, $\hat{\beta}_n$, $\hat{\gamma}_n$, and $\hat{\delta}_n$ of the L_n-matrices of (10.70) with $n = 1, 2$ are related to the 2×2 matrices α, β, γ, and δ of the R-matrix (11.1) by

$$\hat{\alpha}_1 = \alpha \otimes I_2, \qquad \hat{\alpha}_2 = I_2 \otimes \alpha, \qquad \hat{\beta}_1 = \beta \otimes I_2, \qquad \hat{\beta}_2 = I_2 \otimes \beta, \tag{11.2}$$

$$\hat{\gamma}_1 = \gamma \otimes I_2, \qquad \hat{\gamma}_2 = I_2 \otimes \gamma, \qquad \hat{\delta}_1 = \delta \otimes I_2, \qquad \hat{\delta}_2 = I_2 \otimes \delta. \tag{11.3}$$

Show that for this system consisting of only two vertices the transfer matrix becomes

$$T_2 = \hat{\alpha}_1\hat{\alpha}_2 + \hat{\beta}_1\hat{\gamma}_2 + \hat{\gamma}_1\hat{\beta}_2 + \hat{\delta}_1\hat{\delta}_2 \tag{11.4}$$

in two ways:

- graphically by writing down the two vertices and summing over all horizontal spin or arrow variables; and
- by multiplying the two L_n-matrices and performing the trace.

Let us now evaluate the Yang–Baxter relations (10.81) in the $\mathfrak{abc}$-notation

$$\begin{pmatrix} \mathfrak{a}'' & 0 & 0 & 0 \\ 0 & \mathfrak{c}'' & \mathfrak{b}'' & 0 \\ 0 & \mathfrak{b}'' & \mathfrak{c}'' & 0 \\ 0 & 0 & 0 & \mathfrak{a}'' \end{pmatrix} \begin{pmatrix} \alpha\alpha' & \alpha\beta' & \beta\alpha' & \beta\beta' \\ \alpha\gamma' & \alpha\delta' & \beta\gamma' & \beta\delta' \\ \gamma\alpha' & \gamma\beta' & \delta\alpha' & \delta\beta' \\ \gamma\gamma' & \gamma\delta' & \delta\gamma' & \delta\delta' \end{pmatrix} = \begin{pmatrix} \alpha'\alpha & \alpha'\beta & \beta'\alpha & \beta'\beta \\ \alpha'\gamma & \alpha'\delta & \beta'\gamma & \beta'\delta \\ \gamma'\alpha & \gamma'\beta & \delta'\alpha & \delta'\beta \\ \gamma'\gamma & \gamma'\delta & \delta'\gamma & \delta'\delta \end{pmatrix} \begin{pmatrix} \mathfrak{a}'' & 0 & 0 & 0 \\ 0 & \mathfrak{c}'' & \mathfrak{b}'' & 0 \\ 0 & \mathfrak{b}'' & \mathfrak{c}'' & 0 \\ 0 & 0 & 0 & \mathfrak{a}'' \end{pmatrix}. \tag{11.5}$$

As in exercise 10.6, these are $16 \cdot 4 = 64$ linear equations for the three sets of three weight variables each: $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c})$, $(\mathfrak{a}', \mathfrak{b}', \mathfrak{c}')$, and $(\mathfrak{a}'', \mathfrak{b}'', \mathfrak{c}'')$. A great number of these equations are, however, trivially satisfied. Of the remaining non-trivial equations, many turn out to be mutually identical. In the end, only three non-trivial independent equations survive.

EXERCISE 11.2 Linear equations of the six-vertex model Derive the three surviving linear equations of the six-vertex model. This can be done purely algebraically using (11.5) together with (11.1). The reduction in the number of independent equations can also be achieved by invoking the ice rule and by further symmetry arguments (see Baxter, 1982, section 9.6).

The remaining three linear equations are

$$\mathfrak{a}''\mathfrak{c}'\mathfrak{b} = \mathfrak{c}''\mathfrak{b}'\mathfrak{c} + \mathfrak{b}''\mathfrak{c}'\mathfrak{a}, \tag{11.6}$$

$$\mathfrak{a}''\mathfrak{a}'\mathfrak{c} = \mathfrak{c}''\mathfrak{c}'\mathfrak{a} + \mathfrak{b}''\mathfrak{b}'\mathfrak{c}, \tag{11.7}$$

$$\mathfrak{c}''\mathfrak{a}'\mathfrak{b} = \mathfrak{c}''\mathfrak{b}'\mathfrak{a} + \mathfrak{b}''\mathfrak{c}'\mathfrak{c}. \tag{11.8}$$

For given, but different, values of the weights $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c})$ and $(\mathfrak{a}', \mathfrak{b}', \mathfrak{c}')$, we can view this set of equations as a homogeneous set of linear algebraic equations for the weights $(\mathfrak{a}'', \mathfrak{b}'', \mathfrak{c}'')$. The condition for the existence of such a solution is that the determinant of this set of equations vanishes, which is the case if $(\mathfrak{a}, \mathfrak{b}, \mathfrak{c})$ and $(\mathfrak{a}', \mathfrak{b}', \mathfrak{c}')$ satisfy

$$\frac{\mathfrak{a}^2 + \mathfrak{b}^2 - \mathfrak{c}^2}{2\mathfrak{a}\mathfrak{b}} = \frac{\mathfrak{a}'^2 + \mathfrak{b}'^2 - \mathfrak{c}'^2}{2\mathfrak{a}'\mathfrak{b}'} \equiv \Delta \tag{11.9}$$

where we have introduced a factor of 1/2 for later convenience. The quantity Δ is often called Lieb's parameter in honour of Elliott Lieb's seminal work on vertex models.

11.2 Parameterization of the six-vertex model

The condition (11.9), an algebraic curve, and, thus, the local Yang–Baxter relations (10.80) of the six-vertex model are satisfied, for instance, for the trigonometric parameterization

$$\mathfrak{a} = w_3 + w_4 = \sin(u + \eta), \tag{11.10}$$

$$\mathfrak{b} = w_4 - w_3 = \sin(u - \eta), \tag{11.11}$$

$$\mathfrak{c} = 2w_1 = \sin(2\eta), \tag{11.12}$$

i.e. we have two parameters, u and η, whose significance will become clear in the following.

In fact, we shall regard u, the so-called spectral parameter, as the main parameter, which will play a major role in diagonalizing the transfer matrix. The parameter η will turn out to describe different regimes of the model but can also be viewed as related to an interaction strength or an anisotropy parameter.

Other parameterizations are also possible, e.g. using hyperbolic functions

$$\mathfrak{a} = w_3 + w_4 = \sinh(u + \eta), \tag{11.13}$$

$$\mathfrak{b} = w_4 - w_3 = \sinh(\pm(u - \eta)), \tag{11.14}$$

$$\mathfrak{c} = 2w_1 = \sinh(2\eta). \tag{11.15}$$

For the trigonometric and the hyperbolic parameterizations, respectively, we obtain for Lieb's parameter Δ

$$\Delta = \cos(2\eta) \qquad \text{and} \qquad \Delta = \pm\cosh(2\eta). \tag{11.16}$$

Later we shall also encounter, in the limit $\eta \to 0$, a so-called rational parameterization, i.e. one in terms of rational functions of the parameters. After a short section on the interpretation of the six-vertex model to describe two-dimensional ferroelectrics, we thereafter concentrate on the trigonometric parameterization.

11.2.1 Two-dimensional ferroelectric models: phase diagram of the six-vertex model

The two-dimensional vertex models have a natural interpretation as models for two-dimensional ferroelectrics when the arrows on the bonds between vertices are interpreted now as electrical dipoles. In the two-dimensional lattice these electric dipoles can be ferroelectically or antiferroelectrically ordered, or they can be disordered according to the energies or, equivalently, vertex weights of the vertices.

The following gives a qualitative picture of the possible phases of the six-vertex model (cp. figure 11.1). Remember that we are always assuming *periodic* boundary conditions in both directions of the lattice, i.e. *toroidal* boundary conditions. More details going beyond a discussion of the possible phases of the six-vertex model can be found, in Lieb and Wu (1972) and in chapters 8 and 9 of Baxter (1982).

The phases of the six-vertex model depend on the three intervals of Lieb's parameter: $\Delta < -1$, $-1 < \Delta < 1$, and $\Delta > 1$. The free energy of the model has a different analytical form for each of these three intervals (see chapter 8 of Baxter, 1982, for the explicit calculation of the respective free energies). In our qualitative discussion, we distinguish the different phases by considering directly the vertices, their energies, and corresponding weights as shown in the first six panels of figure 10.1.

For all non-negative (and real) energies of the six vertices, the weights are restricted to the interval $0 < \mathfrak{a}, \mathfrak{b}, \mathfrak{c} < 1$ or, according to the discussion in section 10.6, to the interval $0 < \mathfrak{a}/\mathfrak{c}, \mathfrak{b}/\mathfrak{c} < 1$ of these two ratios.

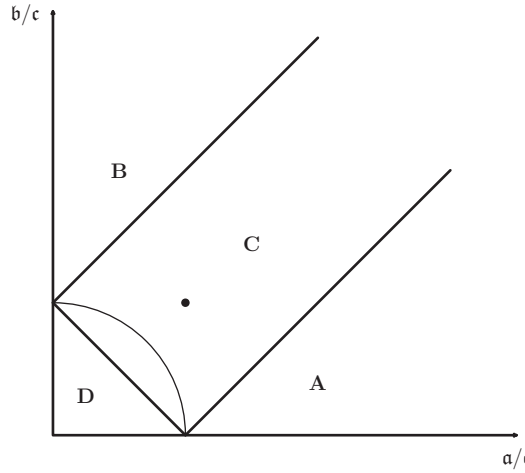

Figure 11.1 *Phase diagram of the six-vertex model. The dot represents the point in the phase diagram where the temperature becomes infinite:* $\mathsf{T} = \infty$, *i.e.* $\mathfrak{a} = \mathfrak{b} = \mathfrak{c} = 1$. *The model is equivalent to a model of free Fermions for* $\Delta = 0$, *which corresponds to the quarter circle of the thin line, i.e.* $(\mathfrak{a}/\mathfrak{c})^2 + (\mathfrak{b}/\mathfrak{c})^2 = 1$.

Lieb's parameter is symmetric under the exchange of $\mathfrak{a} \leftrightarrow \mathfrak{b}$ so that the phase diagram will be symmetric with respect to the line $\mathfrak{a}/\mathfrak{c} = \mathfrak{b}/\mathfrak{c}$.

The phases described in the following can be discerned in figure 11.1.

A,B The *ferroelectric* phase is separated into two parts, depending on whether the ratios of the vertex weights satisfy either $\mathfrak{a}/\mathfrak{c} > \mathfrak{b}/\mathfrak{c} + 1$ or $\mathfrak{b}/\mathfrak{c} > \mathfrak{a}/\mathfrak{c} + 1$. This is the case if either $\epsilon_{\mathfrak{a}} < \epsilon_{\mathfrak{b}}$ and $\epsilon_{\mathfrak{a}} < \epsilon_{\mathfrak{c}}$ or $\epsilon_{\mathfrak{b}} < \epsilon_{\mathfrak{a}}$ and $\epsilon_{\mathfrak{b}} < \epsilon_{\mathfrak{c}}$ and $\Delta > 1$.

At very low temperatures, two ground states are possible. All vertices are either of type $\mathfrak{a}$ or $\mathfrak{b}$. If the vertices are of type $\mathfrak{a}$, all dipoles point either up or to the right if the vertices are of subtype 1, or they point down or to the left if the vertices are of subtype 2. For the subtypes 3 and 4 of vertex type $\mathfrak{b}$, a similar statement holds.

We note in particular that there is a net polarization and thus the system exhibits a ferroelectric order. A detailed analysis reveals that excitations contribute only negligibly to the free energy, which is equal to $\epsilon_{\mathfrak{a}}$ or $\epsilon_{\mathfrak{b}}$, respectively, and therefore the system remains frozen in one of the ground states (see Baxter, 1982, chapter 8).

D The antiferroelectric phase is characterized by $\mathfrak{a}/\mathfrak{c} + \mathfrak{b}/\mathfrak{c} < 1$, $\epsilon_{\mathfrak{c}} < \epsilon_{\mathfrak{a}}$ and $\epsilon_{\mathfrak{c}} < \epsilon_{\mathfrak{b}}$ such that $\Delta < -1$. The lowest energy is now obtained if there are only type $\mathfrak{c}$ vertices where all neighbouring dipoles are antiparallel in either the vertical or horizontal direction. Hence, there are again two ground states that now exhibit antiferroelectric order.

C Arguably, the richest phase in the phase diagram is the one bounded by the three lines $\mathfrak{a}/\mathfrak{c} = \mathfrak{b}/\mathfrak{c} + 1$, $\mathfrak{b}/\mathfrak{c} = \mathfrak{a}/\mathfrak{c} + 1$, and $\mathfrak{a}/\mathfrak{c} + \mathfrak{b}/\mathfrak{c} = 1$. This region can be

circumscribed in the symmetrical form $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} < \frac{1}{2}(\mathfrak{a} + \mathfrak{b} + \mathfrak{c})$. In this region, Lieb's parameter is in the interval $-1 < \Delta < 1$.

In particular, the point $\mathfrak{a} = \mathfrak{b} = \mathfrak{c} = 1$ lies in this region. It corresponds to infinite temperature $T \to \infty$ where we thus expect a disordered phase. In fact, from the analysis of the free energy (see Baxter, 1982), it can be inferred that this whole region is a disordered phase.

Furthermore, the quarter circle $(\mathfrak{a}/\mathfrak{c})^2 + (\mathfrak{b}/\mathfrak{c})^2 = 1$ describes the locus of all points where $\Delta = 0$. In the detailed Bethe ansatz analysis, which can be found e.g. in Baxter (1982), this case is particularly simple and corresponds to a non-interacting spin chain or, equivalently, a chain of free Fermions.

Moreover, the Bethe ansatz solution of the eight-vertex model reveals that the entire region **C** corresponds to the critical eight-vertex model, i.e. the six-vertex model is critical in a whole region of the phase diagram, rather than just at a point or on a line.

We note that the region of $-1 < \Delta < 0$ is bounded by the quarter circle $(\mathfrak{a}/\mathfrak{c})^2 + (\mathfrak{b}/\mathfrak{c})^2 = 1$ from above and the line $\mathfrak{a}/\mathfrak{c} + \mathfrak{b}/\mathfrak{c} = 1$, whereas the region of $0 < \Delta < 1$ is above the quarter circle in the corridor limited by the lines $\mathfrak{b}/\mathfrak{c} = \mathfrak{a}/\mathfrak{c} \pm 1$.

Lastly, in our qualitative analysis, we remark that the point $\mathfrak{a}/\mathfrak{c} = \mathfrak{b}/\mathfrak{c} = 1$ is reached from all other points of the phase diagram as the temperature T increases from 0 to ∞. If the path traced out in this process starts in one of the regions **A**, **B**, or **D**, it will have to cross a phase boundary to eventually reach the point $\mathfrak{a}/\mathfrak{c} = \mathfrak{b}/\mathfrak{c} = 1$ in region **C**, i.e. one of the lines $\mathfrak{b}/\mathfrak{c} = \mathfrak{a}/\mathfrak{c} \pm 1$ and $\mathfrak{b}/\mathfrak{c} = -\mathfrak{a}/\mathfrak{c} + 1$. Again, a more detailed analysis, e.g. Baxter (1982), reveals that these transitions into the critical region possess only one critical transition temperature T_c where the free energy is singular and the correlation length is infinite.

For those statements in the classification of the phases of the six-vertex model that go beyond a qualitative description based on the vertex energies and weights, we again refer to Baxter (1982) and the references cited therein.

11.3 Algebraic Bethe ansatz solution of the six-vertex model

We now return to the Yang–Baxter relations and exploit them finally to solve a vertex model, the six-vertex model. In concrete terms, this means that we shall determine a set of parameters and derive the Bethe ansatz equations that determine these parameters.

The replacements in the parameterizations (11.10–11.12)

$$u \to v \qquad \text{and} \qquad u \to u - v + \eta \tag{11.17}$$

give sets w_j' and w_j'', respectively, such that the local Yang–Baxter relations (10.80) are satisfied. The local Yang–Baxter relations then take the form

$$\tilde{R}(u-v)(L_n(u)\otimes L_n(v)) = (L_n(v)\otimes L_n(u))\tilde{R}(u-v). \tag{11.18}$$

The L_n- and R-matrices can now be given explicitly as

$$L_n(u) = \begin{pmatrix} w_4(u)I_n + w_3(u)\sigma_n^z & \frac{\sin 2\eta}{2}\sigma_n^- \\ \frac{\sin 2\eta}{2}\sigma_n^+ & w_4(u)I_n - w_3(u)\sigma_n^z \end{pmatrix} \tag{11.19}$$

where

$$w_4(u) \pm w_3(u) = \sin(u \pm \eta) \tag{11.20}$$

and

$$\tilde{R}(u) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathfrak{c}(u) & \mathfrak{b}(u) & 0 \\ 0 & \mathfrak{b}(u) & \mathfrak{c}(u) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{11.21}$$

with[2]

$$b(u) = \frac{\sin 2\eta}{\sin(u+2\eta)} \qquad \text{and} \qquad c(u) = \frac{\sin u}{\sin(u+2\eta)}, \tag{11.22}$$

i.e. we have normalized the R-matrix by dividing by $\sin(u + 2\eta)$. Note that the ratio $b(u)/c(u)$ is an odd function of the spectral parameter u. This fact will be of some importance further on.

Now, we are in a position to start the diagonalization of the transfer matrix. The first step is to introduce the local vacuum state at site n of the row of vertices corresponding to the local L_n-matrix

$$\omega_n = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n, \tag{11.23}$$

i.e. the state defined as the eigenstate of the Pauli matrix σ_n^z with eigenvalue $+1$, i.e.

$$\sigma_n^z\omega_n = \sigma_n^z \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = + \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = +\omega_n. \tag{11.24}$$

[2] This change of notation, $\mathfrak{c} \to b$ and $\mathfrak{b} \to c$, will facilitate comparison with R-matrices of other models in later chapters and the conventional notation used there.

This state is annihilated by σ_n^+ while σ_n^- creates a new local state

$$\sigma_n^+\omega_n = \sigma_n^+ \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = 0 \qquad \text{and} \qquad \sigma_n^-\omega_n = \sigma_n^- \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_n, \tag{11.25}$$

the latter being an eigenstate of σ_n^z with eigenvalue (-1)

$$\sigma_n^z \begin{pmatrix} 0 \\ 1 \end{pmatrix}_n = - \begin{pmatrix} 0 \\ 1 \end{pmatrix}_n. \tag{11.26}$$

The L-operator accordingly acts on ω_n to produce

$$L_n(u)\omega_n = \begin{pmatrix} \alpha(u)\omega_n & * \\ 0 & \delta(u)\omega_n \end{pmatrix}. \tag{11.27}$$

which is an upper triangular matrix, where

$$\alpha(u) = \sin(u+\eta), \qquad \text{and} \qquad \delta(u) = \sin(u-\eta). \tag{11.28}$$

Because this is a triangular matrix, the matrix element indicated by an asterisk, *, will not play any further role in the analysis and we thus do not bother to write it down explicitly.

The reference state

$$\Omega = \prod_{n=1}^{N} \otimes \omega_n \equiv |0\rangle, \tag{11.29}$$

a global vacuum state, assumes the role of the local vacuum state for the monodromy matrix. Using the triangular form of (11.27), we immediately obtain for the monodromy matrix (10.71)

$$\mathcal{T}(u)|0\rangle = \prod_{n=1}^{N} L_n(u)|0\rangle = \begin{pmatrix} \alpha^N(u)|0\rangle & * \\ 0 & \delta^N(u)|0\rangle \end{pmatrix}. \tag{11.30}$$

Now, the formal developments introduced via a number of formal quantities are starting to pay off. The transfer matrix was

$$T(u) = \mathrm{Tr}\,(\mathcal{T}(u)) = A(u) + D(u). \tag{11.31}$$

We can now immediately read off an eigenvalue of the transfer matrix and also the eigenvalues of the matrices A and D

$$A(u)\Omega = \alpha^N(u)|0\rangle \qquad \text{and} \qquad D(u)\Omega = \delta^N(u)|0\rangle \tag{11.32}$$

and, finally, an eigenvalue of the transfer matrix itself

$$T(u)|0\rangle = \left(\alpha^N(u) + \delta^N(u)\right)|0\rangle. \tag{11.33}$$

Thus, we can note that the vacuum state $|0\rangle$ is an eigenstate of the transfer matrix with eigenvalue $\alpha^N(u) + \delta^N(u)$.

We now want to construct other eigenstates of the transfer matrix $T(u)$. The key ingredient in this construction is the fundamental Yang–Baxter relation using the R-matrix (11.21)

$$\tilde{R}(u-v)(\mathcal{T}(u) \otimes \mathcal{T}(v)) = (\mathcal{T}(v) \otimes \mathcal{T}(u))\tilde{R}(u-v). \tag{11.34}$$

It is this constructive algebraic approach that gives the whole approach its name: algebraic Bethe ansatz.

Writing the Yang–Baxter relations in components, we obtain 16 commutation relations between the operators A, B, C, and D, often called the fundamental commutation relations. Fortunately, many of them are rather trivial or redundant. In fact, we only need three of the 16 relations for our purpose. The operator B commutes for different values of the spectral parameter

$$[B(u), B(v)] = 0. \tag{11.35}$$

The other two commutation relations are more complicated and involve the special feature of two terms on the right-hand side,

$$A(u)B(v) = \frac{1}{c(v-u)}B(v)A(u) - \frac{b(v-u)}{c(v-u)}B(u)A(v) \tag{11.36}$$

and

$$D(u)B(v) = \frac{1}{c(u-v)}B(v)D(u) - \frac{b(u-v)}{c(u-v)}B(u)D(v). \tag{11.37}$$

They contain on the right-hand side a first term, where the arguments appearing in the operators on the left-hand side are preserved, and a second term where these arguments are exchanged. The two terms with the exchanged arguments in the second term are the major complication we encounter in the further analysis. We note that the operator $C(u)$ does not play a further role in our analysis. It can be shown that $C(u)$ acts as an annihilation operator in the same sense as $B(u)$ acts as a creation operator.

EXERCISE 11.3 Explicitly derive (11.35–11.37) from the Yang–Baxter relation (11.34) of the six-vertex model. You may want to derive further commutation relations.

The commutation relations (11.36) and (11.37) suggest an interpretation of $B(u)$ as a particular kind of creation operator. Acting with B-operators for M distinct values u_j

($j = 1, \ldots, M$) of the spectral parameter on the reference state $|0\rangle$, we may construct the product state

$$|u_1, u_2, \ldots, u_M\rangle \equiv |\{u_j\}\rangle = B(u_1)B(u_2)\ldots B(u_M)|0\rangle = \prod_{j=1}^{M} B(u_j)|0\rangle. \tag{11.38}$$

It is important to note that, due to their commutativity 11.35, we can permute the $B(u_j)$ operators in the product of this state arbitrarily.

A further observation that will play a key role in the following is that acting with the transfer matrix $T(u)$ (or, with the operators $A(u)$ and $D(u)$ separately) on the product state (11.38), only the $1 + M$ states

$$\prod_{j=1}^{M} B(u_j)|0\rangle \tag{11.39}$$

$$\prod_{j=1, j\neq n}^{M} B(u_j)|0\rangle, \qquad n = 1, \ldots, M \tag{11.40}$$

can be generated after commuting $T(u)$ through all operators $B(u_j)$ so that $T(u)$, acting on the reference state $|0\rangle$, creates the vacuum eigenvalue $\alpha^N(u) + \delta^N(u)$.

We expect that the state (11.38) will be an eigenstate of the transfer matrix. We shall see, however, that this is not immediately the case. Rather, we find that the commutation relations (11.36) and (11.37), due to their two terms with the exchanged arguments in the second term, create terms that are not proportional to the state $|\{u_j\}\rangle$. In fact, we obtain a large number of further terms

$$T(u)|\{u_j\}\rangle = (A(u) + D(u))\,|\{u_j\}\rangle = \Lambda(u)|\{u_j\}\rangle + \text{further terms}. \tag{11.41}$$

Provided we can make these further terms, called unwanted or sometimes even 'junk' terms, disappear, this equation would, however, indeed become an eigenvalue equation.

We shall find that the unwanted terms indeed do disappear, provided the distinct values of the spectral parameter $\{u_j\}$ satisfy

$$\frac{\alpha^N(u_j)}{\delta^N(u_j)} = \prod_{k=1, k\neq j}^{M} \frac{c(u_k - u_j)}{c(u_j - u_k)} \qquad j = 1, \ldots, M. \tag{11.42}$$

If this provision can be satisfied, we shall find the eigenvalue

$$\Lambda(u; u_1, \ldots, u_M) = \alpha^N(u) \prod_{j=1}^{M} \frac{1}{c(u_j - u)} + \delta^N(u) \prod_{j=1}^{M} \frac{1}{c(u - u_j)}. \tag{11.43}$$

In order to derive the statements related to (11.42) and (11.43), we use the commutation relations (11.35), (11.36), and (11.37). We may focus on the operator A and thus the

commutation relations (11.35) and (11.36). The arguments will be analogous for the operator D and the commutation relations (11.35) and (11.37).

Since we already know how the operator A acts on the reference state $|0\rangle$, we are led to commute $A(u)$ through the string of operators $B(u_j)$ in the state (11.38) using (11.36) repeatedly until $A(u)$ directly acts on $|0\rangle$ to produce the eigenvalue $\alpha^N(u)$. Thus, if indeed there were only the first term on the right-hand side of the commutation relation (11.36), after M commutations, we would immediately reach the result

$$A(u)B(u_1)\dots B(u_M)|0\rangle = \alpha^N(u)\prod_{j=1}^{M}\frac{1}{c(u_j-u)}B(u_1)B\dots B(u_M)|0\rangle. \tag{11.44}$$

Thus, $B(u_1)B(u_2)\dots B(u_M)|0\rangle = |\{u_j\}\rangle$ would be the eigenvalue of $A(u)$ with eigenvalue $\alpha^N(u)\prod_{j=1}^{M}1/c(u_j-u)$.

However, the second term on the right-hand side of (11.36) spoils this straightforward picture. Commuting $A(u)$ through the operators $B(u_j)$ one by one, two new terms are created each time. After M commutations, we thus would end up with 2^M terms. To handle this large number of terms seems a formidable task.

However, there are two observations that help and that we alluded to when we briefly discussed the $1+M$ states (11.39) and (11.40):

- Only in one of the terms have all arguments remained unchanged, i.e. the argument of the operator A is still u, the arguments of the B operators in the product state (11.38) are still u_j for $j=1,\dots,M$. This term is given by (11.44).
- Many terms can be grouped together to form M combined terms, each of which is characterized by one missing argument, say u_n, in the arguments of the B operators. The number of B operators in the product is, however, unchanged by applying the commutation relation (11.36). The argument u_n has instead become the argument of the operator A, while one of the B's now bears the argument u. We thus have M combined terms proportional to

$$B(u)\prod_{j\neq n}^{M}B(u_j)A(u_n)|0\rangle = \alpha^N(u_n)B(u)\prod_{j\neq n}^{M}B(u_j)|0\rangle \qquad n=1,\dots,M. \tag{11.45}$$

We emphasize again, that, because of the commutation relation (11.36), there can be no types of terms other than the two just described. What remains to be calculated are the proportionality functions for the second type of terms. These functions will depend on u and all $\{u_j\}$.

We can now state the intermediate result that the first part of (11.41), the equation yet to become an eigenvalue equation, pertaining only to the operator $A(u)$ of the transfer matrix $T(u)$ has the form

$$A(u)|\{u_j\}\rangle = \alpha^N(u)\prod_{j=1}^{M}\frac{1}{c(u_j-u)}|\{u_j\}\rangle + \sum_{n=1}^{M}\Lambda_n^{(A)}\left(u,\{u_j\}\right)B(u)\prod_{j=1,\,j\neq n}^{M}B(u_j)|0\rangle. \tag{11.46}$$

The functions $\Lambda_n^{(A)}\left(u,\{u_j\}\right)$ are yet to be determined. A similar expression will hold for the second part of (11.41) pertaining to $D(u)$ with as yet unknown functions $\Lambda_n^{(D)}\left(u,\{u_j\}\right)$.

If we can show that the sum of these two functions

$$\left(\Lambda_n^{(A)}\left(u,\{u_j\}\right)+\Lambda_n^{(D)}\left(u,\{u_j\}\right)\right) \tag{11.47}$$

vanishes and leads to an expression in the form of the conditions (11.42), we shall have attained our aim to demonstrate that (11.41) becomes an eigenvalue equation with the eigenvalue given by (11.43).

In order to achieve this, let us first consider the function with $n=1$: $\Lambda_1^{(A)}\left(u,\{u_j\}\right)$. The part of the state (11.46) having this function as coefficient, i.e.

$$\Lambda_1^{(A)}\left(u,\{u_j\}\right)B(u)B(u_2)\ldots B(u_M)|0\rangle \tag{11.48}$$

is generated by applying the *second* part of the commutation relation (11.36) once to the product state (11.38). We thus obtain

$$-\frac{b(u_1-u)}{c(u_1-u)}B(u)A(u_1)B(u_2)\ldots B(u_M)|0\rangle. \tag{11.49}$$

Next, we pull $A(u_1)$ through all further $(M-1)$ B-operators to its right, but apply only the *first* part of (11.36) this time. This procedure yields

$$A(u_1)B(u_2)\ldots B(u_M)|0\rangle = \frac{1}{c(u_2-u_1)}B(u_2)A(u_1)B(u_3)\ldots B(u_M)|0\rangle, \tag{11.50}$$

$$= \prod_{j=2}^{M}\frac{1}{c(u_j-u_1)}B(u_2)\ldots B(u_M)A(u_1)|0\rangle, \tag{11.51}$$

$$= \alpha^N(u_1)\prod_{j=2}^{M}\frac{1}{c(u_j-u_1)}B(u_2)\ldots B(u_M)|0\rangle. \tag{11.52}$$

Only this procedure, applying the second part of the commutation relation once, the first part $(M-1)$ times, leads to a state proportional to

$$B(u)B(u_2)\dots B(u_M)|0\rangle = B(u)\prod_{j=2}^{M} B(u_j)|0\rangle. \tag{11.53}$$

Combining (11.49) and (11.52), we find the function

$$\Lambda_1^{(A)}\left(u,\{u_j\}\right) = -\alpha^N(u_1)\frac{b(u_1-u)}{c(u_1-u)}\prod_{j=2}^{M}\frac{1}{c(u_j-u_1)}. \tag{11.54}$$

Repeating the same arguments for the operator $D(u)$ applied to the reference state (11.38), we obtain the corresponding function

$$\Lambda_1^{(D)}\left(u,\{u_j\}\right) = -\delta^N(u_1)\frac{b(u-u_1)}{c(u-u_1)}\prod_{j=2}^{M}\frac{1}{c(u_1-u_j)}. \tag{11.55}$$

Finally, we recall the remarks after the definition of the state (11.38), which mean that we can write this state, for instance, in the form

$$|\{u_j\}\rangle = B(u_n)\prod_{j=1,\, j\neq n}^{M} B(u_j)|0\rangle. \tag{11.56}$$

Thus, we can go through the arguments above equally well for $B(u_n)$ instead of $B(u_1)$. In this general case, then, setting the sum of the functions $\Lambda_n^{(A)}$ and $\Lambda_n^{(D)}$ equal to zero implies the condition (11.42), which, in turn, renders (11.41) an eigenvalue equation.

The equations (11.42) are the celebrated Bethe ansatz equations, derived here within the algebraic framework.

The logic of what we have achieved now is the following. Suppose we find a solution, i.e. a set of M Bethe ansatz roots,[3] to the Bethe ansatz equations (11.42), then we succeed in diagonalizing the transfer matrix of the six-vertex model.

We are now only left with the task to show the connection of this approach with the Heisenberg quantum spin-1/2 chain. In order to do this, we have to consider the transfer matrix $T(u)$ in some more detail.

EXERCISE 11.4 Derive these results by directly applying the commutation relation (11.36), (11.35), and (11.37) on the state (11.38) respectively for the cases $M = 1$, which

[3] The solutions of the Bethe ansatz equations are sometimes also called Bethe ansatz roots. We shall be using both expressions.

is straightforward, and $M = 2$, which is already quite tedious. In the latter case, use the explicit forms of the functions $b(u)$ and $c(u)$. Larger values of M will obviously become unwieldy very rapidly.

11.4 Quantum Hamiltonians from the transfer matrix

We now construct quantum Hamiltonians from the transfer matrix $T(u)$. The latter can be viewed as the generating function of the former.

The trick is to find a special value, $u = u_0$, of the spectral parameter at which to evaluate the transfer matrix, or more precisely, its logarithm and logarithmic derivatives with respect to u, with the aim of writing these operators as sums of simpler operators.

The eigenvalue $\Lambda(u, \{u_j\})$ (11.43) suggests that we choose a value u_0 for the spectral parameter such that $\delta(u_0) = 0$, i.e. $u_0 = \eta$ in the six-vertex model.

Then, for this special value of the spectral parameter, $u = u_0$ with $\delta(u_0) = 0$, the operators

$$\log T(u_0) \qquad \text{and} \qquad \frac{d^k}{du^k} \log T(u)|_{u=u_0}, \qquad k = 1, 2, \ldots \tag{11.57}$$

all have an additive spectrum, i.e. a spectrum that corresponds to an operator that is a sum of simpler operators. This was the objective of our choice for the special value of the spectral parameter u_0.

We shall see that these operators are local quantum Hamiltonians, i.e. they can be written as sums of quantum operators on a one-dimensional lattice involving only a few lattice sites.

11.4.1 Isotropic case

We restrict our attention to the isotropic or XXX case (the reason for this terminology will become clear soon), which is obtained from the general six-vertex model after rescaling

$$u \to \kappa u \qquad \text{and} \qquad \eta \to \frac{i}{2}\kappa \tag{11.58}$$

in the limit $\kappa \to 0$.

Thus, with

$$\alpha(u) = u + \frac{i}{2} \qquad \text{and} \qquad \delta(u) = u - \frac{i}{2} \tag{11.59}$$

where also the functions $\alpha(u)$ and $\delta(u)$ have been rescaled, i.e.

$$\frac{1}{\kappa}\alpha(u) \to \alpha(u) \qquad \text{and} \qquad \frac{1}{\kappa}\delta(u) \to \delta(u), \tag{11.60}$$

the L_n-matrix becomes

$$L_n(u) = u + \frac{i}{2}\sum_{a=1}^{3}\sigma^a \otimes \sigma_n^a = \begin{pmatrix} u+\frac{i}{2}\sigma_n^z & \frac{i}{2}\sigma_n^- \\ \frac{i}{2}\sigma_n^+ & u-\frac{i}{2}\sigma_n^z \end{pmatrix}. \tag{11.61}$$

Moreover, with

$$b(u) = \frac{i}{u+i} \qquad \text{and} \qquad c(u) = \frac{u}{u+i} \tag{11.62}$$

we obtain the R-matrix of the isotropic six-vertex model explicitly

$$R(u) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{i}{u+i} & \frac{u}{u+i} & 0 \\ 0 & \frac{u}{u+i} & \frac{i}{u+i} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{2iI_2 \otimes I_2 + u\left(I_2 \otimes I_2 + \sum_\alpha \sigma^\alpha \otimes \sigma^\alpha\right)}{2i+2u}, \tag{11.63}$$

which is the simplest example of a rational R-matrix, i.e. an R-matrix where the matrix elements are rational functions of the spectral parameter u.

The special value of the spectral parameter is now

$$u = u_0 = \frac{i}{2}. \tag{11.64}$$

We have to evaluate $L_n(u)$, $T(u)$, etc, at this special value of $u = u_0 = \frac{i}{2}$.

At $u = u_0 = \frac{i}{2}$ we have

$$L_n\left(\frac{i}{2}\right) = \frac{i}{2}(\mathbb{I} + \boldsymbol{\sigma} \otimes \boldsymbol{\sigma}_n) = i\mathbb{P}_{0n} \tag{11.65}$$

where $\mathbb{I}$ is the 2×2 unit matrix where the matrix elements are themselves $2^N \times 2^N$ unit matrices and $\sigma \otimes \sigma_n$ stands for a 'scalar' product where the products of the components are, however, to be understood as tensor products

$$\boldsymbol{\sigma} \otimes \boldsymbol{\sigma}_n = \boldsymbol{\sigma} \otimes \left(\overbrace{I \otimes I \otimes \ldots \otimes \underbrace{\boldsymbol{\sigma}}_{n} \otimes \ldots \otimes I}^{N}\right). \tag{11.66}$$

The permutation operator $\mathbb{P}_{0n}$ (cp. (10.65)) can now be used to calculate the transfer matrix

$$T\left(\frac{i}{2}\right) = i^N \mathrm{Tr}\,(\mathbb{P}_{01}\ldots\mathbb{P}_{0N}) = i^N U_N^{-1}, \tag{11.67}$$

where U_N is the shift operator on the one-dimensional lattice

$$U_N \boldsymbol{\sigma}_n U_N^{-1} = \boldsymbol{\sigma}_{n+1} \tag{11.68}$$

with periodic boundary conditions understood.

EXERCISE 11.5 Consider the case $N = 2$ and show that

$$U_2 \sigma_1^x U_2^{-1} = \sigma_2^x. \tag{11.69}$$

A good starting point is to use the explicit form (11.61) for L_n at $u = \frac{i}{2}$.

The first logarithmic derivative of the transfer matrix is

$$\left(\frac{d}{du}\log T(u)\right)_{u=\frac{i}{2}} = \left(T^{-1}(u)\frac{d}{du}T(u)\right)_{u=\frac{i}{2}} = i^{-N}U_N\left(\frac{d}{du}T(u)\right)_{u=\frac{i}{2}} \tag{11.70}$$

$$= \frac{1}{i}U_N\sum_{n=1}^{N}\mathrm{Tr}\left(\mathbb{P}_{01}\ldots\underbrace{\mathbb{I}}_{n}\ldots\mathbb{P}_{0N}\right). \tag{11.71}$$

With

$$U_N\mathrm{Tr}\left(\mathbb{P}_{01}\ldots\underbrace{\mathbb{I}}_{n}\ldots\mathbb{P}_{0N}\right) = \mathbb{P}_{n,n+1} = \frac{1}{2}\left(\mathbb{I} + \boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_{n+1}\right) \tag{11.72}$$

we finally get

$$i\left(\frac{d}{du}\log T(u)\right)_{u=\frac{i}{2}} = \sum_{n=1}^{N}\frac{1}{2}\left(\mathbb{I} + \boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_{n+1}\right) \tag{11.73}$$

where again periodic boundary conditions $\sigma_{N+1} = \sigma_1$ are understood. The periodic boundary conditions are, of course, a reflection that we took the trace of the monodromy matrix $\mathcal{T}$ to obtain the transfer matrix T.

The right-hand side of (11.73) is the well-known Hamiltonian

$$\mathcal{H}_{\mathrm{XXX}} = \sum_{n=1}^{N}\frac{1}{2}\left(\mathbb{I} + \boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_{n+1}\right) \tag{11.74}$$

of the isotropic Heisenberg quantum spin chain of spins $\frac{1}{2}$, the so-called XXX quantum spin chain. This explains the terminology we have been using throughout this section.

EXERCISE 11.6 Consider again the case $N = 2$ and show that

$$U_2 \mathrm{Tr}\ (\mathbb{P}_{01}\mathbb{I}) = \mathbb{P}_{1,2} = \frac{1}{2}(\mathbb{I} + \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2). \tag{11.75}$$

The eigenvalues of the transfer matrix determine the eigenvalues of the Heisenberg XXX quantum spin chain. Taking the logarithmic derivative of (11.43), recalling that we choose $\delta(u = \frac{i}{2}) = 0$, we obtain

$$h = -\sum_{j=1}^{M} \frac{1}{u_j^2 + \frac{1}{4}} \tag{11.76}$$

where the parameters u_j are to be determined by the Bethe ansatz equations (11.42).

EXERCISE 11.7 Derive the expression (11.76) for the eigenvalue of the XXX quantum spin chain. You may have to redefine the Heisenberg XXX Hamiltonian to subtract a constant term.

11.5 Inhomogeneous Yang–Baxter quantum integrable models

The previous sections of this chapter showed how, the local L_n matrices play a central oart in the algebraic form of the Bethe ansatz. Their decisive role in the analysis consists in their encoding the vertex weights for a whole row of vertices in such a way that the position at site n of the individual vertex in the row can be specified. However, so far we have only considered the homogeneous case where the vertex weights themselves are not site-dependent. The Yang–Baxter relations, e.g. in the form (10.80), remain valid if the vertex weights themselves that enter the L_n matrices become inhomogeneous, i.e. site-dependent. The inhomogeneities created in this way can be interpreted as local defects or impurities.

Integrable models with such site-dependent defects or impurities have been mentioned throughout the literature (Baxter, 1982; Takhtajan, 1985; Shastry, 1988*a*; Korepin *et al*., 1993). This section only provides a brief overview of such models without detailed derivations.

11.5.1 Higher spin impurities

Andrei and Johannesson (1984) have considered an isotropic Heisenberg quantum spin chain with an impurity carrying a higher spin $S > 1/2$ for which they replaced one of the L_n matrices (11.61) by