

From Vertex Models to Quantum Spin Chains:  
An Introductory Exposition of the Yang–Baxter Equation

by Fedor Prikolab

Summer 2023

# Preface

This document grew out of the summer research project PHY479Y1 at the University of Toronto, carried out under the supervision of Artur Izmaylov during summer of 2023. The initial goal was to understand the transfer-matrix formalism for classical vertex models and to see how it inspired the transfer-matrix approach to quantum spin chains. This was meant as a first step toward the broader question of how one can construct exactly solvable quantum models, with quantum spin chains as the guiding examples.

In practice, the project turned out to be quite challenging. The existing literature on the Yang–Baxter equation and integrable models is very rich, but many expositions skip intermediate steps that are crucial for a reader encountering the subject for the first time. The aim of this work is therefore twofold: to present the flow of ideas in a clear and systematic way, and to keep the level accessible to an advanced undergraduate reader with some familiarity with representation theory.

The exposition is organized as follows. In Chapter 1 we introduce the eight-vertex model and show how the transfer-matrix formalism is used to obtain its partition function. Already here we begin a discussion of symmetries, which serves as a guiding thread for the rest of the text. Chapter 2 develops the parallel formalism for quantum spin chains and introduces the Yang–Baxter equation, ending with a discussion of how new exactly solvable quantum models can, in principle, be generated. Chapter 3 presents the algebraic Bethe ansatz and explains how it is connected to the transfer-matrix construction. Chapter 4 offers a brief introduction to quasi-triangular Hopf algebras and returns to the question of how they provide a general framework for constructing solutions to the Yang–Baxter equation. An appendix contains some simple numerical tests illustrating the discussion of the six-vertex model.

Parts of this document have been revised since the original project as my understanding of the subject has improved, but I have tried to keep the spirit of the original undergraduate research project. I am very grateful to Artur Izmaylov and Smik Patel for their guidance and help throughout this work.

# Contents

<b>Preface</b>	<b>1</b>
<b>1 The Eight-Vertex Model and the Transfer Matrix</b>	<b>4</b>
1.1 From the Ising Model to Arrow Configurations . . . . .	4
1.2 Local Weights and the Partition Function . . . . .	6
1.3 Row-to-Row Transfer Matrix . . . . .	6
1.4 Matrix Form of the Local $L$ -Operator . . . . .	7
1.5 Tensorial Description and Traces over Horizontal and Vertical Space . . . . .	10
1.6 From the Eight-Vertex Model to the XYZ Spin Chain . . . . .	11
<b>2 Introducing Yang-Baxter equation with Quantum Spin Chains</b>	<b>13</b>
2.1 Introduction to quantum spin chains and their Lax operators . . . . .	13
2.1.1 Imposing $G$ -invariance on the Lax operator . . . . .	14
2.1.2 Example 1: XXX spin- $\frac{1}{2}$ chain . . . . .	14
2.1.3 Example 2: XXZ spin- $\frac{1}{2}$ chain . . . . .	15
2.2 The $R$ -matrix, transfer matrices, and the Yang-Baxter equation . . . . .	21
2.2.1 Coming back to transfer matrix formalism . . . . .	21
2.2.2 The $R$ -matrix as a sufficient condition for commutativity . . . . .	21
2.2.3 Consistency and the Yang-Baxter equation . . . . .	22
2.3 Deriving conserved quantities of the Hamiltonian from the transfer matrix . . . . .	23
2.3.1 Commutativity of the transfer matrix with the Hamiltonian . . . . .	27
2.3.2 Conserved quantities from the analytic expansion of the transfer matrix . . . . .	28
2.4 Deriving Solvable models from Yang-Baxter equation . . . . .	28
2.4.1 Symmetry and $G$ -invariance of the $R$ -matrix . . . . .	28
2.4.2 Reverse engineering integrable quantum spin chains . . . . .	30
<b>3 Algebraic Bethe Ansatz</b>	<b>32</b>
3.1 Introduction . . . . .	32
3.2 Ferromagnetic/Antiferromagnetic case . . . . .	34
3.3 Components of Monodromy Matrix and the ferromagnetic pseudo-vacuum . . . . .	34
3.4 Summary of the Algebraic Bethe Ansatz for the XXZ spin chain . . . . .	37
<b>4 Quasi-triangular Hopf algebras and solutions of the Yang-Baxter equation</b>	<b>39</b>
4.1 The Yang-Baxter equation revisited . . . . .	39
4.2 Hopf algebras: algebra and coalgebra in one object . . . . .	39
4.2.1 Algebras and coalgebras . . . . .	39
4.2.2 Hopf algebras . . . . .	40
4.2.3 Representations and the tensor product via the coproduct . . . . .	40
4.3 Quasi-triangular Hopf algebras and the universal $R$ -matrix . . . . .	41
4.3.1 Definition and axioms . . . . .	41
4.3.2 The universal Yang-Baxter equation in $H^{\otimes 3}$ . . . . .	41
4.4 From universal $R$ to $R$ -matrices on representations . . . . .	41
4.4.1 Induced operators on $V \otimes W$ . . . . .	42

4.4.2	The Yang–Baxter equation on $U \otimes V \otimes W$ . . . . .	42
4.5	Hopf algebra symmetry of the spin- $\frac{1}{2}$ chain . . . . .	43
4.5.1	The Lie algebra $\mathfrak{sl}_2$ and its enveloping Hopf algebra . . . . .	43
4.5.2	Spin- $\frac{1}{2}$ representation and tensor product chain . . . . .	43
4.5.3	Heisenberg Hamiltonian and $\mathfrak{sl}_2$ symmetry . . . . .	44
4.5.4	Remark: from $U_q(\mathfrak{sl}_2)$ to non-trivial $R$ -matrices . . . . .	44
<b>5</b>	<b>Appendix: A symbolic two–site experiment</b>	<b>45</b>

# Chapter 1

## The Eight–Vertex Model and the Transfer Matrix

In this chapter we introduce the eight–vertex model as a classical lattice model of statistical mechanics and explain how its local Boltzmann weights can be organised into a row–to–row transfer matrix. Later in the report, this transfer matrix will serve as an entry point to the Yang–Baxter equation and to integrable quantum spin chains, but in this chapter we stay entirely on the classical, combinatorial level.

### 1.1. From the Ising Model to Arrow Configurations

A natural way to motivate the eight–vertex model is to start from the Ising model and gradually change the point of view.

Consider the Ising model on a two–dimensional  $N \times M$  square lattice with periodic boundary conditions. At each site  $(i, j)$  we have a classical spin variable

$$\sigma_{i,j} \in \{+1, -1\},$$

and spins interact along horizontal and vertical edges. For a horizontal edge connecting  $(i, j)$  and  $(i + 1, j)$  the local contribution to the energy depends only on the product  $\sigma_{i,j}\sigma_{i+1,j}$ , and similarly for a vertical edge.

It is convenient to encode *relative* orientations of neighbouring spins by arrows on the edges instead of storing the spins themselves. We encode *relative* orientations of neighbouring spins by introducing, on each edge  $\langle x, y \rangle$ , the *bond variable*

$$\tau_{xy} = \sigma_x \sigma_y \in \{+1, -1\}.$$

Focus on a single elementary plaquette (which we will regard as a single “vertex” of the arrow model). Around this plaquette we have four Ising sites, say

$$\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \{\pm 1\},$$

arranged cyclically, and four incident edges

$$e_1 = \langle 1, 2 \rangle, \quad e_2 = \langle 2, 3 \rangle, \quad e_3 = \langle 3, 4 \rangle, \quad e_4 = \langle 4, 1 \rangle.$$

Each of these edges carries a bond variable (or arrow)  $\tau_{e_k} = \sigma_i \sigma_j$ .

By construction,  $\tau_{xy}$  depends only on whether the two spins are aligned or anti–aligned, not on their common sign. We now represent  $\tau_{xy}$  graphically by an arrow:

$$\begin{aligned} \tau_{xy} = +1 &\iff \text{arrow pointing “right” or “up” along the edge,} \\ \tau_{xy} = -1 &\iff \text{arrow pointing “left” or “down” along the edge.} \end{aligned}$$

In this way each edge carries a binary variable — the direction of the arrow — and the original spin configuration can be reconstructed from the collection of bond variables up to a global spin flip  $\sigma \mapsto -\sigma$ . Since the Ising energy depends only on nearest–neighbour products  $\sigma_x \sigma_y$ , it can be rewritten purely in terms of the bond variables  $\tau_{xy}$ , or equivalently, in terms of the arrow configuration.

### Local arrow patterns at a vertex

Focus on a single vertex of the lattice. There are four edges meeting at this vertex: left, right, up, and down. Before imposing any consistency conditions, each arrow can point in two directions, so there are  $2^4 = 16$  possible arrow patterns at a vertex.

However, not all 16 patterns are consistent with the rule that arrows are induced by differences of neighbouring Ising spins. Two basic constraints arise:

- (i) **Ising  $\mathbb{Z}_2$  symmetry.** Flipping *all spins* at the four corners around a vertex,

$$\sigma \mapsto -\sigma,$$

does not change any nearest-neighbour products and hence does not change the local energy. On the arrow level this operation reverses *all four arrows* at once and must leave the local energy invariant.

- (ii) **Consistency of spins around the vertex.** At a given vertex we have four sites, say  $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \{\pm 1\}$ , arranged cyclically, and four incident edges

$$e_1 = \langle 1, 2 \rangle, \quad e_2 = \langle 2, 3 \rangle, \quad e_3 = \langle 3, 4 \rangle, \quad e_4 = \langle 4, 1 \rangle.$$

The corresponding bond variables are

$$\tau_{e_1} = \sigma_1 \sigma_2, \quad \tau_{e_2} = \sigma_2 \sigma_3, \quad \tau_{e_3} = \sigma_3 \sigma_4, \quad \tau_{e_4} = \sigma_4 \sigma_1.$$

Their product is

$$\tau_{e_1} \tau_{e_2} \tau_{e_3} \tau_{e_4} = (\sigma_1 \sigma_2)(\sigma_2 \sigma_3)(\sigma_3 \sigma_4)(\sigma_4 \sigma_1) = \sigma_1^2 \sigma_2^2 \sigma_3^2 \sigma_4^2 = 1.$$

Thus for any arrow pattern coming from a genuine spin configuration, the product of the four bond variables (or, equivalently, of the four arrow signs) around a vertex must equal +1. This shows that the four arrows cannot be chosen independently: only those arrow patterns for which

$$\prod_{\text{edges } e \text{ incident at the vertex}} \tau_e = 1$$

are admissible. Imposing this local constraint reduces the naive  $2^4 = 16$  patterns to the eight vertex configurations of the model.

Imposing these constraints reduces the 16 naive arrow patterns to 8 *admissible* vertex configurations. They come in four inversion-symmetric pairs (related by reversing all arrows). We denote the corresponding local energies by

$$\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4,$$

and the associated Boltzmann weights by

$$v_i = e^{-\beta \epsilon_i}, \quad i = 1, \dots, 4.$$

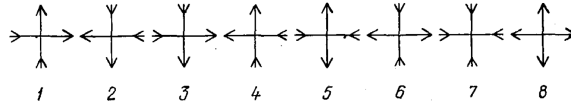


Figure 1.1: The eight allowed arrow configurations at a vertex in the eight-vertex model. Opposite configurations in each pair are related by reversing all arrows, and share the same Boltzmann weight.

The resulting model, defined in terms of these eight arrow patterns, is the *eight-vertex model*. It may be viewed as a generalisation of the Ising model: instead of specifying pairwise couplings between spins, we assign an energy directly to each allowed local arrow pattern at a vertex.

## 1.2. Local Weights and the Partition Function

We now introduce a convenient notation for the arrows around a single vertex and define the partition function of the model.

### Labelling the arrows

We label the four arrows around a vertex by

$$(\gamma, \alpha'; \gamma', \alpha),$$

where:

- $\gamma$  and  $\gamma'$  denote the left and right horizontal arrows,
- $\alpha$  and  $\alpha'$  denote the down and up vertical arrows.

Each of these variables takes two values, which we may call  $\pm 1$  for definiteness (for example,  $+1$  for an arrow pointing right or up,  $-1$  for an arrow pointing left or down).

To every admissible arrow pattern we assign a Boltzmann weight

$$L_{\alpha}^{\alpha'}(\gamma, \gamma') \in \{v_1, v_2, v_3, v_4\},$$

and to each forbidden pattern we assign weight 0. The function  $L$  is the *local Boltzmann weight* (or *L-operator*) of the eight-vertex model.

### Total energy and partition function

Let  $v$  denote a full arrow configuration on the  $N \times M$  lattice, that is, a choice of arrow direction on each edge. The total energy of  $v$  is the sum of local contributions over all vertices:

$$E_v = \sum_{\text{vertices } (i,j)} E_{ij},$$

where  $E_{ij}$  is determined by the local pattern of arrows around vertex  $(i, j)$ .

The partition function is defined in the usual way as

$$\mathcal{Z} = \sum_{\{v\}} e^{-\beta E_v} = \sum_{\{v\}} \prod_{\text{vertices } (i,j)} L_{\alpha_{ij}}^{\alpha'_{ij}}(\gamma_{ij}, \gamma'_{ij}),$$

where the product runs over all vertices, and the sum runs over all arrow configurations compatible with the eight allowed local patterns.

## 1.3. Row-to-Row Transfer Matrix

We next reorganise the partition function in terms of *rows*, which leads naturally to the notion of a transfer matrix.

For simplicity, fix the number of columns to be  $N$  and consider a pair of consecutive rows, labelled by  $j$  and  $j+1$ . Let

$$\{\alpha^j\} = (\alpha_1^j, \dots, \alpha_N^j)$$

denote the vertical arrows pointing from row  $j$  to row  $j+1$  along each column. Similarly, denote by

$$\{\alpha^{j+1}\} = (\alpha_1^{j+1}, \dots, \alpha_N^{j+1})$$

the vertical arrows from row  $j+1$  to row  $j+2$  (or back to row 1 if we impose periodicity in the vertical direction).

### Weight of a single row

Fix two vertical arrow configurations  $\{\alpha^j\}$  and  $\{\alpha^{j+1}\}$ . The arrows within the strip between these two rows are the horizontal arrows

$$\gamma_1, \dots, \gamma_N$$

along the edges connecting the sites in row  $j$  to those in row  $j+1$ . For each  $n$ , the local weight at the vertex in column  $n$  depends only on the four arrows

$$(\gamma_n, \alpha_n^j; \gamma_{n+1}, \alpha_n^{j+1}),$$

with the periodic convention  $\gamma_{N+1} = \gamma_1$  when the lattice has periodic boundary conditions in the horizontal direction.

The Boltzmann weight of the entire row (i.e. of all vertices between rows  $j$  and  $j+1$ ) is obtained by summing over all horizontal arrows:

$$T_{\{\alpha^j\}, \{\alpha^{j+1}\}} = \sum_{\gamma_1=\pm 1} \cdots \sum_{\gamma_N=\pm 1} \prod_{n=1}^N L_{\alpha_n^j}^{\alpha_n^{j+1}}(\gamma_n, \gamma_{n+1}).$$

This quantity depends only on the two boundary configurations  $\{\alpha^j\}$  and  $\{\alpha^{j+1}\}$ , and by construction it is non-negative.

### The transfer matrix

If we think of the  $2^N$  possible vertical arrow configurations  $\{\alpha^j\}$  as forming a basis of a  $2^N$ -dimensional vector space, then the numbers  $T_{\{\alpha^j\}, \{\alpha^{j+1}\}}$  are the matrix elements of a linear operator

$$T_N : \text{span}\{\text{vertical arrow configurations}\} \longrightarrow \text{span}\{\text{vertical arrow configurations}\},$$

called the *row-to-row transfer matrix*.

Let  $M$  be the number of rows. For each row  $j = 1, \dots, M$  let

$$\{\alpha^j\} = (\alpha_1^j, \dots, \alpha_N^j)$$

denote the vertical arrow configuration between rows  $j$  and  $j+1$ , with the periodic identification  $\{\alpha^{M+1}\} = \{\alpha^1\}$ . The matrix element of the row-to-row transfer matrix  $T_N$  is

$$T_{\{\alpha^j\}, \{\alpha^{j+1}\}} = \sum_{\gamma_1=\pm 1} \cdots \sum_{\gamma_N=\pm 1} \prod_{n=1}^N L_{\alpha_n^j}^{\alpha_n^{j+1}}(\gamma_n, \gamma_{n+1}),$$

with  $\gamma_{N+1} = \gamma_1$  for horizontal periodicity.

The partition function can then be written as a sum over all vertical arrow configurations:

$$\begin{aligned} \mathcal{Z} &= \sum_{\{\alpha^1\}} \sum_{\{\alpha^2\}} \cdots \sum_{\{\alpha^M\}} T_{\{\alpha^1\}, \{\alpha^2\}} T_{\{\alpha^2\}, \{\alpha^3\}} \cdots T_{\{\alpha^M\}, \{\alpha^1\}} \\ &= \sum_{\{\alpha^1\}} (T_N^M)_{\{\alpha^1\}, \{\alpha^1\}} \\ &= \text{tr}_v(T_N^M). \end{aligned}$$

In the second line we recognise the usual formula for the diagonal matrix elements of  $T_N^M$ , and  $\text{tr}_v$  denotes the trace over the “vertical” space spanned by all vertical arrow configurations. In words: the transfer matrix  $T_N$  propagates arrow configurations from one row to the next, and the trace implements the periodic identification between the first and last row.

## 1.4. Matrix Form of the Local $L$ -Operator

The local weight  $L_\alpha^{\alpha'}(\gamma, \gamma')$  can itself be viewed as a matrix once we choose an ordering of the input and output arrow pairs.

### Choice of basis and local symmetry

We view each arrow as a two-state variable taking values in  $\{\pm 1\}$ . At a single vertex we have a pair of horizontal arrows  $(\gamma, \gamma')$  and a pair of vertical arrows  $(\alpha, \alpha')$ , so both the horizontal and the vertical degrees of freedom are described by a two-spin space

$$V \cong \mathbb{C}^2 \otimes \mathbb{C}^2.$$

We choose the ordered basis

$$|1\rangle = |-, -\rangle, \quad |2\rangle = |-, +\rangle, \quad |3\rangle = |+, -\rangle, \quad |4\rangle = |+, +\rangle,$$

for both the horizontal and the vertical pairs. In this basis the local  $L$ -operator

$$L : V_{\text{hor}} \longrightarrow V_{\text{vert}},$$

becomes a  $4 \times 4$  matrix and we identify its matrix elements with Boltzmann weights in the usual way.

For the eight-vertex model the local  $L$ -matrix has the well-known form

$$L = \begin{pmatrix} a & 0 & 0 & d \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ d & 0 & 0 & a \end{pmatrix}.$$

The four parameters  $(a, b, c, d)$  encode the Boltzmann weights of the eight allowed arrow configurations. More explicitly,

$$a = L_{-1}^{-1}(-1, -1) = L_{+1}^{+1}(+1, +1),$$

$$b = L_{-1}^{+1}(-1, +1) = L_{+1}^{-1}(+1, -1),$$

$$c = L_{-1}^{+1}(+1, -1) = L_{+1}^{-1}(-1, +1),$$

$$d = L_{-1}^{-1}(+1, +1) = L_{+1}^{+1}(-1, -1),$$

and all other entries vanish because the corresponding arrow patterns are forbidden in the eight-vertex model. The equalities on the right express the fact that configurations related by simultaneous reversal of all four arrows have the same Boltzmann weight.

To describe the local symmetry more systematically it is convenient to introduce Pauli matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

acting on a single spin, and to view

$$V = \mathbb{C}^2 \otimes \mathbb{C}^2$$

as the tensor product of two spin- $\frac{1}{2}$  spaces. On  $V$  we define the following commuting operators:

$$A_1 := \sigma_1^z \sigma_2^z = \text{diag}(1, -1, -1, 1),$$

$$A_2 := |2\rangle\langle 2| + |3\rangle\langle 3| = \text{diag}(0, 1, 1, 0),$$

$$A_3 := |1\rangle\langle 1| + |4\rangle\langle 4| = \text{diag}(1, 0, 0, 1).$$

Then

$$A_2^2 = A_2, \quad A_3^2 = A_3, \quad A_2 A_3 = A_3 A_2 = 0,$$

so  $A_2$  and  $A_3$  are orthogonal projectors onto the subspaces with opposite and equal arrows respectively, while  $A_1$  is an involution,

$$A_1^2 = \mathbb{I}, \quad A_1 = A_3 - A_2, \quad A_2 + A_3 = \mathbb{I}.$$

Thus the algebra generated by  $A_1, A_2, A_3$  is the commutative two-dimensional algebra

$$\mathcal{A} = \text{span}\{\mathbb{I}, A_1\} = \text{span}\{\mathbb{I}, A_2, A_3\},$$

with minimal central idempotents

$$P_+ = A_3 = \frac{\mathbb{I} + A_1}{2}, \quad P_- = A_2 = \frac{\mathbb{I} - A_1}{2}.$$

These project onto the eigenspaces of  $A_1$  with eigenvalues  $\pm 1$ . Geometrically,  $A_1$  distinguishes vertices with an even/odd number of up-arrows in the local configuration.

There is a second natural involution on  $V$  given by simultaneous arrow reversal:

$$F := \sigma_1^x \sigma_2^x.$$

In the chosen basis  $F$  acts as

$$F|1\rangle = |4\rangle, \quad F|4\rangle = |1\rangle, \quad F|2\rangle = |3\rangle, \quad F|3\rangle = |2\rangle,$$

so its matrix is

$$F = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad F^2 = \mathbb{I}.$$

One easily checks that

$$[A_1, F] = 0, \quad [A_2, F] = 0, \quad [A_3, F] = 0,$$

so  $A_1$  and  $F$  are commuting involutions.

We now define the local symmetry group

$$G := \langle g_z, g_x \mid g_z^2 = g_x^2 = e, \ g_z g_x = g_x g_z \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2,$$

with a representation

$$\rho : G \longrightarrow \text{End}(V)$$

given on generators by

$$\rho(g_z) = A_1 = \sigma_1^z \sigma_2^z, \quad \rho(g_x) = F = \sigma_1^x \sigma_2^x.$$

We equip both  $V_{\text{hor}}$  and  $V_{\text{vert}}$  with this same representation of  $G$ , so they are isomorphic  $G$ -modules:

$$(V_{\text{hor}}, \rho_{\text{hor}}) \cong (V_{\text{vert}}, \rho_{\text{vert}}) \cong (V, \rho).$$

In terms of the projectors  $A_2, A_3$  and the flip  $F$  the eight-vertex  $L$ -matrix can be written as

$$L = a A_3 + b A_2 + d A_3 F + c A_2 F. \tag{1.1}$$

Indeed,  $A_3$  projects onto the subspace spanned by  $|1\rangle, |4\rangle$ , on which the block of  $L$  is

$$\begin{pmatrix} a & d \\ d & a \end{pmatrix} = a \mathbb{I}_{(1,4)} + d F_{(1,4)},$$

while  $A_2$  projects onto the subspace spanned by  $|2\rangle, |3\rangle$ , on which the block of  $L$  is

$$\begin{pmatrix} b & c \\ c & b \end{pmatrix} = b \mathbb{I}_{(2,3)} + c F_{(2,3)}.$$

Formula (1.1) simply combines these two blocks.

Since  $A_2, A_3, F$  all commute with  $\rho(g_z)$  and  $\rho(g_x)$ , it follows immediately that

$$[L, \rho(g_z)] = 0, \quad [L, \rho(g_x)] = 0,$$

and hence  $[L, \rho(g)] = 0$  for every  $g \in G$ . Equivalently, if we regard  $L$  as a map

$$L : (V_{\text{hor}}, \rho_{\text{hor}}) \longrightarrow (V_{\text{vert}}, \rho_{\text{vert}}),$$

then for all  $g \in G$  we have

$$\rho_{\text{vert}}(g) L = L \rho_{\text{hor}}(g).$$

Thus the local Boltzmann weight operator  $L$  is a  $G$ -equivariant map, or in other words:

*The local  $L$ -operator of the eight-vertex model is a  $G$ -morphism for the symmetry group  $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$  generated by the charge involution  $\sigma_1^z \sigma_2^z$  and the simultaneous arrow-flip  $\sigma_1^x \sigma_2^x$ .*

This symmetry statement explains the restricted pattern of nonzero entries of  $L$ : a general  $4 \times 4$  matrix commuting with the  $G$ -action must be block-diagonal with two identical  $2 \times 2$  blocks on the subspaces  $A_3 V$  and  $A_2 V$ , each of the form  $\alpha \mathbb{I} + \beta F$ , and this is precisely the structure of the eight-vertex  $L$ -matrix.

## 1.5. Tensorial Description and Traces over Horizontal and Vertical Space

So far, everything has been phrased in terms of classical arrow configurations and combinatorial sums. However, the matrix viewpoint already suggests a natural ‘‘Hilbert-space’’ interpretation that will be useful later when we discuss quantum spin chains.

### Local Hilbert space and local symmetry at a vertex

Each arrow variable takes two values, so it is natural to identify the set  $\{\pm 1\}$  with a basis of a two-dimensional complex vector space:

$$\mathbb{C}^2 = \text{span}\{|-\rangle, |+\rangle\}.$$

If we make this identification, then

- a single vertical edge carries a copy of  $\mathbb{C}^2$ ,
- a single horizontal edge also carries a copy of  $\mathbb{C}^2$ ,
- the four arrows around a vertex live in  $\mathbb{C}^2 \otimes \mathbb{C}^2$  (vertical pair) and  $\mathbb{C}^2 \otimes \mathbb{C}^2$  (horizontal pair).

Thus the local  $L$ -operator at a vertex is a linear map

$$L : V_{\text{hor}} \longrightarrow V_{\text{vert}}, \quad V_{\text{hor}} \cong V_{\text{vert}} \cong \mathbb{C}^2 \otimes \mathbb{C}^2,$$

with matrix elements  $L_{\alpha}^{\alpha'}(\gamma, \gamma')$  in the chosen basis.

Based on the results of previous section we want to emphasize that this local Hilbert space carries a natural *representation of a local symmetry group*

$$G = \mathbb{Z}_2 \times \mathbb{Z}_2 = \langle g_z, g_x \mid g_z^2 = g_x^2 = e, g_z g_x = g_x g_z \rangle.$$

In particular, both the horizontal and vertical two-spin spaces are  $G$ -modules:

$$(V_{\text{hor}}, \rho_{\text{hor}}) \cong (V, \rho), \quad (V_{\text{vert}}, \rho_{\text{vert}}) \cong (V, \rho),$$

where

$$\rho(g_z) = g_z = \sigma^z \otimes \sigma^z, \quad \rho(g_x) = g_x = \sigma^x \otimes \sigma^x.$$

Physically,  $g_x$  implements simultaneous reversal of the two arrows, while  $g_z$  measures the parity (or ‘‘charge’’) of the configuration. Because these two operations are local (they act only on the two arrows attached to a given vertex) and commute with each other, the local Hilbert space at a vertex is naturally a representation of the local symmetry group  $\mathbb{Z}_2 \times \mathbb{Z}_2$ .

Although we started from a purely classical model with discrete arrow configurations and Boltzmann weights, this formulation already places the eight-vertex model into the language of quantum spin systems: each edge carries a two-dimensional quantum spin- $\frac{1}{2}$  space, and each vertex carries a local Hilbert space which is a  $G$ -module for the symmetry group  $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ . In later sections we will see again the condition for local  $L$ -operator to be a  $G$ -equivariant map under the local representation.

## Monodromy matrix and horizontal trace

Along a single row there are  $N$  vertices, hence  $N$  copies of the local  $L$ -operator. We denote by  $L_n$  the operator acting at column  $n$ . It acts nontrivially only on the tensor factor

$$\mathbb{C}_{h,n}^2 \otimes \mathbb{C}_{v,n}^2 \subset \mathcal{H}_h \otimes \mathcal{H}_v,$$

where  $\mathcal{H}_h$  is the space of horizontal arrow configurations in the row, and  $\mathcal{H}_v$  is the space of vertical arrow configurations (one copy of  $\mathbb{C}^2$  per column).

The ordered product

$$\mathcal{T}_N = L_N L_{N-1} \cdots L_1$$

is called the *monodromy matrix* of the row. It acts on  $\mathcal{H}_h \otimes \mathcal{H}_v$  and collects the effect of all  $N$  vertices between two neighbouring rows.

To obtain the row-to-row transfer matrix we sum over all internal horizontal arrows, i.e. we take the trace over the horizontal space:

$$T_N = \text{tr}_h(\mathcal{T}_N).$$

In components, if we write vertical configurations as

$$|\{\alpha^j\}\rangle_v = |\alpha_1^j, \dots, \alpha_N^j\rangle_v$$

and horizontal configurations as

$$|\{\gamma\}\rangle_h = |\gamma_1, \dots, \gamma_N\rangle_h,$$

then

$$\begin{aligned} T_{\{\alpha^j\}, \{\alpha^{j+1}\}} &= \sum_{\{\gamma\}} (\mathcal{T}_N)^{\{\alpha^{j+1}\}, \{\gamma\}}_{\{\alpha^j\}, \{\gamma\}} \\ &= \sum_{\gamma_1=\pm 1} \cdots \sum_{\gamma_N=\pm 1} \prod_{n=1}^N L_{\alpha_n^j}^{\alpha_n^{j+1}}(\gamma_n, \gamma_{n+1}), \end{aligned}$$

with the periodic convention  $\gamma_{N+1} = \gamma_1$ . This reproduces the explicit formula for  $T_N$  used earlier. The partition function on an  $N \times M$  lattice with periodic boundary conditions in the vertical direction can then be written as

$$\mathcal{Z} = \text{tr}_v(T_N^M),$$

a trace over the vertical space  $\mathcal{H}_v$ .

## A hint towards auxiliary and quantum spaces

From the point of view of later quantum-mechanical models, it is useful to keep in mind the following interpretation:

- the vertical space, on which  $T_N$  acts, is analogous to the *quantum space* of a spin chain, typically  $(\mathbb{C}^2)^{\otimes N}$  for a chain of  $N$  spin- $\frac{1}{2}$  sites;
- the horizontal space, over which we take the trace when forming  $T_N$ , plays the role of an *auxiliary space*.

In the classical eight-vertex model these are simply spaces of arrow configurations, but the same structure will reappear in the quantum setting, where arrows are replaced by spin states in a Hilbert space and the local  $L$ -operator becomes a matrix of operators acting on that space.

## 1.6. From the Eight-Vertex Model to the XYZ Spin Chain

We conclude this chapter by briefly explaining how the eight-vertex model is related to quantum spin chains, and in particular to the XYZ Heisenberg chain. The goal here is only to outline the dictionary between the two pictures, leaving technical details for later chapters.

### Arrows as spin- $\frac{1}{2}$ states

Once we interpret the two arrow values as the basis vectors  $\{|-\rangle, |+\rangle\}$  of  $\mathbb{C}^2$ , each vertical edge of the lattice carries a two-dimensional Hilbert space. Along a row with  $N$  columns, the collection of vertical arrows

$$\{\alpha^j\} = (\alpha_1^j, \dots, \alpha_N^j)$$

corresponds to a basis vector in the tensor product

$$\mathcal{H}_{\text{quantum}} = (\mathbb{C}^2)^{\otimes N}.$$

The transfer matrix  $T_N$  then acts on this space, and its matrix elements in the arrow basis are precisely the row weights we obtained by summing over horizontal arrows.

In this language, the classical partition function,

$$\mathcal{Z} = \text{tr}_v(T_N^M),$$

looks formally like the trace of a (discrete) “imaginary-time” evolution operator built from  $T_N$ .

### Family of transfer matrices and effective Hamiltonian

The parameters  $(a, b, c, d)$  entering the local  $L$ -operator can be made to depend smoothly on an additional continuous parameter  $u$ , called the *spectral parameter*. This produces a one-parameter *family* of transfer matrices  $T_N(u)$  acting on the same quantum space  $(\mathbb{C}^2)^{\otimes N}$ .

Under appropriate conditions on the dependence of  $(a, b, c, d)$  on  $u$  (to be discussed later), one can show that the operators  $T_N(u)$  for different values of  $u$  commute with each other. In that situation it will become possible to define an effective Hamiltonian by taking the logarithmic derivative of the transfer matrix at some special value  $u_0$ :

$$H_{\text{XYZ}} \propto \left. \frac{d}{du} \log T_N(u) \right|_{u=u_0}.$$

The resulting operator  $H_{\text{XYZ}}$  acts on  $(\mathbb{C}^2)^{\otimes N}$  and has the structure of a local quantum spin chain Hamiltonian: it can be written as a sum of nearest-neighbour terms involving the Pauli matrices at sites  $j$  and  $j+1$ ,

$$H_{\text{XYZ}} = \sum_{j=1}^N (J_x \sigma_j^x \sigma_{j+1}^x + J_y \sigma_j^y \sigma_{j+1}^y + J_z \sigma_j^z \sigma_{j+1}^z) \quad (\text{with periodic boundary conditions}),$$

for suitable couplings  $J_x, J_y, J_z$  determined by the eight-vertex weights.

Thus, starting from a classical model with eight local arrow configurations and discrete energies, we obtain a quantum Hamiltonian describing interacting spin- $\frac{1}{2}$  particles on a chain.

### Dictionary between the classical model and the quantum chain

To summarise, the bridge between the classical eight-vertex model and the XYZ quantum spin chain can be outlined as follows:

- classical arrow variables  $\pm 1$  along vertical edges  $\longleftrightarrow$  basis states  $|-\rangle, |+\rangle$  in  $\mathbb{C}^2$  at each quantum site;
- row-to-row transfer matrix  $T_N \longleftrightarrow$  linear operator on the quantum space  $(\mathbb{C}^2)^{\otimes N}$ ;
- trace over horizontal arrows to build  $T_N \longleftrightarrow$  trace over an auxiliary space;
- trace over vertical configurations to form  $\mathcal{Z} \longleftrightarrow$  trace over the quantum space, reminiscent of a partition function of a quantum system;
- suitable parametrisation of the local weights  $(a, b, c, d)$  in terms of a spectral parameter  $\longleftrightarrow$  family of commuting transfer matrices  $T_N(u)$  and an associated local Hamiltonian of XYZ type.

In later chapters we will place this correspondence in a broader framework. For now, the key takeaway is that the classical eight-vertex model, formulated in terms of discrete energy configurations at each vertex, already contains the algebraic structure needed to build a Hilbert space at each site and to construct a quantum spin chain Hamiltonian from its transfer matrix.

## Chapter 2

# Introducing Yang-Baxter equation with Quantum Spin Chains

### 2.1. Introduction to quantum spin chains and their Lax operators

In this chapter we pass from the classical eight-vertex model of the previous section to its quantum counterpart: one-dimensional spin chains. Our basic object is a chain of  $N$  spin- $\frac{1}{2}$  sites, with Hilbert space

$$\mathcal{H} = \bigotimes_{k=1}^N \mathfrak{h}_k, \quad \mathfrak{h}_k \cong \mathbb{C}^2,$$

where each local space  $\mathfrak{h}_k$  carries a fixed spin- $\frac{1}{2}$  representation of a symmetry group  $G$ .

We consider nearest-neighbour Hamiltonians of the form

$$H = \sum_{k=1}^{N-1} H_{k,k+1},$$

where each local term  $H_{k,k+1}$  acts on  $\mathfrak{h}_k \otimes \mathfrak{h}_{k+1}$  and is assumed to be invariant under a fixed symmetry group  $G$ :

$$[H_{k,k+1}, \rho_k(g) \otimes \rho_{k+1}(g)] = 0, \quad \forall g \in G, \quad k = 1, \dots, N-1.$$

Here  $\rho : G \rightarrow \text{End}(\mathbb{C}^2)$  is the given representation on spin- $\frac{1}{2}$  module, and  $\rho_k(g)$  denotes its action on the  $k$ -th site.

Our goal is to reinterpret such quantum chains in a way that parallels the transfer-matrix formalism for the classical eight-vertex model. There, the partition function on a two-dimensional lattice was built by stacking row-to-row transfer matrices, each acting on a horizontal tensor product of spins. For a one-dimensional quantum chain we recover a similar picture by *enlarging* the Hilbert space with an auxiliary “horizontal” degree of freedom.

More precisely, we introduce an auxiliary space  $V_{\text{aux}}$  (typically another copy of  $\mathbb{C}^2$  carrying a representation of  $G$ ) and define local  $L$ -operators, or also known as, Lax operators:

$$L_{k,\text{aux}}(u) \in \text{End}(\mathfrak{h}_k \otimes V_{\text{aux}}), \quad k = 1, \dots, N,$$

depending on a spectral parameter  $u$ . In the 8-vertex model,  $L$  operators were parametrized Boltzmann weights and for the spin chain construction we will also assume some form of parametrization. These operators play the role of local vertex weights: they act on a *quantum* space  $\mathfrak{h}_k$  and an *auxiliary* space  $V_{\text{aux}}$ . By composing them along the chain we obtain the monodromy matrix

$$T_{\text{aux}}(u) = L_{N,\text{aux}}(u) \cdots L_{1,\text{aux}}(u)$$

and the transfer matrix

$$t(u) = \text{Tr}_{\text{aux}}(T_{\text{aux}}(u)),$$

from which we will see how the Hamiltonian and higher conserved quantities will eventually be generated.

### 2.1.1. Imposing $G$ -invariance on the Lax operator

Inspired by construction of  $L$ -operators for 8-vertex model, the key structural requirement is that the local Lax operator respects the same symmetry group  $G$  as the local Hamiltonian  $H_{k,k+1}$ . Concretely, we equip the auxiliary space with a representation

$$\rho_{\text{aux}} : G \longrightarrow \text{End}(V_{\text{aux}}),$$

and we require that  $L_{k,\text{aux}}(u)$  be  $G$ -invariant homomorphism, or equivalently an intertwiner for the diagonal  $G$ -action on  $\mathfrak{h}_k \otimes V_{\text{aux}}$ :

$$[L_{k,\text{aux}}(u), \rho_k(g) \otimes \rho_{\text{aux}}(g)] = 0, \quad \forall g \in G, \quad \forall u, \quad k = 1, \dots, N. \quad (2.1)$$

In other words, the same group  $G$  that leaves each local Hamiltonian  $H_{k,k+1}$  invariant is imposed as a symmetry of the local Lax operators. This ensures that the monodromy matrix and eventually a transfer matrix  $t(u)$  inherit the  $G$ -symmetry from the outset.

In the simplest situation we take  $V_{\text{aux}} \cong \mathbb{C}^2$  carrying the same spin- $\frac{1}{2}$  representation as each  $\mathfrak{h}_k$ . Then Schur–Weyl (or Clebsch–Gordan) theory describes explicitly how  $\mathfrak{h}_k \otimes V_{\text{aux}}$  decomposes into irreducible  $G$ -modules, and (2.1) forces  $L_{k,\text{aux}}(u)$  to be a linear combination of the corresponding projectors. We now illustrate this in two standard examples: the XXX and XXZ spin- $\frac{1}{2}$  chains.

### 2.1.2. Example 1: XXX spin- $\frac{1}{2}$ chain

The isotropic Heisenberg (XXX) Hamiltonian is

$$H^{\text{XXX}} = \sum_{k=1}^{N-1} (\sigma_k^x \sigma_{k+1}^x + \sigma_k^y \sigma_{k+1}^y + \sigma_k^z \sigma_{k+1}^z), \quad (2.2)$$

where  $\sigma_k^{x,y,z}$  are Pauli matrices acting on the  $k$ -th site. This model has full spin-rotation symmetry: its symmetry group is

$$G^{\text{XXX}} = SU(2),$$

acting on each local space  $\mathfrak{h}_k \cong \mathbb{C}^2$  via the defining spin- $\frac{1}{2}$  representation. The local Hamiltonian

$$H_{k,k+1}^{\text{XXX}} = \sigma_k^x \sigma_{k+1}^x + \sigma_k^y \sigma_{k+1}^y + \sigma_k^z \sigma_{k+1}^z$$

commutes with the diagonal  $SU(2)$ -action on  $\mathfrak{h}_k \otimes \mathfrak{h}_{k+1}$ .

For the  $L$ -operator we take  $V_{\text{aux}} \cong \mathbb{C}^2$  with the same  $SU(2)$  representation, and we impose (2.1) with  $G = SU(2)$ . The tensor product of two spin- $\frac{1}{2}$  representations decomposes as

$$\mathfrak{h}_k \otimes V_{\text{aux}} \cong V_0 \oplus V_1,$$

where

- $V_0$  is the one-dimensional singlet (total spin 0),
- $V_1$  is the three-dimensional triplet (total spin 1).

Let  $P_{\text{sing}}$  and  $P_{\text{trip}}$  be the projectors onto  $V_0$  and  $V_1$ . By Schur's lemma, any operator commuting with the diagonal  $SU(2)$ -action must act as scalar multiplication on each irreducible component. Hence

$$L_{k,\text{aux}}^{\text{XXX}}(u) = \rho_0(u) P_{\text{sing}} + \rho_1(u) P_{\text{trip}}, \quad (2.3)$$

for some scalar functions  $\rho_0(u), \rho_1(u)$ .

**Relating the matrix form of  $L^{\text{XXX}}$  to the projectors.** Starting from the standard basis

$$\{ |\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle \}.$$

The singlet and triplet subspaces are

$$V_0 = \text{span} \left\{ |s\rangle := \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \right\},$$

$$V_1 = \text{span} \left\{ |\uparrow\uparrow\rangle, \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), |\downarrow\downarrow\rangle \right\}.$$

The corresponding projectors are

$$P_{\text{sing}} = |s\rangle\langle s|, \quad P_{\text{trip}} = \mathbf{1} - P_{\text{sing}}.$$

Writing these in the above basis and substituting into (??), one finds

$$L_{k,\text{aux}}^{\text{XXX}}(u) = \begin{pmatrix} \rho_1(u) & 0 & 0 & 0 \\ 0 & \frac{\rho_0(u)+\rho_1(u)}{2} & \frac{\rho_1(u)-\rho_0(u)}{2} & 0 \\ 0 & \frac{\rho_1(u)-\rho_0(u)}{2} & \frac{\rho_0(u)+\rho_1(u)}{2} & 0 \\ 0 & 0 & 0 & \rho_1(u) \end{pmatrix}.$$

Thus, in the schematic notation

$$L_{k,\text{aux}}^{\text{XXX}}(u) = \begin{pmatrix} a(u) & 0 & 0 & d(u) \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ d(u) & 0 & 0 & a(u) \end{pmatrix},$$

SU(2)-invariance together with the singlet-triplet decomposition forces

$$a(u) = \rho_1(u), \quad d(u) = 0, \quad b(u) = \frac{\rho_0(u) + \rho_1(u)}{2}, \quad c(u) = \frac{\rho_1(u) - \rho_0(u)}{2}.$$

In particular, the off-diagonal corner elements must vanish ( $d(u) = 0$ ), and only two scalar functions  $\rho_0(u), \rho_1(u)$  remain as independent data.

### 2.1.3. Example 2: XXZ spin- $\frac{1}{2}$ chain

The anisotropic XXZ Hamiltonian is

$$H^{\text{XXZ}} = \sum_{k=1}^{N-1} \left( \sigma_k^x \sigma_{k+1}^x + \sigma_k^y \sigma_{k+1}^y + \Delta \sigma_k^z \sigma_{k+1}^z \right), \quad (2.4)$$

where  $\Delta$  is the anisotropy parameter. In this case the full SU(2) symmetry of the XXX chain is broken; the Hamiltonian conserves only the  $z$ -component of total spin.

**Step 1: the obvious  $U(1)$  symmetry.** The first symmetry group we examine is

$$G_{\text{vis}}^{\text{XXZ}} = U(1),$$

generated by rotations around the  $z$ -axis. The local representation of  $U(1)$  on  $\mathfrak{h}_k \cong \mathbb{C}^2$  is given by

$$\rho(e^{i\theta}) = \exp\left(i\frac{\theta}{2} \sigma^z\right), \quad \theta \in \mathbb{R},$$

so that  $|\uparrow\rangle$  and  $|\downarrow\rangle$  have weights  $+\frac{1}{2}$  and  $-\frac{1}{2}$ , respectively. The local Hamiltonian  $H_{k,k+1}^{\text{XXZ}}$  commutes with the *diagonal*  $U(1)$ -action on  $\mathfrak{h}_k \otimes \mathfrak{h}_{k+1}$ , i.e.

$$[H_{k,k+1}^{\text{XXZ}}, S_k^z + S_{k+1}^z] = 0.$$

We again take  $V_{\text{aux}} \cong \mathbb{C}^2$  with the same  $U(1)$  representation and impose the  $G$ -invariance condition (2.1) with  $G = U(1)$ . The tensor product decomposes into weight spaces of the *total*  $z$ -spin:

$$\mathfrak{h}_k \otimes V_{\text{aux}} = V_{+1} \oplus V_0 \oplus V_{-1},$$

where

$$\begin{aligned} V_{+1} &= \text{span}\{|\uparrow\uparrow\rangle\}, \\ V_0 &= \text{span}\{|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle\}, \\ V_{-1} &= \text{span}\{|\downarrow\downarrow\rangle\}. \end{aligned}$$

Let  $P_{+1}$ ,  $P_0$  and  $P_{-1}$  be the projectors onto these weight spaces. Any operator commuting with the diagonal  $U(1)$ -action must preserve each total-weight subspace, so Schur–Weyl theory gives a *block* decomposition

$$L_{k,\text{aux}}^{\text{XXZ}}(u) = \lambda_+(u) P_{+1} + \Lambda_0(u) + \lambda_-(u) P_{-1},$$

where

$$\Lambda_0(u) \in \text{End}(V_0)$$

is a  $2 \times 2$  matrix acting within the total-weight 0 subspace, and  $\lambda_{\pm}(u)$  are scalar functions on  $V_{\pm 1}$ .

In the  $\sigma^z$ -basis

$$\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\},$$

this  $U(1)$ -invariant structure becomes

$$L_{k,\text{aux}}^{\text{XXZ}}(u) = \begin{pmatrix} a(u) & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c'(u) & b'(u) & 0 \\ 0 & 0 & 0 & d(u) \end{pmatrix},$$

with  $a(u) = \lambda_+(u)$ ,  $d(u) = \lambda_-(u)$  and the  $2 \times 2$  middle block forming  $\Lambda_0(u)$ . At the level of  $U(1)$  symmetry alone, the four entries of this middle block are *a priori* independent; In this case the symmetry only forces a block-diagonal structure.

## Step 2: combining discrete and continuous symmetries.

In the XXZ case the local Hilbert space at a bond is

$$V := \mathbb{C}^2 \otimes \mathbb{C}^2 = \text{span}\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}.$$

We have already seen that the local Hamiltonian

$$H_{12}^{\text{XXZ}} = J_x \sigma_1^x \sigma_2^x + J_y \sigma_1^y \sigma_2^y + J_z \sigma_1^z \sigma_2^z = \begin{pmatrix} J_z & 0 & 0 & J_x - J_y \\ 0 & -J_z & J_x + J_y & 0 \\ 0 & J_x + J_y & -J_z & 0 \\ J_x - J_y & 0 & 0 & J_z \end{pmatrix}, \quad J_x = J_y,$$

has a matrix form similar to the 8-vertex  $L$ -matrix and therefore can be written in terms of the operators

$$A_1 = \sigma_1^z \sigma_2^z, \quad A_2 = M_{-+} + M_{+-}, \quad A_3 = M_{--} + M_{++},$$

so that

$$H_{12}^{\text{XXZ}} = J_z A_1 + 2J_x A_2.$$

From the 8-vertex exposition we know that on this space the Hamiltonian is invariant under action of the Klein group representation

$$G := \langle g_z, g_x \mid g_z^2 = g_x^2 = e, g_z g_x = g_x g_z \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2,$$

given by

$$\rho(g_z) = A_1 = \sigma_1^z \sigma_2^z, \quad \rho(g_x) = F := \sigma_1^x \sigma_2^x.$$

We already have checked directly that  $H_{12}^{\text{XXZ}}$  commutes with both  $A_1$  and  $F$ , hence with the whole image of  $G$ :

$$[H_{12}^{\text{XXZ}}, \rho(g)] = 0 \quad \text{for all } g \in G.$$

Independently of this discrete symmetry, we know that XXZ is partially isotropic:

$$[S^z, H] = 0, \quad S^z = \sum_{n=1}^N S_n^z,$$

so the local density commutes with the two-site total spin

$$S_{\text{tot}}^z := S_1^z + S_2^z = \frac{1}{2}(\sigma_1^z + \sigma_2^z).$$

This yields a continuous  $U(1)$ -symmetry of the local Hamiltonian, implemented by

$$U(\phi) := \exp(i\phi S_{\text{tot}}^z), \quad U(\phi) H_{12}^{\text{XXZ}} U(\phi)^{-1} = H_{12}^{\text{XXZ}}.$$

We now combine the discrete  $G$ -symmetry and the  $U(1)$ -symmetry into a single group acting on  $V$ . The key observation is how  $G$  acts by conjugation on the  $U(1)$  rotations.

- The element  $A_1 = \sigma_1^z \sigma_2^z$  commutes with  $S_{\text{tot}}^z$  (both are diagonal in the computational basis), so

$$A_1 U(\phi) A_1^{-1} = U(\phi).$$

- The element  $F = \sigma_1^x \sigma_2^x$  implements a  $\pi$ -rotation around the  $x$ -axis on each spin. This flips the sign of  $S^z$  and hence of  $S_{\text{tot}}^z$ :

$$F S_{\text{tot}}^z F^{-1} = -S_{\text{tot}}^z,$$

and therefore

$$F U(\phi) F^{-1} = U(-\phi).$$

Thus the subgroup  $\langle F \rangle \cong \mathbb{Z}_2$  acts on  $U(1)$  by inversion  $\phi \mapsto -\phi$ , whereas the subgroup  $\langle A_1 \rangle \cong \mathbb{Z}_2$  acts trivially. This leads to a semidirect product description of the combined local symmetry group.

Let us write

$$H := \langle F \rangle \cong \mathbb{Z}_2, \quad Z := \langle A_1 \rangle \cong \mathbb{Z}_2.$$

Define a homomorphism

$$\alpha : H \longrightarrow \text{Aut}(U(1))$$

by

$$\alpha(e)(e^{i\phi}) = e^{i\phi}, \quad \alpha(F)(e^{i\phi}) = e^{-i\phi}.$$

Then the semidirect product  $U(1) \rtimes_{\alpha} H$  has underlying set  $U(1) \times H$  and group law

$$(e^{i\phi_1}, h_1) \cdot (e^{i\phi_2}, h_2) = (e^{i\phi_1} \alpha(h_1)(e^{i\phi_2}), h_1 h_2) = (e^{i(\phi_1 + \varepsilon_1 \phi_2)}, h_1 h_2),$$

where  $\varepsilon_1 = +1$  if  $h_1 = e$  and  $\varepsilon_1 = -1$  if  $h_1 = F$ .

The full local symmetry group of the XXZ Hamiltonian density can then be written as

$$G_{\text{ext}} := (U(1) \rtimes_{\alpha} H) \times Z \cong O(2) \times \mathbb{Z}_2.$$

The action of  $(e^{i\phi}, h, z) \in G_{\text{ext}}$  on  $V$  is given by

$$(e^{i\phi}, h, z) \longmapsto \pi(e^{i\phi}, h, z) := U(\phi) \rho(hz),$$

and by construction  $H_{12}^{\text{XXZ}}$  commutes with the image of  $G_{\text{ext}}$ :

$$[H_{12}^{\text{XXZ}}, \pi(g)] = 0 \quad \text{for all } g \in G_{\text{ext}}.$$

In this sense the XXZ local Hamiltonian is invariant under the combined action of the discrete Klein group and the continuous  $U(1)$ -symmetry.

**Constraints on the Lax operator from  $G_{\text{ext}}$ -invariance.**

We now apply this enlarged symmetry to the Lax operator. Let  $V_{\text{aux}}$  denote the two-dimensional auxiliary space and  $V_n$  the two-dimensional quantum space at site  $n$ . For the XXZ six-vertex model we take both to carry the same  $G_{\text{ext}}$ -module structure as above, so that

$$(V_{\text{aux}}, \pi_{\text{aux}}) \cong (V_n, \pi_n) \cong (V, \pi).$$

The Lax operator is an operator

$$L(u) \in \text{End}(V_{\text{aux}} \otimes V_n)$$

depending on a spectral parameter  $u$ , and we require it to be  $G_{\text{ext}}$ -invariant, i.e.

$$(\pi_{\text{aux}}(g) \otimes \pi_n(g)) L(u) (\pi_{\text{aux}}(g) \otimes \pi_n(g))^{-1} = L(u) \quad \text{for all } g \in G_{\text{ext}}. \quad (2.5)$$

First, invariance under the  $U(1)$  subgroup forces  $L(u)$  to preserve the eigenspaces of the total  $S^z$  on  $V_{\text{aux}} \otimes V_n$ . In the canonical basis

$$|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$$

these are the subspaces

$$V_{+2} = \text{span}\{|\uparrow\uparrow\rangle\}, \quad V_0 = \text{span}\{|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle\}, \quad V_{-2} = \text{span}\{|\downarrow\downarrow\rangle\}.$$

Hence  $L(u)$  must be block-diagonal with respect to this decomposition; in coordinates this means

$$L(u) = \begin{pmatrix} a_+(u) & 0 & 0 & 0 \\ 0 & b_{11}(u) & b_{12}(u) & 0 \\ 0 & b_{21}(u) & b_{22}(u) & 0 \\ 0 & 0 & 0 & a_-(u) \end{pmatrix}.$$

Next, we use the discrete part of  $G_{\text{ext}}$ . The operator  $F = \sigma^x \otimes \sigma^x$  interchanges

$$F : V_{+2} \longleftrightarrow V_{-2}, \quad F : |\uparrow\downarrow\rangle \longleftrightarrow |\downarrow\uparrow\rangle.$$

Condition (2.5) for  $g = g_x$  is therefore precisely the statement that  $L(u)$  is an intertwiner for the  $G$ -action on  $V_{\text{aux}} \otimes V_n$ :

$$L(u) (\pi_{\text{aux}} \otimes \pi_n)(g_x) = (\pi_{\text{aux}} \otimes \pi_n)(g_x) L(u).$$

By Schur–Weyl duality, the commutant of the  $G$ -action on  $V \otimes V$  is generated by the algebra  $\mathcal{A} = \text{span}\{I, A_1, A_2, A_3\}$ , and in particular the restriction of  $L(u)$  to each isotypic component of the  $G$ -module  $V \otimes V$  must be a linear combination of the identity and the permutations between equivalent irreducible subrepresentations. Concretely:

- The spaces  $V_{+2}$  and  $V_{-2}$  carry equivalent one-dimensional representations of  $G$  and are interchanged by  $F$ . By Schur's lemma, the restriction of  $L(u)$  to  $V_{+2} \oplus V_{-2}$  must be proportional to the identity in an  $F$ -invariant basis. In our matrix form this forces

$$a_+(u) = a_-(u) =: a(u).$$

- On the zero-magnetisation subspace  $V_0$  the action of  $G$  is generated by the flip  $|\uparrow\downarrow\rangle \leftrightarrow |\downarrow\uparrow\rangle$ . Again, the algebra of intertwiners is spanned by the identity and this flip, so the restriction of  $L(u)$  to  $V_0$  must be of the form

$$L(u)|_{V_0} = b(u) \text{id}_{V_0} + c(u) P_0,$$

where  $P_0$  swaps the two basis vectors. In matrix form this gives

$$\begin{pmatrix} b_{11}(u) & b_{12}(u) \\ b_{21}(u) & b_{22}(u) \end{pmatrix} = \begin{pmatrix} b(u) & c(u) \\ c(u) & b(u) \end{pmatrix}.$$

Putting these constraints together, we conclude that any Lax operator  $L(u) \in \text{End}(V_{\text{aux}} \otimes V_n)$  satisfying the  $G_{\text{ext}}$ -invariance condition (2.5) must have the form

$$L(u) = \begin{pmatrix} a(u) & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & a(u) \end{pmatrix}.$$

This is exactly the usual  $L$ -matrix (or Boltzmann weight matrix) of the six-vertex model, with three independent functions  $a(u), b(u), c(u)$ .

Thus, by enlarging the symmetry from the pure  $U(1)$  generated by  $S^z$  to the full group  $G_{\text{ext}}$ , we significantly restrict the possible form of the Lax operator: instead of a generic  $4 \times 4$  matrix compatible only with charge conservation, we are left with the three-parameter six-vertex ansatz. In the next step we will see that the XXZ model enjoys an even richer hidden symmetry, described by a deformed Hopf algebra, which further constrains the dependence of  $a(u), b(u), c(u)$  on the spectral parameter.

**Step 3: upgrading to the full quantum-group symmetry.** If we are determined to dig truly deep then we find that the actual symmetry of the XXZ Hamiltonian is described by a *larger* algebraic object: a  $q$ -deformed Hopf algebra  $U_q(\mathfrak{sl}_2)$  (a quantum group), which is a deformation of the universal enveloping algebra  $U(\mathfrak{sl}_2)$  and underlies the integrability of the model.<sup>1</sup> The  $U(1)$  symmetry generated by  $S^z$  can be viewed as the “Cartan” part of this quantum group;  $U_q(\mathfrak{sl}_2)$  contains additional generators (raising and lowering operators) together with a compatible coproduct.

On the local two-site space one considers the tensor product

$$\mathfrak{h}_k \otimes V_{\text{aux}} \cong \mathbb{C}^2 \otimes \mathbb{C}^2,$$

where each factor carries the fundamental representation of  $U_q(\mathfrak{sl}_2)$ . For generic  $q$  the representation theory of  $U_q(\mathfrak{sl}_2)$  is a  $q$ -deformation of that of  $\mathfrak{sl}_2$ , and the Clebsch–Gordan rule becomes

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \cong V_1 \oplus V_0,$$

where  $V_1$  is the 3-dimensional “spin-1” (triplet) representation and  $V_0$  is the 1-dimensional “spin-0” (singlet) representation. In particular, the  $U(1)$  weight-zero subspace  $V_0^{(\text{weight})} = \text{span}\{|\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle\}$  is *reducible* under  $U_q(\mathfrak{sl}_2)$  and splits into a  $q$ -deformed singlet plus the zero-weight component of the triplet.

Let  $P_1$  and  $P_0$  be the projectors onto  $V_1$  and  $V_0$ , respectively. Since the decomposition

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \cong V_1 \oplus V_0$$

has no multiplicities, Schur–Weyl theory now tells us that the commutant

$$\text{End}_{U_q(\mathfrak{sl}_2)}(\mathbb{C}^2 \otimes \mathbb{C}^2)$$

is 2-dimensional, spanned by  $P_1$  and  $P_0$ . In other words, any  $U_q(\mathfrak{sl}_2)$ -invariant  $L$ -operator must be of the form

$$L_{k,\text{aux}}^{(q)}(u) = \alpha(u) P_1 + \beta(u) P_0, \tag{2.6}$$

for some scalar functions  $\alpha(u)$  and  $\beta(u)$ . Compared with the purely  $U(1)$ -invariant case, the larger symmetry group  $U_q(\mathfrak{sl}_2)$  *reduces* the freedom in  $L$ : the general  $4 \times 4$  block form with six scalar functions collapses to a two-parameter family determined by the projectors onto the irreducible quantum-group components.

**Explicit matrix form.** To make (2.6) concrete, fix the standard basis

$$\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$$

---

<sup>1</sup>We will not need the explicit definition of  $U_q(\mathfrak{sl}_2)$  here; see, for example, Chari–Pressley, *A Guide to Quantum Groups*, Cambridge Univ. Press (1994).

of  $\mathbb{C}^2 \otimes \mathbb{C}^2$ , and let  $q$  be the deformation parameter. A convenient choice of  $q$ -deformed singlet vector is

$$v_{\text{sing}} = \frac{1}{\sqrt{q+q^{-1}}} (\sqrt{q} |\uparrow\downarrow\rangle - q^{-1/2} |\downarrow\uparrow\rangle),$$

so that the projector onto the singlet is

$$P_0 = |v_{\text{sing}}\rangle\langle v_{\text{sing}}|.$$

In the chosen basis this has the matrix

$$P_0 = \frac{1}{q+q^{-1}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & q & -1 & 0 \\ 0 & -1 & q^{-1} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The triplet projector is  $P_1 = \mathbb{I}_4 - P_0$ . Substituting into (2.6) gives

$$L_{k,\text{aux}}^{(q)}(u) = \alpha(u) \mathbb{I}_4 + (\beta(u) - \alpha(u)) P_0,$$

i.e.

$$L_{k,\text{aux}}^{(q)}(u) = \begin{pmatrix} \alpha(u) & 0 & 0 & 0 \\ 0 & \alpha(u) + \frac{q}{q+q^{-1}}(\beta(u) - \alpha(u)) & -\frac{1}{q+q^{-1}}(\beta(u) - \alpha(u)) & 0 \\ 0 & -\frac{1}{q+q^{-1}}(\beta(u) - \alpha(u)) & \alpha(u) + \frac{q^{-1}}{q+q^{-1}}(\beta(u) - \alpha(u)) & 0 \\ 0 & 0 & 0 & \alpha(u) \end{pmatrix}.$$

In a basis adapted to the decomposition  $V_1 \oplus V_0$ , this  $L$ -operator is simply

$$L_{k,\text{aux}}^{(q)}(u) \cong \text{diag}(\alpha(u), \alpha(u), \alpha(u), \beta(u)),$$

with  $\alpha(u)$  acting on the triplet and  $\beta(u)$  on the singlet. This illustrates the general philosophy of this section: by enlarging the symmetry group from the abelian  $U(1)$  to the full quantum group  $U_q(\mathfrak{sl}_2)$ , the tensor product  $\mathbb{C}^2 \otimes \mathbb{C}^2$  becomes more highly reducible, and the corresponding  $L$ -operator is correspondingly more tightly constrained (a two-parameter family instead of a generic  $U(1)$ -invariant block matrix). Standard references for the representation theory of  $U_q(\mathfrak{sl}_2)$  and the decomposition  $\mathbb{C}^2 \otimes \mathbb{C}^2 \cong V_1 \oplus V_0$  include Chari–Pressley, *A Guide to Quantum Groups*, and lecture notes on quantum groups and  $R$ -matrices such as Losev *Representations of  $U_q(\mathfrak{g})$  and  $R$ -matrices*.

**Remark 1** (Changing the spin representation). *In both examples above we worked with  $\text{spin}-\frac{1}{2}$  local spaces, i.e. with the fundamental representation of the corresponding symmetry group:*

$$\mathfrak{h}_k \cong V_{\frac{1}{2}} \quad \text{for } G^{\text{XXX}} = SU(2), \quad \mathfrak{h}_k \cong V_{\frac{1}{2}} \quad \text{for } G^{\text{XXZ}} = U(1).$$

*If instead we choose a higher spin (for instance, a  $\text{spin}-1$  chain for XXX or XXZ), the local Hilbert space becomes a higher-dimensional irreducible representation  $V_s$  of the same symmetry group  $G$ , and the relevant tensor products are larger:*

$$V_{\text{aux}} \otimes V_s \cong \bigoplus_{\lambda} V_{\lambda},$$

*with more irreducible components  $V_{\lambda}$ . The  $G$ -invariance condition (2.1) still holds, but the  $L$ -operator is now expressed as a sum of projectors onto these components:*

$$L_{k,\text{aux}}(u) = \sum_{\lambda} \rho_{\lambda}(u) P_{\lambda},$$

*where  $P_{\lambda}$  projects onto  $V_{\lambda}$ . In other words, changing the spin does not change the symmetry group of the model; it changes the representation in which the model is realized, and correspondingly enlarges the decomposition of  $\mathfrak{h}_k \otimes V_{\text{aux}}$  and the list of projectors appearing in the  $L$ -operator.*

## 2.2. The $R$ -matrix, transfer matrices, and the Yang–Baxter equation

In the transfer-matrix formulation of quantum integrable systems, the central object is a one-parameter family of transfer matrices

$$t(u) \in \text{End}(\mathcal{H}), \quad u \in \mathbb{C},$$

acting on the quantum space  $\mathcal{H}$  of the spin chain. The key structural property we want is the *commutativity* of this family:

$$[t(u), t(v)] = 0 \quad \text{for all } u, v. \quad (2.7)$$

This commutativity implies the existence of an infinite family of conserved quantities obtained, for example, by expanding  $\log t(u)$  in the spectral parameter  $u$ . Thus:

*Our main goal is to identify a parametrized commuting family of transfer matrices  $\{t(u)\}_{u \in \mathbb{C}}$ .*

In this section we show that the *existence of a suitable  $R$ -matrix* is a *sufficient condition* for the commutativity (2.7), and that consistency of this construction naturally leads to the Yang–Baxter equation.

### 2.2.1. Coming back to transfer matrix formalism

Let  $\mathcal{H} = V_1 \otimes \cdots \otimes V_N$  be the quantum space of a spin chain with  $N$  sites, and let  $V_a$  be an auxiliary space. A Lax operator is a family of operators

$$L_{aj}(u) \in \text{End}(V_a \otimes V_j), \quad j = 1, \dots, N,$$

depending meromorphically on a complex spectral parameter  $u$ . The *monodromy matrix* is defined by the ordered product

$$T_a(u) = L_{aN}(u) L_{a,N-1}(u) \cdots L_{a1}(u) \in \text{End}(V_a \otimes \mathcal{H}),$$

where the label  $a$  indicates that  $T_a(u)$  acts non-trivially on the auxiliary space  $V_a$  and on the chain  $\mathcal{H}$ , and as the identity on any other auxiliary spaces.

The *transfer matrix* is defined as the partial trace over the auxiliary space:

$$t(u) = \text{tr}_a(T_a(u)) \in \text{End}(\mathcal{H}).$$

The problem is now to determine conditions on the Lax operators  $L_{aj}(u)$  under which the transfer matrices  $t(u)$  commute for different values of  $u$ .

### 2.2.2. The $R$ -matrix as a sufficient condition for commutativity

To formulate a sufficient condition for (2.7), consider two *distinct* auxiliary spaces  $V_a$  and  $V_b$ , and introduce an operator

$$R_{ab}(u, v) \in \text{End}(V_a \otimes V_b),$$

depending on two spectral parameters  $u, v$ . We assume that  $R_{ab}(u, v)$  satisfies the *local intertwining relation* (often called the *RLL* relation)

$$R_{ab}(u, v) L_{aj}(u) L_{bj}(v) = L_{bj}(v) L_{aj}(u) R_{ab}(u, v), \quad j = 1, \dots, N. \quad (2.8)$$

Note that each  $L_{aj}(u)$  acts non-trivially only on  $V_a \otimes V_j$ , and each  $L_{bj}(v)$  only on  $V_b \otimes V_j$ , while  $R_{ab}(u, v)$  acts only on  $V_a \otimes V_b$ . In particular, all these operators act as the identity on tensor factors where they are not explicitly involved.

**Proposition 1** (Existence of  $R$  is sufficient for commutativity). *If there exists  $R_{ab}(u, v)$  satisfying (2.8) for all sites  $j = 1, \dots, N$ , then the transfer matrices  $t(u)$  and  $t(v)$  commute:*

$$[t(u), t(v)] = 0.$$

*Proof.* Define monodromy matrices  $T_a(u)$  and  $T_b(v)$  in the auxiliary spaces  $V_a$  and  $V_b$ :

$$T_a(u) = L_{aN}(u) \cdots L_{a1}(u), \quad T_b(v) = L_{bN}(v) \cdots L_{b1}(v).$$

Using the local relation (2.8) at each site  $j$  and the fact that  $R_{ab}(u, v)$  commutes with  $L_{ak}(u)$  and  $L_{bk}(v)$  for  $k \neq j$  (since they act on different factors), we obtain

$$\begin{aligned} R_{ab}(u, v) T_a(u) T_b(v) &= R_{ab}(u, v) (L_{aN}(u) \cdots L_{a1}(u)) (L_{bN}(v) \cdots L_{b1}(v)) \\ &= R_{ab}(u, v) (L_{aN} L_{bN}) \cdots (L_{a1} L_{b1}) \\ &= (L_{bN} L_{aN}) \cdots (L_{b1} L_{a1}) R_{ab}(u, v) \\ &= T_b(v) T_a(u) R_{ab}(u, v), \end{aligned}$$

where we have suppressed the spectral parameters in the intermediate steps for readability. Thus

$$R_{ab}(u, v) T_a(u) T_b(v) = T_b(v) T_a(u) R_{ab}(u, v). \quad (2.9)$$

Now take the trace over  $V_a \otimes V_b$  and use cyclicity of the trace:

$$\begin{aligned} \text{tr}_{a,b}(T_a(u) T_b(v)) &= \text{tr}_{a,b}(R_{ab}(u, v) T_a(u) T_b(v) R_{ab}(u, v)^{-1}) \\ &= \text{tr}_{a,b}(T_b(v) T_a(u)). \end{aligned}$$

Factoring the trace over the two auxiliary spaces yields

$$\text{tr}_a(T_a(u)) \text{tr}_b(T_b(v)) = \text{tr}_b(T_b(v)) \text{tr}_a(T_a(u)),$$

that is,

$$t(u) t(v) = t(v) t(u).$$

This proves the commutativity of the family  $\{t(u)\}$ . □

**Remark 2.** We emphasize that:

1. The main structural interest lies in the commuting one-parameter family  $\{t(u)\}$ , which encodes infinitely many conserved quantities.
2. The existence of an operator  $R_{ab}(u, v)$  satisfying (2.8) is a sufficient condition for this commutativity. It need not be unique, and, in general, it is not obvious that it is a necessary condition.

### 2.2.3. Consistency and the Yang–Baxter equation

So far we have seen that any  $R_{ab}(u, v)$  satisfying the local relation (2.8) provides a sufficient mechanism for constructing a commuting family  $\{t(u)\}$ . However, if we want this construction to be *consistent* when we consider *three* different auxiliary spaces, we obtain an additional constraint on  $R$ , namely the Yang–Baxter equation.

Consider three auxiliary spaces  $V_1, V_2, V_3$  with monodromy matrices

$$T_1(u), \quad T_2(v), \quad T_3(w),$$

and suppose we have  $R$ -matrices

$$R_{12}(u, v), \quad R_{13}(u, w), \quad R_{23}(v, w) \in \text{End}(V_i \otimes V_j)$$

satisfying the corresponding local relations

$$R_{ij}(u_i, u_j) L_{ik}(u_i) L_{jk}(u_j) = L_{jk}(u_j) L_{ik}(u_i) R_{ij}(u_i, u_j)$$

for  $(i, j) = (1, 2), (1, 3), (2, 3)$  and all sites  $k$ . As above, these imply global intertwining relations

$$\begin{aligned} R_{12}(u, v) T_1(u) T_2(v) &= T_2(v) T_1(u) R_{12}(u, v), \\ R_{13}(u, w) T_1(u) T_3(w) &= T_3(w) T_1(u) R_{13}(u, w), \\ R_{23}(v, w) T_2(v) T_3(w) &= T_3(w) T_2(v) R_{23}(v, w), \end{aligned}$$

with each  $R_{ij}$  commuting with monodromies  $T_k$  for  $k \neq i, j$ .

We now consider two different ways of reordering the product  $T_1(u)T_2(v)T_3(w)$  into  $T_3(w)T_2(v)T_1(u)$  using the  $R$ -matrices.

**First route.** Start from

$$T_1 T_2 T_3,$$

and interchange  $T_1$  and  $T_2$  using  $R_{12}$ , then  $T_1$  and  $T_3$  using  $R_{13}$ , and finally  $T_2$  and  $T_3$  using  $R_{23}$ :

$$\begin{aligned} R_{23} R_{13} R_{12} T_1 T_2 T_3 &= R_{23} R_{13} T_2 T_1 T_3 R_{12} \\ &= R_{23} T_2 T_3 T_1 R_{13} R_{12} \\ &= T_3 T_2 T_1 R_{23} R_{13} R_{12}. \end{aligned}$$

**Second route.** Alternatively, we can first interchange  $T_2$  and  $T_3$  (using  $R_{23}$ ), then  $T_1$  and  $T_3$  (using  $R_{13}$ ), and finally  $T_1$  and  $T_2$  (using  $R_{12}$ ). This yields

$$\begin{aligned} R_{12} R_{13} R_{23} T_1 T_2 T_3 &= R_{12} R_{13} T_1 T_3 T_2 R_{23} \\ &= R_{12} T_3 T_1 T_2 R_{13} R_{23} \\ &= T_3 T_2 T_1 R_{12} R_{13} R_{23}. \end{aligned}$$

Since both routes must implement the same transformation  $T_1 T_2 T_3 \mapsto T_3 T_2 T_1$  on the tensor product  $V_1 \otimes V_2 \otimes V_3 \otimes \mathcal{H}$ , the two intertwiners must coincide:

$$R_{23}(v, w) R_{13}(u, w) R_{12}(u, v) = R_{12}(u, v) R_{13}(u, w) R_{23}(v, w).$$

This is precisely the *Yang–Baxter equation*:

$$R_{12}(u, v) R_{13}(u, w) R_{23}(v, w) = R_{23}(v, w) R_{13}(u, w) R_{12}(u, v), \quad (2.10)$$

written as an identity in  $\text{End}(V_1 \otimes V_2 \otimes V_3)$ .

From this perspective, the Yang–Baxter equation expresses the *consistency of pairwise intertwiners*  $R_{ij}(u_i, u_j)$  when we attempt to reorder products of monodromy matrices with more than two auxiliary spaces. Once again, the conceptual chain is:

$R$  solves the Yang–Baxter equation (2.10)

$\Downarrow$

$R$  defines a consistent system of local relations (2.8)

$\Downarrow$

$R$  provides a sufficient mechanism for constructing a commuting family of transfer matrices  $\{t(u)\}$ .

In summary, the Yang–Baxter equation appears as the structural condition ensuring that the sufficient condition for commutativity (the existence of  $R$  satisfying the local *RLL* relations) extends consistently to arbitrarily many auxiliary spaces.

## 2.3. Deriving conserved quantities of the Hamiltonian from the transfer matrix

After introducing the Yang–Baxter equation and the *RLL* relations, we now explain how one uses the resulting family of commuting transfer matrices to generate an infinite sequence of conserved quantities for a quantum spin chain Hamiltonian. We focus on the XXZ spin- $\frac{1}{2}$  chain with periodic boundary conditions, whose Hamiltonian is

$$\mathcal{H} = -\frac{J}{2} \sum_{i=1}^N (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z), \quad \sigma_{N+1}^{x,y,z} = \sigma_1^{x,y,z}, \quad J > 0.$$

The Hilbert space of the model is the  $N$ -fold tensor product

$$\mathcal{H}_N = \bigotimes_{i=1}^N \mathfrak{h}_i, \quad \mathfrak{h}_i \cong \mathbb{C}^2 \quad \forall i.$$

As discussed before, we impose the six-vertex (XXZ) symmetry by requiring the Lax operators to be invariant under the local symmetry group of the XXZ Hamiltonian, which we can write as

$$G_{\text{ext}} := (U(1) \rtimes_{\alpha} H) \times Z \cong O(2) \times \mathbb{Z}_2.$$

This symmetry requirement determines the form of the Lax operator up to scalar functions of the spectral parameter and ensures that the algebraic construction of integrability is compatible with the physical symmetries of the model.

**Remark 3.** *Given the Hilbert space for the spin chain with  $N$  sites,  $\mathcal{H} = V^{\otimes N}$ , and an operator  $A \in \text{End}(V)$ , we use the standard notation*

$$A_i := \mathbb{I}_1 \otimes \cdots \otimes \mathbb{I}_{i-1} \otimes A \otimes \mathbb{I}_{i+1} \otimes \cdots \otimes \mathbb{I}_N,$$

*to indicate that  $A$  acts non-trivially only on the  $i$ -th factor and as the identity on all other tensor factors.*

For the XXZ / six-vertex model, the Lax operator acting on a quantum space  $V \cong \mathbb{C}^2$  and an auxiliary space  $V_{\text{aux}} \cong \mathbb{C}^2$  has the form

$$L(u) = \begin{pmatrix} a(u) & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & a(u) \end{pmatrix},$$

written in the basis  $|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$ . For a fixed value of the spectral parameter  $u$ , this is equivalently characterized by a triple of Boltzmann weights  $(a, b, c) := (a(u), b(u), c(u))$ .

The monodromy matrix on a chain of length  $N$  is

$$T_N(u) = L_{N,a}(u) L_{N-1,a}(u) \cdots L_{1,a}(u),$$

and the associated transfer matrix is defined by tracing over the auxiliary space:

$$t_N(u) := \text{tr}_a T_N(u).$$

By virtue of the Yang–Baxter equation, the monodromy matrices satisfy the fundamental relation

$$R_{ab}(u-v) T_a(u) T_b(v) = T_b(v) T_a(u) R_{ab}(u-v),$$

which in turn implies the commutativity of the transfer matrices for different values of the spectral parameter:

$$[t_N(u), t_N(v)] = 0 \quad \forall u, v.$$

Equivalently, instead of using  $u$  as parameter, we may label the transfer matrix by its Boltzmann weights and write  $t(a, b, c)$ .

We now specialize to the following situation. Consider two Lax operators with (possibly different) weights  $(a, b, c)$  and  $(a', b', c')$ , and suppose there exists an intertwiner  $L''_{aa'}$  (playing the role of an  $R$ -matrix between two auxiliary spaces) such that the mixed  $RLL$ -relation

$$L''_{aa'} L_{n,a'}(a', b', c') L_{n,a}(a, b, c) = L_{n,a}(a, b, c) L_{n,a'}(a', b', c') L''_{aa'}$$

holds for every site  $n$ . A direct computation (for instance, with a symbolic equation solver) shows that a non-trivial solution  $L''_{aa'}$  exists if and only if the two triples  $(a, b, c)$  and  $(a', b', c')$  share the same value of the anisotropy parameter

$$\Delta(a, b, c) := \frac{a^2 + b^2 - c^2}{2ab} = \frac{(a')^2 + (b')^2 - (c')^2}{2a'b'} = \Delta(a', b', c').$$

In other words, transfer matrices with different Boltzmann weights commute precisely when the corresponding six-vertex data describe the same XXZ anisotropy  $\Delta$ . We denote by

$$\mathcal{F} := \{ t(a, b, c) \mid \Delta(a, b, c) = \Delta_0 \}$$

the commuting family of transfer matrices with fixed anisotropy  $\Delta_0$ .

In this setting, the mixed  $RLL$ -relation above can be viewed as a system of functional equations in the Boltzmann weights  $(a, b, c)$  and  $(a', b', c')$ , together with the coefficients of the intertwiner  $L''_{aa'}$ , which are themselves parametrized by another triple of weights  $(a'', b'', c'')$  for the auxiliary-auxiliary  $R$ -matrix. These functional equations were first systematically written down and solved by Baxter in the context of the six-vertex (ice-type / Ising-type) model, and they constitute the starting point for deriving the explicit form of the  $R$ -matrix and the associated transfer matrices.

We adopt the standard hyperbolic parametrization of the six-vertex Boltzmann weights

$$a(u) = \sinh(u + \gamma), \quad b(u) = \sinh(u), \quad c(u) = \sinh(\gamma),$$

where  $\gamma$  is the anisotropy (or crossing) parameter. The quantity

$$\Delta(a, b, c) = \frac{a(u)^2 + b(u)^2 - c(u)^2}{2a(u)b(u)}$$

is independent of the spectral parameter  $u$  and evaluates to

$$\Delta(a(u), b(u), c(u)) = \cosh \gamma.$$

For real  $\gamma$  we thus have  $\Delta \geq 1$ , while for purely imaginary  $\gamma = i\mu$  one finds  $\Delta = \cos \mu \in [-1, 1]$ ; these two regimes correspond to the massive and critical phases of the XXZ Hamiltonian (up to an overall sign convention for  $J$ ).

With this parametrization, the unnormalised six-vertex  $R$ -matrix acting on  $V \otimes V$  takes the form

$$R(u) = \begin{pmatrix} \sinh(u + \gamma) & 0 & 0 & 0 \\ 0 & \sinh(u) & \sinh(\gamma) & 0 \\ 0 & \sinh(\gamma) & \sinh(u) & 0 \\ 0 & 0 & 0 & \sinh(u + \gamma) \end{pmatrix}.$$

At  $u = 0$  we have

$$R(0) = \sinh \gamma P,$$

where  $P$  is the permutation operator on  $V \otimes V$ . Rescaling by the scalar factor  $\sinh \gamma$  we obtain the normalised  $R$ -matrix

$$\widehat{R}(u) := \frac{1}{\sinh \gamma} R(u), \quad \text{so that} \quad \widehat{R}(0) = P.$$

The corresponding Lax operator is defined by

$$L_{a,k}(u) := \widehat{R}_{a,k}(u) \in \text{End}(V_a \otimes \mathfrak{h}_k),$$

and satisfies the regularity condition

$$L_{a,k}(0) = P_{a,k}.$$

Expanding around the regular point  $u = 0$  we write

$$\widehat{R}(u) = P + u \widehat{R}'(0) + O(u^2), \quad L_{a,k}(u) = P_{a,k} + u \widehat{R}'_{a,k}(0) + O(u^2).$$

In the basis

$$\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$$

one finds

$$\widehat{R}'(0) = \frac{1}{\sinh \gamma} \text{diag}(\cosh \gamma, 1, 1, \cosh \gamma) = \alpha \mathbf{1} + \beta \sigma^z \otimes \sigma^z,$$

for suitable constants  $\alpha, \beta$  depending on  $\gamma$ .

From this we obtain the explicit matrix

$$P_{k,a} \widehat{R}'_{k,a}(0) = \begin{pmatrix} \alpha + \beta & 0 & 0 & 0 \\ 0 & 0 & \alpha - \beta & 0 \\ 0 & \alpha - \beta & 0 & 0 \\ 0 & 0 & 0 & \alpha + \beta \end{pmatrix}.$$

It is sometimes convenient to separate out an arbitrary multiple  $\delta \in \mathbb{C}$  of the identity:

$$P_{k,a} \widehat{R}'_{k,a}(0) = \begin{pmatrix} \alpha + \delta + \beta & 0 & 0 & 0 \\ 0 & \delta & \alpha - \beta & 0 \\ 0 & \alpha - \beta & \delta & 0 \\ 0 & 0 & 0 & \alpha + \delta + \beta \end{pmatrix} - \delta \mathbf{1}.$$

In terms of Pauli matrices this operator can be written as

$$h_{k,a} := P_{k,a} \widehat{R}'_{k,a}(0) = J_x (\sigma_k^x \sigma_a^x + \sigma_k^y \sigma_a^y) + J_z \sigma_k^z \sigma_a^z + (\text{constant}),$$

for suitable constants  $J_x, J_z$  expressed in terms of  $\alpha, \beta, \delta$ . In particular, the derivative of the (normalised)  $R$ -matrix at the regular point, conjugated by the permutation  $P$ , has exactly the local XXZ Hamiltonian density structure.

We are now in a position to relate the family of commuting transfer matrices to the Hamiltonian. Consider the monodromy matrix

$$T_a(u) = L_{N,a}(u) \cdots L_{2,a}(u) L_{1,a}(u), \quad t(u) = \text{tr}_a T_a(u).$$

Using the expansion

$$L_{k,a}(u) = P_{k,a} + u \widehat{R}'_{k,a}(0) + O(u^2) = P_{k,a} + u h_{k,a} + O(u^2),$$

we obtain

$$T_a(u) = T_a(0) + u T'_a(0) + O(u^2),$$

with

$$T_a(0) = P_{N,a} \cdots P_{2,a} P_{1,a},$$

which is the operator that cyclically shifts the quantum spaces along the chain when traced over the auxiliary space; thus  $t(0)$  is (up to a possible scalar factor) the one-site shift operator  $U$  on  $\mathcal{H}_N$ . The derivative  $T'_a(0)$  is a sum over terms in which one  $L_{k,a}$  is replaced by  $h_{k,a}$  and all others by  $P_{j,a}$ :

$$T'_a(0) = \sum_{k=1}^N P_{N,a} \cdots P_{k+1,a} h_{k,a} P_{k-1,a} \cdots P_{1,a}.$$

Taking the trace over the auxiliary space and conjugating by  $t(0)$ , one finds that

$$t^{-1}(0) t'(0) \propto \sum_{k=1}^N h_{k,k+1},$$

where  $h_{k,k+1}$  is the local XXZ Hamiltonian density acting on sites  $k$  and  $k+1$  (with periodic identification  $N+1 \equiv 1$ ), obtained from the same matrix as  $h_{k,a}$  after identifying the auxiliary space with site  $k+1$ . In particular, up to an additive constant and an overall normalization of the spectral parameter, the physical Hamiltonian  $\mathcal{H}$  is obtained as the first logarithmic derivative of the transfer matrix at the regular point,

$$\mathcal{H} = \text{const} \cdot t^{-1}(0) t'(0) + (\text{scalar}).$$

By construction, this operator commutes with  $t(u)$  for all  $u$  and therefore belongs to the commuting family of conserved quantities generated by the transfer matrix.

### 2.3.1. Commutativity of the transfer matrix with the Hamiltonian

One last result about transfer matrix that remains to show is how it is in fact a generating function for the conserved quantities of our Hamiltonian. Let  $t(u)$  be the transfer matrix of an integrable quantum spin chain, defined as the auxiliary trace of the monodromy matrix,

$$t(u) = \text{tr}_a T_a(u),$$

and suppose that the corresponding *RLL*-relations imply the commutativity of transfer matrices for different values of the spectral parameter:

$$[t(u), t(v)] = 0 \quad \forall u, v. \quad (2.11)$$

Assume furthermore that the Hamiltonian  $H$  of the spin chain is obtained from the regularity of the  $R$ -matrix and is given by the logarithmic derivative of the transfer matrix at some regular point, which we take for simplicity to be  $u = 0$ :

$$H = c \left. \frac{d}{du} \right|_{u=0} \log t(u) + \text{const} = c t'(0) t(0)^{-1} + \text{const}, \quad (2.12)$$

where  $c$  is a model-dependent constant (for instance,  $c \sim J$  in spin chain notation).

**Lemma 1.** *For all  $u$ , the transfer matrix  $t(u)$  commutes with both  $t(0)$  and  $t'(0)$ :*

$$[t(u), t(0)] = 0, \quad [t(u), t'(0)] = 0.$$

*Proof.* The first commutation relation is just a special case of (2.11) with  $v = 0$ :

$$[t(u), t(0)] = 0.$$

To obtain the second, differentiate (2.11) with respect to  $v$ :

$$\frac{d}{dv} [t(u), t(v)] = [t(u), t'(v)] = 0 \quad \forall u, v.$$

In particular, setting  $v = 0$  gives

$$[t(u), t'(0)] = 0.$$

□

**Theorem 1.** *The transfer matrix commutes with the Hamiltonian:*

$$[t(u), H] = 0 \quad \forall u.$$

*Proof.* Using (2.12) and ignoring the additive constant (which trivially commutes with everything), we compute

$$[t(u), H] = c [t(u), t'(0) t(0)^{-1}].$$

Since  $t(u)$  commutes with  $t'(0)$  and with  $t(0)$ , it also commutes with any polynomial (and hence any convergent power series) in these operators. In particular,  $[t(u), t(0)^{-1}] = 0$ , so

$$[t(u), t'(0) t(0)^{-1}] = [t(u), t'(0)] t(0)^{-1} + t'(0) [t(u), t(0)^{-1}] = 0,$$

and therefore  $[t(u), H] = 0$  for all  $u$ .

□

### 2.3.2. Conserved quantities from the analytic expansion of the transfer matrix

Suppose that the transfer matrix is analytic in a neighbourhood of  $u = 0$ . Then we can expand it as a power series

$$t(u) = \sum_{n=0}^{\infty} I_n u^n, \quad I_n \in \text{End}(\mathcal{H}), \quad (2.13)$$

where  $\mathcal{H}$  is the quantum Hilbert space of the chain.

From the theorem above we know that

$$[H, t(u)] = 0 \quad \forall u.$$

Substituting (2.13) into this commutator, we obtain

$$0 = [H, t(u)] = \left[ H, \sum_{n=0}^{\infty} I_n u^n \right] = \sum_{n=0}^{\infty} [H, I_n] u^n.$$

Since this identity holds as an analytic function of  $u$  in a neighbourhood of 0, all coefficients must vanish:

$$[H, I_n] = 0 \quad \forall n \geq 0.$$

**Corollary 1.** *Each coefficient  $I_n$  in the expansion (2.13) is a conserved quantity for the Hamiltonian  $H$ :*

$$[H, I_n] = 0 \quad \forall n.$$

*Equivalently, in the Heisenberg picture,*

$$\frac{d}{dt} I_n(t) = i[H, I_n(t)] = 0,$$

*so the operators  $I_n$  are integrals of motion. Moreover, the commuting family of transfer matrices implies*

$$[I_m, I_n] = 0,$$

*so these conserved charges commute among themselves.*

Thus, starting from the commuting family of transfer matrices  $t(u)$  and the Hamiltonian defined through the logarithmic derivative (2.12), we obtain an infinite family of pairwise commuting conserved quantities  $\{I_n\}_{n \geq 0}$ , which is a hallmark of quantum integrability.

## 2.4. Deriving Solvable models from Yang-Baxter equation

### 2.4.1. Symmetry and $G$ -invariance of the $R$ -matrix

Up to now we have focused on the role of the  $R$ -matrix in producing a commuting family of transfer matrices. In many concrete models the Lax operators and monodromy matrix also respect a global symmetry given by a group (or Hopf algebra)  $G$ . In this subsection we explain how this symmetry forces the  $R$ -matrix itself to be  $G$ -invariant.

#### Setup: $G$ -action and intertwiners

Let  $G$  be a (quantum) group acting on each local quantum space  $V_j$  and on each auxiliary space  $V_a, V_b$  via representations

$$\pi_j : G \rightarrow \text{End}(V_j), \quad \pi_a : G \rightarrow \text{End}(V_a), \quad \pi_b : G \rightarrow \text{End}(V_b).$$

We write  $\pi^{(N)}$  for the diagonal action of  $G$  on the chain  $\mathcal{H} = V_1 \otimes \cdots \otimes V_N$ :

$$\pi^{(N)}(g) = \pi_1(g) \otimes \cdots \otimes \pi_N(g), \quad g \in G.$$

**Definition 1.** We say that a Lax operator  $L_{aj}(u) \in \text{End}(V_a \otimes V_j)$  is a  $G$ -intertwiner if

$$L_{aj}(u) (\pi_a(g) \otimes \pi_j(g)) = (\pi_a(g) \otimes \pi_j(g)) L_{aj}(u) \quad \forall g \in G, \forall u. \quad (2.14)$$

Similarly for  $L_{bj}(v) \in \text{End}(V_b \otimes V_j)$ .

If all Lax operators  $L_{aj}(u)$  are  $G$ -intertwiners, then the monodromy matrix

$$T_a(u) = L_{aN}(u) \cdots L_{a1}(u) \in \text{End}(V_a \otimes \mathcal{H})$$

is also a  $G$ -intertwiner:

$$T_a(u) (\pi_a(g) \otimes \pi^{(N)}(g)) = (\pi_a(g) \otimes \pi^{(N)}(g)) T_a(u) \quad \forall g \in G, \forall u. \quad (2.15)$$

In this sense we say that the monodromy is  $G$ -invariant.

### Statement

Recall the local  $RLL$ -relation for two auxiliary spaces  $V_a$  and  $V_b$ :

$$R_{ab}(u, v) L_{aj}(u) L_{bj}(v) = L_{bj}(v) L_{aj}(u) R_{ab}(u, v), \quad j = 1, \dots, N. \quad (2.16)$$

Here  $R_{ab}(u, v)$  acts only on  $V_a \otimes V_b$ .

**Proposition 2** (Symmetry forces  $G$ -invariance of  $R$ ). *Assume:*

1. Each Lax operator  $L_{aj}(u)$  and  $L_{bj}(v)$  is a  $G$ -intertwiner in the sense of (2.14).
2. For each pair  $(u, v)$ , the space of operators

$$X_{ab}(u, v) \in \text{End}(V_a \otimes V_b)$$

satisfying the  $RLL$  relation (2.16) is one-dimensional.<sup>2</sup>

Then any  $R$ -matrix  $R_{ab}(u, v)$  entering the  $RLL$  relation is itself  $G$ -invariant:

$$(\pi_a(g) \otimes \pi_b(g)) R_{ab}(u, v) = R_{ab}(u, v) (\pi_a(g) \otimes \pi_b(g)) \quad \forall g \in G, \forall u, v. \quad (2.17)$$

*Proof.* Fix  $u, v$  and  $g \in G$ , and define the conjugated operator

$$\tilde{R}_{ab}(u, v; g) := (\pi_a(g) \otimes \pi_b(g)) R_{ab}(u, v) (\pi_a(g) \otimes \pi_b(g))^{-1}.$$

We claim that  $\tilde{R}_{ab}(u, v; g)$  is also a solution of the  $RLL$  relation (2.16).

Using the intertwining property (2.14) we have

$$(\pi_a(g) \otimes \pi_j(g)) L_{aj}(u) = L_{aj}(u) (\pi_a(g) \otimes \pi_j(g)), \quad (\pi_b(g) \otimes \pi_j(g)) L_{bj}(v) = L_{bj}(v) (\pi_b(g) \otimes \pi_j(g)).$$

Hence on  $V_a \otimes V_b \otimes V_j$  the diagonal action  $\pi_a(g) \otimes \pi_b(g) \otimes \pi_j(g)$  commutes with both  $L_{aj}(u)$  and  $L_{bj}(v)$ . Starting from the original  $RLL$  relation, conjugate by  $\pi_a(g) \otimes \pi_b(g) \otimes \pi_j(g)$ :

$$\begin{aligned} & (\pi_a \otimes \pi_b \otimes \pi_j) R_{ab} L_{aj} L_{bj} (\pi_a \otimes \pi_b \otimes \pi_j)^{-1} \\ &= (\pi_a \otimes \pi_b \otimes \pi_j) L_{bj} L_{aj} R_{ab} (\pi_a \otimes \pi_b \otimes \pi_j)^{-1}. \end{aligned}$$

Using commutativity of the representations with  $L_{aj}, L_{bj}$  we can move  $\pi_a \otimes \pi_b \otimes \pi_j$  through the Lax operators, obtaining

$$\tilde{R}_{ab}(u, v; g) L_{aj}(u) L_{bj}(v) = L_{bj}(v) L_{aj}(u) \tilde{R}_{ab}(u, v; g).$$

Thus  $\tilde{R}_{ab}(u, v; g)$  satisfies the same  $RLL$  relation as  $R_{ab}(u, v)$ .

---

<sup>2</sup>This is the typical situation when  $V_a \otimes V_b$  decomposes as a direct sum of pairwise non-isomorphic irreducible  $G$ -modules, so that Schur's lemma applies.

By assumption, the space of solutions to (2.16) is one-dimensional, so there exists a scalar  $c(u, v; g) \in \mathbb{C}^\times$  such that

$$\tilde{R}_{ab}(u, v; g) = c(u, v; g) R_{ab}(u, v).$$

If we fix a normalization of  $R_{ab}(u, v)$  (for example, by imposing a regularity condition  $R_{ab}(u, u) = P_{ab}$  or a similar condition at a reference point), the scalar must be  $c(u, v; g) = 1$ . Therefore

$$(\pi_a(g) \otimes \pi_b(g)) R_{ab}(u, v) (\pi_a(g) \otimes \pi_b(g))^{-1} = R_{ab}(u, v)$$

for all  $g \in G$ , which is equivalent to (2.17). This is exactly the statement that  $R_{ab}(u, v)$  is  $G$ -invariant.  $\square$

**Remark 4.** *Conceptually, the proposition expresses the following idea:*

- *The Lax operators are assumed to respect the symmetry  $G$ , so the monodromy (and hence the transfer matrix) is  $G$ -invariant.*
- *Any  $R$ -matrix that interchanges monodromies via the RTT/RLL relations must be compatible with this symmetry.*
- *Under mild uniqueness assumptions, this compatibility forces  $R$  to commute with the  $G$ -action; in other words,  $R$  lives in the commutant  $\text{End}_G(V_a \otimes V_b)$ .*

*Thus, once the symmetry group  $G$  of the local Hamiltonian is fixed, it is natural (and, in typical situations, unavoidable) to search for  $R$ -matrices inside the  $G$ -invariant subalgebra of  $\text{End}(V_a \otimes V_b)$ .*

#### 2.4.2. Reverse engineering integrable quantum spin chains

The transfer matrix formalism can be used not only to analyse a given quantum spin chain, but also to *construct* new exactly solvable models in a systematic way. Conceptually, one reverses the usual order of arguments: instead of starting from a Hamiltonian and searching for its commuting integrals of motion, one begins with an  $R$ -matrix satisfying the Yang–Baxter equation and then *derives* both the transfer matrix and the Hamiltonian from it.

Fix a symmetry algebra or group  $G$  and a finite-dimensional representation  $V$  which will serve as the local Hilbert space of a spin chain. We look for a family of operators

$$R(u) \in \text{End}(V \otimes V),$$

depending on a complex spectral parameter  $u$ , such that

(i) ( $G$ -invariance)

$$[R(u), g \otimes g] = 0 \quad \forall g \in G, \forall u, \quad (2.18)$$

so  $R(u)$  intertwines the diagonal action of  $G$  on  $V \otimes V$ ; and

(ii) ( $Yang$ – $Baxter$  equation)

$$R_{12}(u - v) R_{13}(u - w) R_{23}(v - w) = R_{23}(v - w) R_{13}(u - w) R_{12}(u - v) \quad (2.19)$$

as operators on  $V \otimes V \otimes V$ .

The  $G$ -invariance condition (2.18) allows us to use Schur–Weyl duality. The tensor product  $V \otimes V$  decomposes into irreducible  $G$ -modules as

$$V \otimes V \cong \bigoplus_{\lambda} V_{\lambda} \otimes M_{\lambda},$$

where  $V_{\lambda}$  are pairwise non-isomorphic irreducible  $G$ -modules and  $M_{\lambda}$  are multiplicity spaces. Any  $G$ -equivariant operator on  $V \otimes V$  acts as a scalar on each irreducible component, hence every  $G$ -invariant  $R$ -matrix has the form

$$R(u) = \sum_{\lambda} \rho_{\lambda}(u) P_{\lambda}, \quad (2.20)$$

where  $P_\lambda$  is the  $G$ -equivariant projector onto the isotypic component  $V_\lambda \otimes M_\lambda$  and  $\rho_\lambda(u)$  are scalar functions (“eigenvalues” of  $R(u)$  on each irreducible channel).

Substituting (2.20) into the Yang–Baxter equation (2.19) reduces the operator equation on  $V^{\otimes 3}$  to a finite set of functional equations for the scalar functions  $\rho_\lambda(u)$ . These equations are then supplemented by additional analytic and physical conditions, such as *regularity* and *unitarity*. Regularity requires that at some value of the spectral parameter (usually  $u = 0$ ) the two local spaces become indistinguishable, so that  $R(u)$  reduces to the permutation operator:

$$R(0) = P, \quad \implies \quad \rho_\lambda(0) = 1 \text{ for all } \lambda. \quad (2.21)$$

Unitarity expresses the invertibility of two-body scattering and is typically imposed in the form

$$R(u) R(-u) = \phi(u) \text{Id}_{V \otimes V}, \quad \implies \quad \rho_\lambda(u) \rho_\lambda(-u) = \phi(u), \quad (2.22)$$

where  $\phi(u)$  is a scalar function independent of  $\lambda$ . Together, the Yang–Baxter equation, Schur–Weyl decomposition, and constraints such as (2.21)–(2.22) often fix the functions  $\rho_\lambda(u)$  up to a small number of parameters, thereby determining the  $R$ -matrix explicitly.

Once an  $R$ -matrix  $R(u)$  with the properties above is known, one can construct the Lax operators and transfer matrices. For a chain of length  $N$  with local space  $V$ , introduce an auxiliary space  $V_a \cong V$  and define Lax operators by

$$L_{aj}(u) = R_{aj}(u - \xi_j), \quad (2.23)$$

where  $\xi_j$  are inhomogeneity parameters and  $R_{aj}$  acts on  $V_a \otimes V_j$ . Consequently, we proceed as with XXZ spin chain example from above.

It is important to stress that this “reverse engineering” procedure does not work for an arbitrary choice of symmetry group  $G$  and representation  $V$ . The existence of a non-trivial  $G$ -invariant solution of the Yang–Baxter equation satisfying regularity, unitarity and similar conditions is highly restrictive. In most known examples, the relevant symmetry is not an ordinary Lie group but a *quantum group* (or more generally a quasi-triangular Hopf algebra), whose universal  $R$ -matrix gives rise, upon evaluation in finite-dimensional representations, to families of  $R$ -matrices of the form (2.20). An introduction to quantum groups and their role in the classification of  $R$ -matrices will be presented in a later chapter; here we only emphasize that they provide the natural algebraic framework for constructing new integrable quantum spin chains via the transfer matrix formalism.

## Chapter 3

# Algebraic Bethe Ansatz

### 3.1. Introduction

**Guiding question:** Why do we introduce the transfer matrix to construct conserved quantities and diagonalize the Hamiltonian?

The algebraic Bethe ansatz (ABA) is one of the main realizations of the quantum inverse scattering method. Rather than attacking a given Hamiltonian directly, one first embeds it into a larger algebraic structure built from an  $R$ -matrix, a monodromy matrix and, finally, a *transfer matrix*. For concreteness, keep in mind a spin- $\frac{1}{2}$  Heisenberg chain with  $L$  sites, whose Hilbert space  $\mathcal{H} \simeq (\mathbb{C}^2)^{\otimes L}$  has dimension  $2^L$ . In principle, one could try to diagonalize the Hamiltonian  $H$  as a  $2^L \times 2^L$  matrix, but this quickly becomes impractical as  $L$  grows.

The transfer matrix

$$t(u) = \text{tr}_a T_a(u)$$

is an operator on  $\mathcal{H}$  depending on a *spectral parameter*  $u$ , obtained by tracing over an auxiliary space  $V_a$  of the monodromy matrix  $T_a(u)$ . If the underlying  $R$ -matrix satisfies the Yang–Baxter equation, one can show that

$$[t(u), t(v)] = 0 \quad \forall u, v,$$

so that  $\{t(u)\}$  forms a commuting family of operators. This observation answers our guiding question in two ways:

- **Systematic construction of conserved charges.** Expanding  $\ln t(u)$  around a regular point (typically  $u = 0$ ) produces an infinite hierarchy of mutually commuting operators

$$\ln t(u) = \sum_{n \geq 0} u^n I_n,$$

where  $I_0$  is proportional to the identity and  $I_1$  is (up to normalization and additive constants) the Hamiltonian  $H$  of the spin chain. Higher  $I_n$  then give conserved quantities in involution, realizing quantum integrability in a very explicit and systematic way.

- **Diagonalization via ABA.** Since all  $t(u)$  commute, they can be diagonalized simultaneously. The ABA constructs eigenvectors of  $t(u)$  (and hence of  $H$ ) algebraically, by acting with suitable creation operators (traditionally denoted  $B(u)$ ) on a simple reference state or “pseudovacuum”. The eigenvalue problem for an exponentially large matrix is thereby reduced to solving a finite set of algebraic equations for the spectral parameters (the Bethe roots).

In this chapter we introduce the and implement the algebraic Bethe ansatz to diagonalize the Hamiltonian of the Heisenberg-type spin chain.

The Hamiltonian we will consider is for XXZ periodic spin chain

$$\mathcal{H} = -\frac{J}{2} \sum_{i=1}^N (\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \Delta \sigma_i^z \sigma_{i+1}^z), \quad \text{where} \quad \sigma_{N+1}^{x,y,z} = \sigma_1^{x,y,z} \quad \text{and} \quad J > 0$$

The Hilbert space for this model is

$$\mathcal{H}_N = \bigotimes_{i=1}^N \mathfrak{h}_i, \quad \text{where } \forall i, \mathfrak{h}_i = \mathbb{C}^2$$

**Remark 5.** In the above expression, we adopted the following notation. Given the Hilbert space for the spin chain model with  $N$  sites as  $\mathcal{H} = V \otimes V \otimes \dots \otimes V$ , for an operator  $A \in \text{End}(V)$  the notation  $A_i$  means it acts as  $A$  only on space  $i^{\text{th}}$  space  $V$  and trivially on otherwise, i.e.

$$A_i = \mathbb{I}_1 \otimes \dots \otimes A \otimes \dots \otimes \mathbb{I}_n$$

Let's make the following recall from the last time: After defining the Lax operator for XYZ model we proceed further to obtain the Yang-Baxter equations in the special case of XXZ spin chain.

- Similar to a 6-vertex model we define the transfer matrix as

$$t_N = \text{tr}_{aux}(L_{N,aux} \cdots L_{1,aux})$$

and obtain the condition for the transfer matrices with different weights  $t(a, b, c)$ ,  $t(a', b', c')$  in Lax operators to commute

$$\Delta(a, b, c) = \frac{a^2 + b^2 - c^2}{2ab} = \frac{(a')^2 + (b')^2 - (c')^2}{2a'b'} = \Delta(a', b', c')$$

This can be deduced using the MatLab equation solver for the expression

$$L''_{aux,aux} \cdot L'_{n,aux} \cdot L_{n,aux} = L_{n,aux} \cdot L'_{n,aux} \cdot L''_{aux,aux}$$

where  $L''_{aux,aux}$  plays the role of intertwiner  $R$ -matrix. Denote this family of transfer matrices as  $\mathcal{F}$ .

- We also know that for a fixed value of  $\Delta$  in the Hamiltonian  $H_{XXZ}$ , it commutes with transfer matrices in the family  $\mathcal{F}$  which have  $\Delta(a, b, c) = \Delta$ .
- For variables  $a, b, c$  we can consider the parametrization

$$a = \sinh(u + \gamma), \quad b = \sinh(u), \quad c = \sinh(\gamma)$$

which then yields

$$\Delta(\sinh(u + \gamma), \sinh(u), \sinh(\gamma)) = -\cosh(\gamma)$$

If  $\gamma \in \mathbb{R}$  then  $\Delta \in \mathbb{R} - (-1, 1)$ , and in case  $\gamma$  is imaginary,  $\Delta \in [-1, 1]$ . This parametrization thus covers the antiferromagnetic and ferromagnetic cases of  $H_{XXZ}$ .

- Due to our parametrization we have the Lax matrix in form

$$R = \begin{pmatrix} \sinh(u + \gamma) & 0 & 0 & 0 \\ 0 & \sinh(u) & \sinh(\gamma) & 0 \\ 0 & \sinh(\gamma) & \sinh(u) & 0 \\ 0 & 0 & 0 & \sinh(u + \gamma) \end{pmatrix}$$

and can choose  $R$ -matrix to be in the same form.

- The conserved quantities from the transfer matrix  $t(u)$  for our Hamiltonian can be found as

$$\mathcal{Q}_l = \frac{d^l \log t(u)}{du^l} \Big|_{u=0}$$

where in fact it can be shown that

$$H = \mathcal{Q}_1 = \frac{d \log t(u)}{du} \Big|_{u=0} = t^{-1}(0)t'(0)$$

### 3.2. Ferromagnetic/Antiferromagnetic case

We want to study the eigenvectors and eigenvalues of  $\mathcal{H}$  and this depends on the sign of  $\Delta$  parameter.

- (a) For  $\Delta > 1$ , the system is ferromagnetic, i.e., our ground state is of the form

$$|0\rangle = |\uparrow\uparrow\uparrow \dots \uparrow\rangle$$

- (b) For  $\Delta < 1$ , the system is antiferromagnetic, i.e., our ground state is of the form

$$|0\rangle = \sum_{\sigma \in S} a_{\sigma} |\sigma\rangle$$

where for even  $N$ , the set  $S$  is made of all states which have half of the spins up and half down. For odd  $N$ , there are two possible ground states: one with  $\frac{N+1}{2}$  spins up and  $\frac{N-1}{2}$  spins down, and the opposite configuration.

**Remark 6.** Important to note that for the antiferromagnetic case, the highest excited state is the same as the ground state for the ferromagnetic case.

### 3.3. Components of Monodromy Matrix and the ferromagnetic pseudo-vacuum

1. Let's take a closer look at the monodromy matrix

$$\mathcal{T}_N(u) = L_{aux,N}(u) \cdot \dots \cdot L_{aux,1}(u) = \begin{pmatrix} \mathcal{A}_N(u) & \mathcal{B}_N(u) \\ \mathcal{C}_N(u) & \mathcal{D}_N(u) \end{pmatrix}$$

where for pedagogical purposes we will use  $N = 2$  and  $N = 3$ .

2. The monodromy matrix is acting on Hilbert space  $\mathcal{H}_{+aux} = \mathcal{H}_{chain} \otimes V_{aux} = \mathcal{H}_{chain} \otimes \mathbb{C}^2$  which implies that the matrices  $\mathcal{A}_N(u), \mathcal{B}_N(u), \mathcal{C}_N(u), \mathcal{D}_N(u)$  act on the spin chain Hilbert space  $\mathcal{H}_{chain}$  of dimension  $2^N$ . The matrix above is represented in such form that implies

$$t_N(u) = \text{tr}_{aux}(\mathcal{T}_N(u)) = \mathcal{A}_N(u) + \mathcal{D}_N(u)$$

3. Our goal is to diagonalize the transfer matrix  $t_N(u)$ . It starts from considering the state

$$|0\rangle = |\uparrow\uparrow \dots \uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

which is a ground state for a ferromagnetic case and the highest energy for the antiferromagnetic case.

4. Using the MatLab, we find that for  $N = 2, 3$

$$\mathcal{A}_2(u) |0\rangle = \sinh^2(u + \gamma) |0\rangle, \mathcal{A}_3(u) |0\rangle = \sinh^3(u + \gamma) |0\rangle, \dots \implies \mathcal{A}_N(u) |0\rangle = \sinh^N(u + \gamma) |0\rangle$$

and similarly

$$\mathcal{D}_2(u) |0\rangle = \sinh^2(u) |0\rangle, \mathcal{D}_3(u) |0\rangle = \sinh^3(u) |0\rangle, \dots \implies \mathcal{D}_N(u) |0\rangle = \sinh^N(u) |0\rangle$$

and similarly

$$\mathcal{C}_2(u) |0\rangle = 0, \mathcal{C}_3(u) |0\rangle = 0, \dots \implies \mathcal{C}_N(u) |0\rangle = 0$$

and similarly

$$\mathcal{B}_N(u) |0\rangle = \sinh(\gamma) \sum_{i=1}^N \sinh^{N-i}(u + \gamma) \sinh^{i-1}(u) |i\rangle$$

where

$$|i\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{i^{th} \text{ site}} \otimes \dots \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

is the state with the  $i^{th}$  spin down. This last result was checked for  $N = 2, 3$  and therefore assumed to hold for all  $N > 3$ .

5. The main insight from these results comes as that we can treat  $\mathcal{B}(u)$  and  $\mathcal{C}(u)$  as creation or annihilation operators depending on ferromagnetic/antiferromagnetic cases and consider  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  as generators of non-commutative algebra called Yang-Baxter algebra. In fact, we can check that some of the commutation relations are

$$\begin{cases} [\mathcal{B}(u), \mathcal{B}(v)] = 0, \\ \sinh(u-v)\mathcal{A}(u)\mathcal{B}(v) = \sinh(\gamma)\mathcal{B}(u)\mathcal{A}(v) + \sinh(u-v-\gamma)\mathcal{B}(v)\mathcal{A}(u), \\ \sinh(u-v)\mathcal{D}(u)\mathcal{B}(v) = -\sinh(\gamma)\mathcal{D}(u)\mathcal{B}(v) + \sinh(u-v+\gamma)\mathcal{B}(v)\mathcal{D}(u) \end{cases}$$

6. To simplify the language, from now on we will consider the ferromagnetic case where  $|0\rangle$  is a ground state,  $\mathcal{B}(u)$  is a creation operator, and  $\mathcal{C}(u)$  is an annihilation operator. We use the ground state  $|0\rangle$  to build excited states called **Bethe states**

$$|\beta u_1, \dots, u_k\rangle = \mathcal{B}(u_1) \cdots \mathcal{B}(u_k) |0\rangle$$

Where  $u_1, \dots, u_k$  are distinct choices of spectral parameter and  $k \leq N$  and where the number  $k$  corresponds to  $k$  spins being up in the state  $|\beta u_1, \dots, u_k\rangle$ .

7. Our main goal is to diagonalize the transfer matrix  $t(u)$ , where  $u$  varies. Ideally, we want  $|\beta u_1, \dots, u_k\rangle$  to be eigenstates of  $t(u)$  such that

$$t(u) |\beta u_1, \dots, u_k\rangle = (\mathcal{A}(u) + \mathcal{D}(u)) |\beta u_1, \dots, u_k\rangle = \beta u, u_1, \dots, u_k |\beta u_1, \dots, u_k\rangle$$

but not for all choices of  $u_1, \dots, u_k$  this holds, because of non-commutativity relations given above.

8. Let's demonstrate how non-commutativity affects the eigenstate status for  $|\beta u_1, \dots, u_k\rangle$

$$\begin{aligned} t(u) |\beta u_1, \dots, u_k\rangle &= \mathcal{A}(u) + \mathcal{D}(u) |\beta u_1, \dots, u_k\rangle \\ &= \frac{\sinh(\gamma)}{\sinh(u-u_1)} \mathcal{B}(u) \mathcal{A}(u_1) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle + \\ &\quad - \frac{\sinh(\gamma)}{\sinh(u-u_1)} \mathcal{B}(u) \mathcal{D}(u_1) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle + \\ &\quad + \frac{\sinh(u-u_1-\gamma)}{\sinh(u-u_1)} \mathcal{B}(u_1) \mathcal{A}(u) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle + \\ &\quad + \frac{\sinh(u-u_1-\gamma)}{\sinh(u-u_1)} \mathcal{B}(u_1) \mathcal{D}(u) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle \end{aligned} \tag{3.1}$$

If we want to pass  $\mathcal{A}(u)$  and  $\mathcal{D}(u)$  through  $\mathcal{B}(u_1) \cdots \mathcal{B}(u_k)$  i.e.

$$\mathcal{A}(u) \mathcal{B}(u_1) \cdots \mathcal{B}(u_k) \rightarrow \mathcal{B}(u_1) \cdots \mathcal{B}(u_k) \mathcal{A}(u) \quad \text{and} \quad \mathcal{D}(u) \mathcal{B}(u_1) \cdots \mathcal{B}(u_k) \rightarrow \mathcal{B}(u_1) \cdots \mathcal{B}(u_k) \mathcal{D}(u)$$

then the first two terms in the above sum fail to have the desired exchange relations between  $\mathcal{A}(u), \mathcal{D}(u)$  and  $\mathcal{B}(u_1)$ , so we need the canceled if want to make  $|0\rangle$  as the eigenvector. Therefore the unwanted term is

$$\text{"unwanted term 1"} = \frac{\sinh(\gamma)}{\sinh(u-u_1)} \mathcal{B}(u) \mathcal{A}(u_1) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle - \frac{\sinh(\gamma)}{\sinh(u-u_1)} \mathcal{B}(u) \mathcal{D}(u_1) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle$$

9. Thus our main strategy can be summarized as follows

- Keep passing  $\mathcal{A}(u)$  and  $\mathcal{B}(u)$  and collecting all the wanted terms. Read off the final eigenvalue  $\beta\{u_1, \dots, u_k\}$  for the state proportional to  $|0\rangle$ .
- Demand that the unwanted terms cancel to get the so-called Bethe Ansatz equations for the allowed values of  $u$ .

10. Let's consider the unwanted term above. We can keep applying the commutation relations from Yang-Baxter algebra

$$\begin{aligned}
\text{"unwanted term 1"} &= \frac{\sinh(\gamma)}{\sinh(u - u_1)} \mathcal{B}(u) \mathcal{A}(u_1) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle - \frac{\sinh(\gamma)}{\sinh(u - u_1)} \mathcal{B}(u) \mathcal{D}(u_1) \mathcal{B}(u_2) \dots \mathcal{B}(u_k) |0\rangle \\
&= \frac{\sinh(\gamma)}{\sinh(u - u_1)} \frac{\sinh(u_1 - u_2 - \gamma)}{\sinh(u_1 - u_2)} \mathcal{B}(u) \mathcal{B}(u_2) \mathcal{A}(u_1) \dots \mathcal{B}(u_k) |0\rangle - \\
&\quad - \frac{\sinh(\gamma)}{\sinh(u - u_1)} \frac{\sinh(u_1 - u_2 + \gamma)}{\sinh(u_1 - u_2)} \mathcal{B}(u) \mathcal{B}(u_2) \mathcal{D}(u_1) \dots \mathcal{B}(u_k) |0\rangle \\
&\quad + \text{"unwanted term 2 containing } \mathcal{A}(u_2) \text{ and } \mathcal{D}(u_2)\text{"}
\end{aligned} \tag{3.2}$$

Where the unwanted term now is

$$\begin{aligned}
\text{"unwanted term 2"} &= \frac{\sinh^2(\gamma)}{\sinh(u - u_1) \sinh(u_1 - u_2)} \mathcal{B}(u) \mathcal{B}(u_1) \mathcal{A}(u_2) \dots \mathcal{B}(u_k) |0\rangle - \\
&\quad - \frac{\sinh^2(\gamma)}{\sinh(u - u_1) \sinh(u_1 - u_2)} \mathcal{B}(u) \mathcal{B}(u_1) \mathcal{D}(u_2) \dots \mathcal{B}(u_k) |0\rangle
\end{aligned}$$

This complicated relation gives a recursion which we can apply later. If we continue to expand the first unwanted term, by moving  $\mathcal{A}(u_1)$  and  $\mathcal{B}(u_2)$  to the right, then we will get

$$\begin{aligned}
&\cdot \\
&\cdot \\
&\cdot \\
&= \left( \frac{\sinh(\gamma)}{\sinh(u - u_1)} \prod_{i=2}^k \frac{\sinh(u_1 - u_i - \gamma)}{\sinh(u_1 - u_i)} \right) \sinh^N(u_1 + \gamma) |0\rangle + \\
&\quad - \left( \frac{\sinh(\gamma)}{\sinh(u - u_1)} \prod_{i=2}^k \frac{\sinh(u_1 - u_i + \gamma)}{\sinh(u_1 - u_i)} \right) \sinh^N(u_1) |0\rangle + \\
&\quad + \text{"unwanted term containing } \{\mathcal{A}(u_i)\}_{i=2}^k \text{ and } \{\mathcal{D}(u_i)\}_{i=2}^k \text{"}
\end{aligned} \tag{3.3}$$

Where  $\mathcal{A}(u_1) |0\rangle = \sinh^N(u_1 + \gamma) |0\rangle$  and  $\mathcal{D}(u_1) |0\rangle = \sinh^N(u_1) |0\rangle$ . The unwanted term looks like a very complicated expression but due to commutativity  $[\mathcal{B}(u), \mathcal{B}(v)] = 0$  and the recursive relation above, it actually can be reduced to a sum of the form

$$\sum_{i=2}^k \mathcal{B}(u) \mathcal{B}(u_1) \dots \widehat{\mathcal{B}(u_i)} \dots \mathcal{B}(u_k) (\sinh^N(u_i + \gamma) \Gamma_{i,A} + \sinh^N(u_i) \Gamma_{i,D}) |0\rangle$$

where  $\mathcal{A}(u_i) |0\rangle = \sinh^N(u_i + \gamma) |0\rangle$ ,  $\mathcal{D}(u_i) |0\rangle = \sinh^N(u_i) |0\rangle$   $u_0 = u$  such that

$$\Gamma_{i,A} = \frac{\sinh^i(\gamma)}{\prod_{j=1}^i \sinh(u_{j-1} - u_j)} \prod_{j \neq i}^k \frac{\sinh(u_i - u_j - \gamma)}{\sinh(u_i - u_j)}$$

and

$$\Gamma_{i,D} = (-1)^i \frac{\sinh^i(\gamma)}{\prod_{j=1}^i \sinh(u_{j-1} - u_j)} \prod_{j \neq i}^k \frac{\sinh(u_i - u_j + \gamma)}{\sinh(u_i - u_j)}$$

11. To cancel the unwanted terms we will require that the parameters  $u_1, \dots, u_k$  satisfy

$$\sinh^N(u_i + \gamma) \Gamma_{i,A} + \sinh^N(u_i) \Gamma_{i,D} = 0, \quad \forall i = 1, 2, \dots, k$$

which is equivalent to

$$\left( \frac{\sinh(u_i + \gamma)}{\sinh(u_i)} \right)^N = \prod_{j \neq i}^k \frac{\sinh(u_i - u_j - \gamma)}{\sinh(u_i - u_j + \gamma)}$$

These are called Bethe Ansatz equations and the eigenvalue corresponding to  $\mathcal{B}(u_1) \cdots \mathcal{B}(u_k) |0\rangle$  was defined already defined above

$$\begin{aligned} \beta\{u_1, \dots, u_k\} &= \left( \frac{\sinh(\gamma)}{\sinh(u - u_1)} \prod_{i=2}^k \frac{\sinh(u_1 - u_i - \gamma)}{\sinh(u_1 - u_i)} \right) \sinh^N(u_1 + \gamma) + \\ &\quad - \left( \frac{\sinh(\gamma)}{\sinh(u - u_1)} \prod_{i=2}^k \frac{\sinh(u_1 - u_i + \gamma)}{\sinh(u_1 - u_i)} \right) \sinh^N(u_1) \end{aligned}$$

such that

$$t(u) |\beta\{u_1, \dots, u_k\}\rangle = \beta\{u_1, \dots, u_k\} |\beta\{u_1, \dots, u_k\}\rangle$$

Since the conserved quantities from the transfer matrix can be expressed as

$$\mathcal{Q}_l = \frac{d^l \log t(u)}{du^l} \Big|_{u=0}$$

then their eigenvalues will be equal to

$$q_l = \frac{d^l \log \beta\{u_1, \dots, u_k\}}{du^l}, \quad \forall k = 1, \dots, N$$

12. The resulting equation for the eigenvalues of the transfer matrix is defined as

$$\Lambda(u, \{u_1, \dots, u_m\}) = \sinh^N(u + \gamma) \prod_{i=1}^m \frac{\sinh(u - u_i - \gamma)}{\sinh(u - u_i)} + \sinh^N(u) \prod_{i=1}^m \frac{\sinh(u - u_i + \gamma)}{\sinh(u - u_i)}$$

### 3.4. Summary of the Algebraic Bethe Ansatz for the XXZ spin chain

- The monodromy matrix for the spin- $\frac{1}{2}$  XXZ chain

$$T_a(u) = \begin{pmatrix} \mathcal{A}(u) & \mathcal{B}(u) \\ \mathcal{C}(u) & \mathcal{D}(u) \end{pmatrix}_a$$

has operator-valued entries  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  acting on the quantum space  $\mathcal{H} \simeq (\mathbb{C}^2)^{\otimes N}$ . They satisfy a non-abelian algebra of commutation relations (the *Yang-Baxter algebra*) derived from the fundamental *RLL*-relation.

- The chosen reference (pseudo-vacuum) state is

$$|0\rangle = |\uparrow\uparrow \cdots \uparrow\rangle,$$

which is an eigenvector of  $\mathcal{A}(u)$  and  $\mathcal{D}(u)$  and is annihilated by  $\mathcal{C}(u)$ :

$$\mathcal{A}(u) |0\rangle = a(u) |0\rangle, \quad \mathcal{D}(u) |0\rangle = d(u) |0\rangle, \quad \mathcal{C}(u) |0\rangle = 0.$$

- The operator  $\mathcal{B}(u)$  acts as a creation operator for down-spins relative to the pseudo-vacuum. For each rapidity  $u$ , the vector  $\mathcal{B}(u) |0\rangle$  lies in the subspace with one spin flipped (one down-spin). More generally,

$$|\psi_m(u_1, \dots, u_m)\rangle = \mathcal{B}(u_1) \cdots \mathcal{B}(u_m) |0\rangle$$

lies in the sector with  $m$  down-spins (fixed total  $S^z = \frac{N}{2} - m$ ).

- By acting with products of  $\mathcal{B}(u)$  on  $|0\rangle$  and allowing the rapidities  $u_1, \dots, u_m$  to vary, one generates candidate vectors in each fixed-magnetization sector of the Hilbert space. Among these candidates we seek eigenvectors of the transfer matrix

$$t(u) = \text{tr}_a T_a(u) = \mathcal{A}(u) + \mathcal{D}(u)$$

(for periodic boundary conditions).

- The goal of the algebraic Bethe ansatz is to find rapidities  $u_1, \dots, u_m$  such that the Bethe vector  $|\psi_m(u_1, \dots, u_m)\rangle$  becomes a common eigenvector of  $t(u)$  for all  $u$ , i.e.

$$t(u) |\psi_m(u_1, \dots, u_m)\rangle = \Lambda(u; u_1, \dots, u_m) |\psi_m(u_1, \dots, u_m)\rangle.$$

- This would be trivial if  $\mathcal{A}(u)$  and  $\mathcal{D}(u)$  commuted with  $\mathcal{B}(v)$  for all  $u, v$ , but the Yang–Baxter algebra gives non-trivial commutation relations. Using these relations, one can move  $\mathcal{A}(u)$  and  $\mathcal{D}(u)$  through the string of  $\mathcal{B}$ ’s until they hit the pseudo-vacuum:

$$\mathcal{A}(u) |\psi_m\rangle = A_{\text{wanted}}(u; u_1, \dots, u_m) |\psi_m\rangle + \sum_{i=1}^m |\text{unwanted term } i\rangle,$$

and an analogous formula for  $\mathcal{D}(u) |\psi_m\rangle$ .

- The Bethe equations arise as the conditions that all “unwanted terms” vanish. For the XXZ model with anisotropy parameter  $\gamma$  (in a common convention), the Bethe equations take the form

$$\left( \frac{\sinh(u_i + \gamma)}{\sinh(u_i)} \right)^N = \prod_{\substack{k=1 \\ k \neq i}}^m \frac{\sinh(u_i - u_k + \gamma)}{\sinh(u_i - u_k - \gamma)}, \quad i = 1, \dots, m,$$

where  $\gamma$  is the parameter appearing in the  $R$ -matrix.

- When the rapidities  $\{u_1, \dots, u_m\}$  satisfy the Bethe equations, the unwanted terms cancel and the Bethe vector  $|\psi_m(u_1, \dots, u_m)\rangle$  becomes an eigenstate of  $t(u)$  with eigenvalue

$$\Lambda(u; u_1, \dots, u_m) = a(u) \prod_{k=1}^m f(u, u_k) + d(u) \prod_{k=1}^m f(u_k, u),$$

for an explicitly known function  $f$  determined by the  $R$ -matrix.

- For each fixed  $m$  (number of down-spins), the dimension of the corresponding magnetization sector is  $\binom{N}{m}$ . Under suitable assumptions (generic parameters and boundary conditions), the set of Bethe solutions is expected to yield a complete set of eigenstates in each sector, with cardinality matching  $\binom{N}{m}$ .
- The Bethe equations do not depend on the auxiliary spectral parameter  $u$ , while the eigenvalue  $\Lambda(u; u_1, \dots, u_m)$  is a function of  $u$ . Since the family  $\{t(u)\}_u$  is commuting, each Bethe eigenstate simultaneously diagonalizes all conserved charges obtained from the expansion of  $\ln t(u)$ , in particular the XXZ Hamiltonian  $H$ .
- The crucial symmetry used in this construction is the  $U(1)$  symmetry generated by the total spin  $S^z$ :

$$[H, \Delta^{(N)}(S^z)] = 0.$$

This decomposes  $\mathcal{H}$  into magnetization sectors and guarantees the existence of a reference state  $|0\rangle$  which is annihilated by  $\mathcal{C}(u)$  and serves as a highest-weight vector for the Yang–Baxter algebra. Full  $SU(2)$  symmetry is present only in the isotropic limit (XXX model).

- For more general models, such as the XYZ spin chain, the same ABA construction with a simple pseudo-vacuum is no longer directly available (e.g. with periodic boundary conditions). These models can still be solved by related methods (Baxter’s eight-vertex model,  $Q$ -operator, vertex–IRF transformations, etc.), but the structure of the reference state and the Bethe ansatz is more involved than in the XXZ case.

## Chapter 4

# Quasi-triangular Hopf algebras and solutions of the Yang–Baxter equation

In this final chapter we explain a general mechanism that produces solutions of the Yang–Baxter equation (YBE). The input is a *quasi-triangular Hopf algebra*  $(H, R)$ , and the output is, for every finite-dimensional representation  $V$  of  $H$ , a linear operator  $\check{R}_{V,V} \in \text{End}(V \otimes V)$  satisfying the (constant) Yang–Baxter equation. In the last section we illustrate how the Hopf algebra structure of  $U(\mathfrak{sl}_2)$  controls the  $\mathfrak{sl}_2$  symmetry of the spin- $\frac{1}{2}$  chain.

Throughout we work over  $\mathbb{C}$  for simplicity.

### 4.1. The Yang–Baxter equation revisited

Recall that given a vector space  $V$ , a (constant) *R-matrix* is an invertible linear map

$$\check{R}: V \otimes V \longrightarrow V \otimes V.$$

We use the standard leg-notation on  $V^{\otimes 3}$ :

$$\check{R}_{12} = \check{R} \otimes \text{id}_V \in \text{End}(V^{\otimes 3}), \quad (4.1)$$

$$\check{R}_{23} = \text{id}_V \otimes \check{R} \in \text{End}(V^{\otimes 3}), \quad (4.2)$$

$$\check{R}_{13} = (\text{id}_V \otimes P)(\check{R} \otimes \text{id}_V)(\text{id}_V \otimes P), \quad (4.3)$$

where  $P: V \otimes V \rightarrow V \otimes V$  is the flip  $P(x \otimes y) = y \otimes x$ .

**Definition 2.** A linear map  $\check{R} \in \text{End}(V \otimes V)$  is said to satisfy the Yang–Baxter equation if

$$\check{R}_{12}\check{R}_{23}\check{R}_{12} = \check{R}_{23}\check{R}_{12}\check{R}_{23} \quad \text{in } \text{End}(V^{\otimes 3}). \quad (4.4)$$

This is sometimes called the *braid form* of the YBE, because (4.4) is exactly the braid relation for the generators of the braid group  $B_3$  acting on  $V^{\otimes 3}$ .

In what follows we will construct such  $\check{R}$  starting from an algebraic object  $H$  equipped with extra structure.

### 4.2. Hopf algebras: algebra and coalgebra in one object

#### 4.2.1. Algebras and coalgebras

A (unital associative) *algebra* over  $\mathbb{C}$  is a vector space  $H$  equipped with

- a multiplication  $m: H \otimes H \rightarrow H$ , usually written  $m(a \otimes b) = ab$ ;
- and a unit map  $u: \mathbb{C} \rightarrow H$ ,  $u(1) = 1_H$ ,

such that  $m$  is associative and  $1_H$  is a two-sided unit:

$$(ab)c = a(bc), \quad 1_H a = a 1_H = a.$$

Dually, a *coalgebra* is a vector space  $H$  with

- a comultiplication  $\Delta: H \rightarrow H \otimes H$ ,
- and a counit  $\varepsilon: H \rightarrow \mathbb{C}$ ,

such that  $\Delta$  is coassociative and  $\varepsilon$  is a counit:

$$(\Delta \otimes \text{id}) \circ \Delta = (\text{id} \otimes \Delta) \circ \Delta, \quad (\varepsilon \otimes \text{id}) \circ \Delta = (\text{id} \otimes \varepsilon) \circ \Delta = \text{id}.$$

#### 4.2.2. Hopf algebras

A Hopf algebra combines these two structures in a compatible way.

**Definition 3.** A Hopf algebra is a quintuple

$$(H, m, u, \Delta, \varepsilon, S)$$

such that:

1.  $(H, m, u)$  is an algebra;
2.  $(H, \Delta, \varepsilon)$  is a coalgebra;
3.  $\Delta$  and  $\varepsilon$  are algebra homomorphisms, i.e.

$$\Delta(ab) = \Delta(a)\Delta(b), \quad \Delta(1_H) = 1_H \otimes 1_H,$$

and similarly for  $\varepsilon$ ;

4.  $S: H \rightarrow H$  is a linear map (the antipode) such that

$$m \circ (S \otimes \text{id}) \circ \Delta = m \circ (\text{id} \otimes S) \circ \Delta = u \circ \varepsilon.$$

We will not need the antipode for the derivation of the YBE from the quasi-triangular structure, but it is part of the standard definition and is important in representation theory.

#### 4.2.3. Representations and the tensor product via the coproduct

A representation (or module) of  $H$  is a vector space  $V$  together with an algebra homomorphism

$$\rho_V: H \rightarrow \text{End}(V).$$

The key role of the coproduct  $\Delta$  is that it allows us to define a consistent action of  $H$  on tensor products of representations.

**Definition 4.** Let  $V$  and  $W$  be  $H$ -modules with actions  $\rho_V$  and  $\rho_W$ . Define an action of  $H$  on  $V \otimes W$  by the requirement

$$\rho_{V \otimes W}(h) := (\rho_V \otimes \rho_W)(\Delta(h)), \quad h \in H.$$

Because  $\Delta$  is an algebra homomorphism, the map  $\rho_{V \otimes W}$  is again an algebra homomorphism:

$$\rho_{V \otimes W}(h_1 h_2) = \rho_{V \otimes W}(h_1) \rho_{V \otimes W}(h_2),$$

so  $V \otimes W$  is indeed an  $H$ -module. Coassociativity of  $\Delta$  implies that the action on triple tensor products  $V \otimes W \otimes U$  does not depend on the way we parenthesize:

$$\rho_{(V \otimes W) \otimes U} = \rho_{V \otimes (W \otimes U)}.$$

Thus a Hopf algebra is precisely the structure that allows us to *tensor* its representations in a coherent way.

### 4.3. Quasi-triangular Hopf algebras and the universal $R$ -matrix

We now introduce the additional structure that will produce solutions to the Yang–Baxter equation.

#### 4.3.1. Definition and axioms

Let  $H$  be a Hopf algebra with coproduct  $\Delta$ . We write  $\Delta^{\text{op}}: H \rightarrow H \otimes H$  for the *opposite coproduct*,

$$\Delta^{\text{op}}(h) := \tau \circ \Delta(h),$$

where  $\tau: H \otimes H \rightarrow H \otimes H$  is the flip  $\tau(a \otimes b) = b \otimes a$ .

**Definition 5.** A quasi-triangular Hopf algebra is a Hopf algebra  $H$  together with an invertible element

$$R \in H \otimes H$$

(the universal  $R$ -matrix) such that:

1. For all  $h \in H$ ,

$$\Delta^{\text{op}}(h) = R \Delta(h) R^{-1}. \quad (4.5)$$

2. The following compatibility relations hold:

$$(\Delta \otimes \text{id})(R) = R_{13} R_{23}, \quad (4.6)$$

$$(\text{id} \otimes \Delta)(R) = R_{13} R_{12}. \quad (4.7)$$

Here

$$R_{12} = R \otimes 1, \quad R_{23} = 1 \otimes R, \quad R_{13} = (\text{id} \otimes \tau)(R \otimes 1),$$

as elements of  $H^{\otimes 3}$ .

Intuitively, equation (4.5) says that  $R$  implements a “twisted commutativity” of the coproduct: one can pass from the coproduct to the opposite coproduct by conjugating with  $R$ . The equations (4.6)–(4.7) express the compatibility of this twist with tensor products of three factors.

#### 4.3.2. The universal Yang–Baxter equation in $H^{\otimes 3}$

A key fact is that the axioms above imply a Yang–Baxter type identity directly in  $H^{\otimes 3}$ .

**Proposition 3.** If  $(H, R)$  is a quasi-triangular Hopf algebra, then the universal  $R$ -matrix satisfies the universal Yang–Baxter equation

$$R_{12} R_{13} R_{23} = R_{23} R_{13} R_{12} \quad \text{in } H^{\otimes 3}. \quad (4.8)$$

*Sketch of proof.* Apply  $(\Delta \otimes \text{id})$  to (4.6) in two different ways using coassociativity. On the one hand,

$$(\Delta \otimes \text{id} \otimes \text{id})((\Delta \otimes \text{id})(R)) = (\Delta \otimes \text{id} \otimes \text{id})(R_{13} R_{23}),$$

and on the other hand, using coassociativity  $(\Delta \otimes \text{id}) \circ \Delta = (\text{id} \otimes \Delta) \circ \Delta$  and (4.7), one can rewrite the same expression in a different order. Carefully expanding both sides and using that  $(\Delta \otimes \text{id})$  is an algebra homomorphism, one eventually arrives at (4.8). A complete proof can be found in any standard reference on quantum groups; here we only need the existence of (4.8).  $\square$

Thus, before even talking about representations, a quasi-triangular Hopf algebra carries a distinguished element  $R$  obeying a Yang–Baxter type equation in  $H^{\otimes 3}$  itself.

### 4.4. From universal $R$ to $R$ -matrices on representations

We now explain how the element  $R \in H \otimes H$  produces, for each pair of  $H$ -modules, a linear operator solving the YBE.

#### 4.4.1. Induced operators on $V \otimes W$

Let  $V$  and  $W$  be finite-dimensional representations of  $H$  with actions  $\rho_V: H \rightarrow \text{End}(V)$  and  $\rho_W: H \rightarrow \text{End}(W)$ .

The tensor product representation  $\rho_{V \otimes W}$  is defined by

$$\rho_{V \otimes W}(h) = (\rho_V \otimes \rho_W)(\Delta(h)).$$

Since  $R$  is an element of  $H \otimes H$ , we can apply  $\rho_V \otimes \rho_W$  to  $R$ :

**Definition 6.** The  $R$ -matrix on  $V \otimes W$  associated to the quasi-triangular Hopf algebra  $(H, R)$  is

$$R_{V,W} := (\rho_V \otimes \rho_W)(R) \in \text{End}(V \otimes W).$$

Frequently, one inserts a flip and considers the braid  $R$ -matrix

$$\check{R}_{V,W} := P_{V,W} \circ R_{V,W},$$

where  $P_{V,W}: V \otimes W \rightarrow W \otimes V$  is the flip.

The definition is natural:  $R$  lives in  $H \otimes H$ , and for any representation of  $H$ , we obtain a representation of  $H \otimes H$  on  $V \otimes W$  by applying  $\rho_V \otimes \rho_W$ . Evaluating the universal element  $R$  in that representation gives a concrete operator on  $V \otimes W$ .

#### 4.4.2. The Yang–Baxter equation on $U \otimes V \otimes W$

Let  $U, V, W$  be three  $H$ -modules. We want to show that the operators  $R_{U,V}, R_{U,W}, R_{V,W}$  satisfy a Yang–Baxter relation on  $U \otimes V \otimes W$ .

Consider the tensor representation

$$\rho_{U \otimes V \otimes W}: H^{\otimes 3} \longrightarrow \text{End}(U \otimes V \otimes W)$$

given by

$$\rho_{U \otimes V \otimes W} = \rho_U \otimes \rho_V \otimes \rho_W.$$

Applying  $\rho_{U \otimes V \otimes W}$  to the universal Yang–Baxter equation (4.8), we obtain

$$\rho_{U \otimes V \otimes W}(R_{12}) \rho_{U \otimes V \otimes W}(R_{13}) \rho_{U \otimes V \otimes W}(R_{23}) = \rho_{U \otimes V \otimes W}(R_{23}) \rho_{U \otimes V \otimes W}(R_{13}) \rho_{U \otimes V \otimes W}(R_{12}).$$

Unwinding the definitions, one checks that

$$\begin{aligned} \rho_{U \otimes V \otimes W}(R_{12}) &= R_{U,V} \otimes \text{id}_W, \\ \rho_{U \otimes V \otimes W}(R_{23}) &= \text{id}_U \otimes R_{V,W}, \\ \rho_{U \otimes V \otimes W}(R_{13}) &= \text{the operator acting as } R_{U,W} \text{ on the first and third factors.} \end{aligned}$$

Thus the equality above is precisely the statement that

$$R_{U,V}^{12} R_{U,W}^{13} R_{V,W}^{23} = R_{V,W}^{23} R_{U,W}^{13} R_{U,V}^{12} \quad \text{on } U \otimes V \otimes W,$$

where the superscript indicates on which pair of factors the operator acts.

If we specialize to  $U = V = W$ , we obtain an  $R$ -matrix  $R_{V,V} \in \text{End}(V \otimes V)$  which satisfies the (non-braided) form of the YBE

$$R_{12} R_{13} R_{23} = R_{23} R_{13} R_{12} \quad \text{in } \text{End}(V^{\otimes 3}).$$

Finally, if we pass from  $R$  to the braided version  $\check{R}_{V,V} = P \circ R_{V,V}$ , a short computation shows that  $\check{R}_{V,V}$  satisfies the braid form (4.4):

$$\check{R}_{12} \check{R}_{23} \check{R}_{12} = \check{R}_{23} \check{R}_{12} \check{R}_{23}.$$

We have therefore proved the desired statement.

**Theorem 2.** Let  $(H, R)$  be a quasi-triangular Hopf algebra and  $V$  an  $H$ -module. Then the operator

$$\check{R}_{V,V} = P \circ (\rho_V \otimes \rho_V)(R) \in \text{End}(V \otimes V)$$

satisfies the Yang–Baxter equation on  $V^{\otimes 3}$ . In particular, every representation of a quasi-triangular Hopf algebra produces an  $R$ -matrix solving the YBE.

## 4.5. Hopf algebra symmetry of the spin- $\frac{1}{2}$ chain

We now illustrate how Hopf algebra structures appear naturally in quantum spin chains. We focus on the simplest case, the  $\mathfrak{sl}_2$  symmetry of the spin- $\frac{1}{2}$  chain. For the purposes of this section we do not yet need a non-trivial universal  $R$ -matrix; it is enough to see how the coproduct encodes the global symmetry.

### 4.5.1. The Lie algebra $\mathfrak{sl}_2$ and its enveloping Hopf algebra

Recall that the Lie algebra  $\mathfrak{sl}_2$  is generated by elements  $E, F, H$  with relations

$$[H, E] = 2E, \quad [H, F] = -2F, \quad [E, F] = H.$$

Its universal enveloping algebra  $U(\mathfrak{sl}_2)$  is an associative algebra generated by  $E, F, H$  with these relations, together with all products of  $E, F, H$ .

This algebra carries a natural Hopf algebra structure:

$$\begin{aligned} \Delta(H) &= H \otimes 1 + 1 \otimes H, \\ \Delta(E) &= E \otimes 1 + 1 \otimes E, \\ \Delta(F) &= F \otimes 1 + 1 \otimes F, \end{aligned}$$

extended as an algebra homomorphism to all of  $U(\mathfrak{sl}_2)$ . The counit and antipode are given on generators by

$$\begin{aligned} \varepsilon(H) &= 0, & S(H) &= -H, \\ \varepsilon(E) &= 0, & S(E) &= -E, \\ \varepsilon(F) &= 0, & S(F) &= -F. \end{aligned}$$

Note that  $\Delta$  is *cocommutative*, i.e.  $\Delta^{\text{op}} = \Delta$ , so the “trivial” choice  $R = 1 \otimes 1$  makes  $U(\mathfrak{sl}_2)$  into a quasi-triangular Hopf algebra. This leads to the permutation operator as the associated  $\check{R}$ , which indeed solves the YBE but is physically uninteresting. Non-trivial  $R$ -matrices relevant for integrable spin chains come from *deformations* such as the quantum group  $U_q(\mathfrak{sl}_2)$  or Yangians; we briefly comment on this below.

### 4.5.2. Spin- $\frac{1}{2}$ representation and tensor product chain

The fundamental (spin- $\frac{1}{2}$ ) representation of  $\mathfrak{sl}_2$  is  $V \cong \mathbb{C}^2$  with basis  $|\uparrow\rangle, |\downarrow\rangle$  and action

$$\begin{aligned} H|\uparrow\rangle &= +|\uparrow\rangle, & H|\downarrow\rangle &= -|\downarrow\rangle, \\ E|\uparrow\rangle &= 0, & E|\downarrow\rangle &= |\uparrow\rangle, \\ F|\uparrow\rangle &= |\downarrow\rangle, & F|\downarrow\rangle &= 0. \end{aligned}$$

Up to normalization, these are the usual Pauli matrices:

$$H \sim \sigma^z, \quad E \sim \sigma^+, \quad F \sim \sigma^-.$$

Consider a chain of  $N$  spins, modeled by the Hilbert space

$$\mathcal{H}_N = V^{\otimes N}.$$

The Hopf algebra structure on  $U(\mathfrak{sl}_2)$  tells us how to extend the action of  $\mathfrak{sl}_2$  from a single site to the entire chain: we use the iterated coproduct

$$\Delta^{(N)}: U(\mathfrak{sl}_2) \longrightarrow U(\mathfrak{sl}_2)^{\otimes N},$$

defined recursively by

$$\Delta^{(1)} = \text{id}, \quad \Delta^{(2)} = \Delta, \quad \Delta^{(N)} = (\Delta \otimes \text{id}^{\otimes N-2}) \circ \Delta^{(N-1)}.$$

For example,

$$\Delta^{(N)}(H) = H \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes H \otimes \cdots \otimes 1 + \cdots + 1 \otimes \cdots \otimes 1 \otimes H.$$

The representation of  $U(\mathfrak{sl}_2)$  on  $\mathcal{H}_N$  is then given by

$$\rho_{\mathcal{H}_N}(X) = (\rho_V^{\otimes N})(\Delta^{(N)}(X)), \quad X \in U(\mathfrak{sl}_2).$$

Concretely, this means that the total spin operators are the sums of the single-site spin operators:

$$S_{\text{tot}}^z = \sum_{j=1}^N S_j^z, \quad S_{\text{tot}}^+ = \sum_{j=1}^N S_j^+, \quad S_{\text{tot}}^- = \sum_{j=1}^N S_j^-.$$

#### 4.5.3. Heisenberg Hamiltonian and $\mathfrak{sl}_2$ symmetry

The Hamiltonian of the isotropic Heisenberg (XXX) spin- $\frac{1}{2}$  chain with nearest-neighbor interactions can be written as

$$H_{\text{XXX}} = \sum_{j=1}^{N-1} \left( S_j^x S_{j+1}^x + S_j^y S_{j+1}^y + S_j^z S_{j+1}^z \right),$$

where  $(S^x, S^y, S^z)$  are the standard spin- $\frac{1}{2}$  operators on  $V$  and  $S_j^\alpha$  means that  $S^\alpha$  acts on the  $j$ -th factor and as the identity on all other factors.

The crucial property is that this Hamiltonian is invariant under the global  $\mathfrak{sl}_2$  action:

$$[H_{\text{XXX}}, \rho_{\mathcal{H}_N}(X)] = 0, \quad \forall X \in \mathfrak{sl}_2.$$

From the Hopf algebra viewpoint, this is the statement that each local interaction term  $h_{j,j+1}$  commutes with the coproduct action  $\Delta(X)$  on  $V \otimes V$ , and hence the sum of all local terms commutes with the iterated coproduct action  $\Delta^{(N)}(X)$  on  $V^{\otimes N}$ .

Thus  $U(\mathfrak{sl}_2)$ , with its Hopf algebra structure, provides the algebraic framework for the  $\mathfrak{sl}_2$  symmetry of the spin chain.

#### 4.5.4. Remark: from $U_q(\mathfrak{sl}_2)$ to non-trivial $R$ -matrices

For the integrable structure of the XXX/XXZ chains one replaces  $U(\mathfrak{sl}_2)$  by its  $q$ -deformation  $U_q(\mathfrak{sl}_2)$  or by the Yangian  $Y(\mathfrak{sl}_2)$ . These are quasi-triangular Hopf algebras with non-trivial universal  $R$ -matrices. Evaluating the universal  $R$  in the spin- $\frac{1}{2}$  representation  $V \cong \mathbb{C}^2$  produces a non-trivial  $4 \times 4$   $R$ -matrix  $R_{V,V}$  solving the YBE. This  $R$ -matrix is precisely the one used to build the monodromy and transfer matrices of the spin chain; the associated transfer matrix generates a commuting family of operators whose logarithmic derivative at a special value of the spectral parameter gives back the Hamiltonian  $H_{\text{XXX}}$  (or  $H_{\text{XXZ}}$  in the anisotropic case).

In this way, quasi-triangular Hopf algebras provide the natural algebraic origin of the  $R$ -matrices and Yang–Baxter structures that underlie integrable spin chains.

## Chapter 5

# Appendix: A symbolic two-site experiment

To make the abstract transfer matrix formalism more concrete, it is useful to work out a small example in full detail. In this subsection we describe a symbolic experiment (implemented in MATLAB) that realises the six-vertex / XXZ spin- $\frac{1}{2}$  model on a chain of two sites and checks the basic integrability identities at the level of explicit matrices.

The starting point is a standard trigonometric parametrisation of the six-vertex Boltzmann weights. We introduce a spectral parameter  $u$  and an anisotropy parameter  $y$  and define

$$a(u, y) = \sinh(u + y), \quad b(u) = \sinh(u), \quad c(y) = \sin(y).$$

These are the familiar weights of the XXZ  $R$ -matrix, written here as scalar functions. In addition, one introduces a function  $d(y)$  which later appears in the diagonal part of the Hamiltonian.

On the quantum side, the local Hilbert space is  $V \cong \mathbb{C}^2$ , and we identify

$$s^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad s^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad s^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

the usual spin- $\frac{1}{2}$  operators and the identity on  $V$ . Using Kronecker products, these operators are lifted to the two- and three-fold tensor products  $V^{\otimes 2}$  and  $V^{\otimes 3}$ , so that expressions of the form

$$s^+ \otimes s^- + s^- \otimes s^+, \quad s^z \otimes s^z$$

and their embeddings into different tensor slots can be used to model local XXZ-type interactions on neighbouring sites.

**Lax operators and monodromy.** Using the scalar functions  $a, b, c$  and the lifted spin operators, the experiment constructs two Lax operators  $L_1(u)$  and  $L_2(u)$  acting on  $V^{\otimes 3}$  of the form

$$L_j(u) \sim \frac{1}{2}(a(u, y) + b(u)) \text{Id} + c(y) (s^+ s^- + s^- s^+) + \frac{1}{2}(a(u, y) - b(u)) s^z s^z,$$

embedded into different pairs of tensor factors. These are nothing but the six-vertex  $R$ -matrices written in operator form, placed so that each  $L_j$  couples the auxiliary space to one of the two quantum sites. The two-site monodromy matrix  $M(u)$  is then defined as the ordered product

$$M(u) = L_2(u) L_1(u).$$

**Two-site transfer matrix and Hamiltonian.** In parallel, the code constructs a  $4 \times 4$  transfer matrix  $T(u)$  acting on  $V^{\otimes 2}$  by writing out explicitly the matrix elements in the basis

$$\{ |\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle \}$$

in terms of the weights  $a(u, y)$ ,  $b(u)$  and  $c(y)$ . This  $T(u)$  is the transfer matrix for a chain of length 2 with periodic boundary conditions, built from the same six-vertex data.

Next, a local two-site Hamiltonian  $H$  is introduced as a  $4 \times 4$  matrix with diagonal entries involving  $d(y)$  and off-diagonal entries implementing spin flips,

$$H \sim \begin{pmatrix} d(y) & 0 & 0 & 0 \\ 0 & -d(y) & 2 & 0 \\ 0 & 2 & -d(y) & 0 \\ 0 & 0 & 0 & d(y) \end{pmatrix}.$$

This has precisely the structure of the local XXZ Hamiltonian on two sites.

**Integrability checks.** The main goal of the experiment is to verify, in this simple 2-site setting, the characteristic commutation relations of the transfer matrix formalism:

- First, one computes the commutator

$$[H, T(u)] = HT(u) - T(u)H,$$

and checks symbolically that all entries simplify to zero. This confirms that the Hamiltonian  $H$  is generated by the transfer matrix  $T(u)$  (more precisely, that  $H$  belongs to the commuting family obtained from  $T(u)$  by differentiation in the spectral parameter).

- Second, one introduces an independent spectral parameter  $v$  and constructs a second transfer matrix  $T(v)$  using the same weights but with  $u$  replaced by  $v$ . The commutator

$$[T(u), T(v)] = T(u)T(v) - T(v)T(u)$$

is then computed symbolically and shown to vanish entrywise. This verifies, on this small example, the fundamental commutativity property

$$[T(u), T(v)] = 0 \quad \forall u, v,$$

which underlies the integrability of the model.

**Big picture.** Conceptually, this symbolic computation realises the following programme in the simplest possible nontrivial case:

1. Start from a choice of (trigonometric) six-vertex weights  $a(u, y)$ ,  $b(u)$ ,  $c(y)$ , which define an  $R$ -matrix of six-vertex / XXZ type.
2. Use these weights to construct Lax operators  $L_j(u)$  and the monodromy matrix  $M(u)$  for a chain of length 2.
3. Extract from this data an explicit transfer matrix  $T(u)$  acting on the two-site quantum space  $V^{\otimes 2}$ .
4. Define a candidate local Hamiltonian  $H$  with XXZ structure and verify that:
  - $[H, T(u)] = 0$ , so that  $H$  is part of the commuting family generated by  $T(u)$ ;
  - $[T(u), T(v)] = 0$  for all  $u, v$ , confirming the existence of a one-parameter family of commuting transfer matrices even at this finite size.

This two-site experiment thus serves as a concrete “laboratory” where one can see, at the level of explicit matrices, how the standard ingredients of the transfer matrix formalism —  $R$ -matrix, Lax operators, monodromy, transfer matrix and Hamiltonian — fit together and realise integrability for the six-vertex / XXZ spin- $\frac{1}{2}$  model.

**Remark 7.** *The code is presented in the `ybe_project_tests.py` in my GitHub repo*

## References

### (A) General introductions to integrability and the algebraic Bethe ansatz

- A. L. Retore, “Introduction to classical and quantum integrability,” *J. Phys. A: Math. Theor.* **55** (2022), 173001.
- L. D. Faddeev, “How Algebraic Bethe Ansatz works for integrable models,” in *Quantum Symmetries*, Les Houches Summer School, Session LXIV (1995), 149–219.
- A. Doikou, S. Evangelisti, G. Feverati, N. Karaiskos, “Introduction to Quantum Integrability,” *Int. J. Mod. Phys. A* **25** (2010), 3307–3351.
- J. Lamers, “A pedagogical introduction to quantum integrability with a view towards theoretical high-energy physics,” lecture notes (2015).
- Y. Jiang, “Lecture 2: Algebraic Bethe Ansatz,” lecture notes.
- Course notes: *Notes on Lie groups and Lie algebras for PHYS5030*.

### (B) Yangian symmetry, $R$ -matrices and quantum groups

- V. Chari, A. Pressley, *A Guide to Quantum Groups*, Cambridge University Press, Cambridge, 1995.
- F. Loebbert, “Lectures on Yangian Symmetry,” *J. Phys. A: Math. Theor.* **49** (2016), 323002.
- L. D. Faddeev, “Integrable models, quantum groups, and conformal field theory,” lecture notes.
- P. P. Kulish, N. Yu. Reshetikhin, E. K. Sklyanin, “Yang–Baxter Equation and Representation Theory: I,” *Lett. Math. Phys.* **5** (1981), 393–403.
- M. A. Semenov-Tian-Shansky, “What is a classical  $r$ -matrix?,” *Funct. Anal. Appl.* **17** (1983), 259–272.
- Y. Cheng, M. L. Ge, K. Xue, “Yang–Baxterization of braid group representations,” *Commun. Math. Phys.* **136** (1991), 195–208.
- H. A. Dye, “Unitary Solutions to the Yang–Baxter Equation in Dimension Four,” *Quantum Inf. Process.* **2** (2003), 117–152.
- V. V. Bazhanov, “Trigonometric solutions of triangle equations and classical Lie algebras,” *Phys. Lett. B* **159** (1985), 321–324.

### (C) Quantum inverse scattering, spin chains and concrete models

- E. K. Sklyanin, L. A. Takhtadzhyan, L. D. Faddeev, “Quantum inverse problem method. I,” *Theor. Math. Phys.* **40** (1979), 688–706.
- L. A. Takhtadzhyan, L. D. Faddeev, “The quantum method of the inverse problem and the Heisenberg XYZ model,” *Russ. Math. Surv.* **34** (1979), 11–68.
- F. H. L. Essler, H. Frahm, F. Göhmann, A. Klümper, V. E. Korepin, *The One-Dimensional Hubbard Model*, Cambridge University Press, 2005.
- K. Sogo, M. Uchinami, Y. Akutsu, M. Wadati, “Classification of Exactly Solvable Two-Component Models,” *Prog. Theor. Phys.* **68** (1982), 508–523.
- N. Beisert, “Integrability in Gauge and String Theory: Spin Chains,” ETH Zürich lecture notes (FS 2013); see in particular the section “ $R$ -matrix from Yangian symmetry”.
- L. D. Faddeev, “Quantum Inverse Problem,” original article (Russian); English translation available via the MathNet link given in the text.

**(D) Online discussions and informal sources**

- Physics Stack Exchange, “R-matrix for spin chains,” discussion thread.
- Quora, answer to “What is the intuition for affine Lie algebras (which are, if I understand it correctly, centrally extended loop algebras)?”