Universiteit Utrecht

Utrecht University

Institute for Theoretical Physics
& Mathematical Institute

Master's Thesis

Spin chains and Yangians

Author:
Laurens Stronks
3864502

Supervisors:
Dr. Dirk Schuricht
Dr. Johan van de Leur

June 29, 2016

Abstract

In this thesis we theoretically investigate the close relation between spin chains and Yangians. First, we investigate the Haldane-Shastry model, which symmetry algebra is a representation of the Yangian. We show this in full detail for the spin-$\frac{1}{2}$ model. We also give an argument why a spin-1 chain with Yangian symmetry cannot exist.

After that, we turn to integrable spin chains which are constructed using the quantum inverse scattering method. Using this method, an integrable spin chain can be constructed from each so called R-matrix, which is a solution of the Yang-Baxter equation. Our main result is the construction of an R-matrix that is invariant under the adjoint representation of $SU(n)$, which is found using an intertwining operator for the corresponding Yangian representation. The integrable spin chain that corresponds to this R-matrix is non-Hermitian, making its physical interpretation unclear. These results have been published recently as a preprint under arXiv:1606.02516.

Finally, integrable spin chains with defects are studied. We present the Hamiltonian for the Heisenberg model and the Babujian-Takhtajan model with defects, such that these models keep there integrability while staying Lie algebra invariant.

Contents

Chapter 1

Introduction

Materials in nature often have very complicated structures consisting of different kinds of particles. For instance, molecules are built out of atoms, that themselves consist of electrons, protons and neutrons. In practice it is impossible to exactly solve physical models that describe real-world materials in full detail, for instance due to the high number of particles and the variety of particles and interactions that are present. However, models that are exactly solvable can give enormous insight in various physical phenomena. One of the clearest examples of this is the two-dimensional square lattice Ising model which is the paradigm for phase transitions. Therefore, it is one of the best studied models in statistical physics, from both the theoretical and experimental point of view. Sometimes it is very interesting to study such 'toy' models to look for interesting physical phenomena rather than describing real physical systems. An important subclass of such models is given by the spin lattices, which are lattice models whose sites carry an irreducible representation of a Lie algebra. The name 'spin' is the label for the irreducible representations of the Lie algebra $\mathfrak{su}(2)$.

In this thesis we restrict ourselves to one-dimensional spin lattice models, which are called spin chains. The main interest for such spin chains is describing magnetism in a simple way. The paradigm of a one-dimensional spin chain is the spin-$\frac{1}{2}$ Heisenberg model whose Hamiltonian is

$$H = J \sum_{i=1}^{N} \mathbf{S}_i \cdot \mathbf{S}_{i+1}$$

with $\mathbf{S}$ the vector of spin operators (that equal half the Pauli matrices) and J a constant. The label i distinguishes the N sites. A schematic picture of the model is given in Fig. 1.1. The Heisenberg model is one the best studied spin chains from both the theoretical [19, 41] as well as the experimental perspective [34, 48]. If one wants to solve such a system, the primary task is to find the eigenvalues and eigenstates of the system from which one can compute all thermodynamic quantities. Since these quantities can be measured experimentally, they can be used to compare the theoretical and experimental results on the model, which could of course corroborate other theoretical predictions that may be impossible to check in experiments. Other quantities that can be measured experimentally are so called correlation functions, although it is much harder to find explicit theoretical results for those since they cannot be directly found from the eigenvalues and eigenstates. Even for spin chains, which form a particularly simple subclass of lattice models, it can be very difficult to find exact expressions for the eigenstates and eigenvalues of the model, as this often involves operators on vector spaces of very large dimension. Furthermore, the dimension of the vector spaces grows exponentially with the number

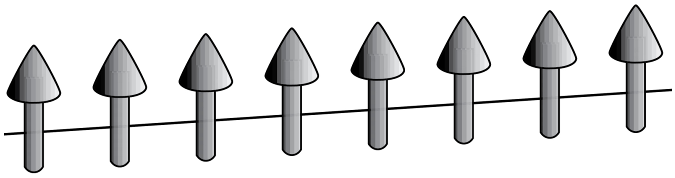

Figure 1.1: A schematic picture of the Heisenberg model, where on each site the arrow depicts the state. This is an edited version of a figure in Ref. [34].

of lattice sites on the spin chain. Therefore, only a limited number of spin chains can be solved analytically. The Heisenberg model was essentially solved by Bethe in 1931 using a method that is now known as the Coordinate Bethe Ansatz [7]. However, this method is only applicable to a select type of spin chains. For instance, if one replaces the spin-$\frac{1}{2}$ operators in the Heisenberg chain by the matrices corresponding to higher spin, the method no longer works. The fact that the Heisenberg can be solved using Bethe's method is a result of one of the key properties of the Heisenberg model, which is its integrability. Loosely speaking this means that there is an infinite set of mutually commuting operators that all commute with the Hamiltonian as well. In fact, the Heisenberg model is the key example of an integrable spin chain and of an integrable system in general.

In this thesis we will often take a broader view than only systems describing 'spin', which are spin chains that are invariant under $\mathfrak{su}(2)$. We will also consider spin chains that are invariant under $\mathfrak{su}(n)$, which allows different kinds of lattice sites in these models. This extends the possibilities from both the physical and mathematical perspective, as the structure of $\mathfrak{su}(n)$ is richer than that of $\mathfrak{su}(2)$. A real physical application of this is given in systems of ultracold atoms, where lasers can be used to simulate these high symmetries [13].

An other method that can be used to solve the Heisenberg model is the Algebraic Bethe Ansatz (ABA) which was developed by Faddeev, Sklyanin and Takhtajian in the early 1980s[19]. The idea of the ABA is to abstractly construct an infinite set of commuting operators that also commutes with the Hamiltonian, and then using the fact that commuting operators have a common set of eigenstates. This infinite set of commuting operators is constructed using the Quantum Inverse Scattering Method (QISM) that was invented around the same time as the ABA. Here, the idea is that a so called R-matrix, which is a solution of the Yang-Baxter equation (YBE), allows for the construction of such a commuting family of operators. Solutions of the YBE play a central role in various other fields of theoretical physics and mathematics such as the theory of quantum groups [12], statistical field theories with factorised scattering [36] or quantum information theory [17, 25]. In the 1980s and 1990s many solutions of the YBE were constructed, see e.g. Refs. [1, 27, 28, 29, 38]. In the QISM every independent solution of the YBE gives essentially rise to a model that can be solved by the ABA, by constructing a Hamiltonian from the infinite set of commuting operators [41].

An example of such a model is the spin-1 integrable spin chain that is now known as the Babujian-Takhtajan model, which Hamiltonian is [3, 47]

$$H = J \sum_{i=1}^{N} \left[\mathbf{S}_i \cdot \mathbf{S}_{i+1} - (\mathbf{S}_i \cdot \mathbf{S}_{i+1})^2 \right].$$

where the vector of operators $\mathbf{S}$ is, up to a constant, a higher spin generalizations of the vector of Pauli-matrices. The fact that this model was found and solved was one of the great achievements of the QISM and the ABA. Various kinds of models have arisen in this way, such as a model for every irreducible representation of $\mathfrak{su}(2)$ [3, 47] and for the fundamental representation of various other Lie algebras [38]. Variations on these models, e.g. spin ladders for tensor product representations or anisotropic models were also found [6, 51]. In fact, it is possible to construct an integrable spin chain for every finite-dimensional irreducible representation of $SU(n)$ [10], although explicit expressions are not known for every representation.

In 1985 Drinfeld discovered an intimate relationship between R-matrices and the Yangian, an associative algebra corresponding to a simple Lie algebra [15, 16]. From the mathematical point of view, the Yangian is a noncocommutative Hopf algebra, that contains (when viewed itself as a Lie algebra with the commutator bracket) the Lie algebra as a Lie subalgebra. Drinfeld proved that every algebra representation of the Yangian leads to a (not necessarily unique) solution of the YBE on this representation, although he did not provide a method to construct these R-matrices. This limitation was partly overcome by Chari and Pressley, who developed a method to construct rational solutions of the YBE for every Lie algebra representation of $\mathfrak{su}(n)$ and the fundamental representation of every Lie algebra [10, 11]. Furthermore, they provided explicit solutions of the YBE on the direct sum of the adjoint representation and the trivial representation for every simple Lie algebra except $\mathfrak{su}(\mathfrak{n})$.

The main result obtained in this thesis is a solution of the YBE that is invariant under the adjoint representation of $SU(3)$ and a conjecture for $SU(n)$. This result has been published recently as a preprint [46]. The R-matrix is found using the method by Chari and Pressley, involving the intertwining operator for Yangian representations. First, this intertwining operator is found using explicit computations for the action of the Yangian generators on various elements of the adjoint representation. The R-matrix can be found straightforwardly from this.

Using the QISM it is possible to construct an integrable spin chain that is invariant under the adjoint representation of $SU(n)$. The adjoint representation of $SU(n)$ turns up in various areas of physics, most notably in Quantum Chromodynamics (QCD), where a special kind of elementary particles called gluons are vectors in the adjoint representation of $SU(3)$. In the so called large-n limit of QCD, these gluons are described by the adjoint representation of $SU(n)$. The spin chain that we construct out of our R-matrix could therefore heuristically be seen as a spin chain of gluons.

However, it turns out that the constructed spin chain is non-Hermitian with nonreal eigenvalues. Therefore, the physical meaning of this spin chain is unclear. We will point out why this spin chain in non-Hermitian in this case en why this does not happen for the adjoint representation of $SU(2)$, which is the spin-1 representation. In that case, one finds the Babujian-Takhtajan model, which is perfectly Hermitian. The fact that the constructed spin chain is non-Hermitian will be a consequence of a property of the adjoint representation of $\mathfrak{su}(n)$.

This thesis is organized as follows. We start with some well-known results on the theory of Lie algebras and its representation theory. In particular, we note some results on the Lie algebra $\mathfrak{su}(3)$ that we will need later on. In Chapter 3 we introduce the Yangian and review some of the key results concerning it. We start with Drinfeld's definition in terms of the generators and generating relations. We point out some useful mathematical properties of the Yangian, such as the comultiplication and the evaluation homomorphism from the Yangian to the universal enveloping algebra, which is a map that only exists for $\mathfrak{su}(n)$. Finally, we note some important properties concerning the representation theory of the Yangian. In particular, the construction of the tensor product representation for the Yangian is shown. Furthermore, we note how one extends a representation of $\mathfrak{su}(n)$ to its Yangian.

In Chapter 4 we study the Haldane-Shastry model. The Haldane-Shastry model is a spin chain for the fundamental representation of $\mathfrak{su}(n)$ with long range interactions that was independently introduced by Haldane and Shastry in 1988 [21, 45]. It is a key example of multiple physical phenomena, such as fractional quantization and anyon statistics [43, 44]. Furthermore, the symmetry algebra of the Haldane-Shastry model is a non-trivial representation of the Yangian as was shown by Haldane et al. [22]. We will explicitly show this in full detail in the case of $n = 2$. Finally, we investigate a possible generalization of the Haldane-Shastry model to a spin-1 model with Yangian symmetry. We give an argument why such a spin chain cannot exist.

We start Chapter 5 with an introduction to the Yang-Baxter equation and we state some key results that can be used to find solutions. Then we present two well-known solutions, corresponding to the two lowest-dimensional non-trivial representations of $\mathfrak{su}(2)$. After that, we try to construct new solutions, starting with the adjoint representation of $\mathfrak{su}(3)$. We present the construction of this solution in full detail. Furthermore, we give a conjectured solution for the adjoint representation of $\mathfrak{su}(n)$ for $n > 3$. This solution is based on formulas we have checked up to $n = 7$. These R-matrices are the main results that are obtained in the thesis. Finally, we investigate Hamiltonians on reducible representations of $\mathfrak{su}(2)$.

Chapter 6 is concerned with one of the main applications of R-matrices in the form of integrable spin chains. We review the construction of integrable spin chains out of solutions of the Yang-Baxter equation. Furthermore, we present the two best-known examples of this construction in the form of the Heisenberg model and the Babujian-Takhtajan model. We will also derive the Bethe equations, which are the most important reason to study these models. The Bethe equations give an abstract but exact description of the eigenvalues and the eigenstates of the model in question. We try to mimic this method for the adjoint representation of $\mathfrak{su}(n)$, using the R-matrix that was obtained in the previous chapter. However, the constructed Hamiltonian turns out to be non-Hermitian. We also tried to construct the Bethe equations, but our approach was unsuccessful. We tried to make sense of our Hamiltonian by considering non-Hermitian operators that can still be interpreted as a Hamiltonian because they have real eigenvalues. However, in our case the Hamiltonian does not have real eigenvalues and therefore its physical meaning is unclear. Finally, we indicate why integrable Hamiltonians on reducible representations have certain unphysical properties.

In Chapter 7 we are concerned with integrable spin chains with defects. These spin chains are constructed using a method that is very similar to those in the previous chapter, but now there are defects on some of the sites. We will be concerned with defects that arise by taking a shift of the R-matrix on some of the sites. We construct a general expression for spin chains with one or two defects, provided some assumptions hold. In the case of one defect, interaction between three sites arise. Two defects behave differently if they are placed on neighbouring sites or not. In particular, defects that are not next to each other give an independent contribution to the Hamiltonian. Therefore, it is possible to construct integrable Hamiltonians with defects at at most half of the sites, which contain at most three-site interactions. Finally, we present the defect Hamiltonians for the Heisenberg and the Babujian-Takhtajan model and compute the Bethe equations for those.

In the final chapter, we summarize all our results that we found in this thesis. We conclude that chapter with a brief outlook in which we mention two openings for further research.

Chapter 2

Preliminaries on Lie algebras

2.1 Introduction

In many areas of physics, Lie algebras turn up naturally. In classical mechanics for instance, they describe a set of variables that close under the Poisson bracket. In quantum mechanics, this Poisson bracket is replaced by the commutator bracket and the Lie algebras describe sets of operators that close under this bracket. In particular, the spin property of a quantum mechanical particle can be described by the representation theory of the simplest non-trivial Lie algebra, $\mathfrak{su}(2)$. In later chapters, we will be interested in particles with such a spin property, but in some cases we will allow a more general set-up where we are interested in the representations of other Lie algebras, in particular $\mathfrak{su}(3)$. In this chapter we will review some facts from the theory of Lie algebras that will be useful in later chapters, in particular when dealing with the Yangian. We start with some general remarks on Lie algebras and then give some specific results in the case of $\mathfrak{su}(3)$, which we will use later on when investigating $\mathfrak{su}(3)$-invariant systems.

2.2 Lie algebras

Let us start with some general facts concerning the theory of Lie algebras. Let $\mathfrak{g}$ be a Lie algebra. We will only be interested in the case where the underlying field is $\mathbb{C}$, so we will always assume this. We recall that $\mathfrak{g}$ is called simple if its only ideals are $\{0\}$ and itself. Furthermore, $\mathfrak{g}$ is called semisimple if it is the direct sum of simple Lie algebras. In particular we will be interested in the simple Lie algebra $\mathfrak{su}(n, \mathbb{C})$, which is (the complexification of) the Lie algebra of the special unitary group $SU(n)$. It can be identified with the the set of complex n-dimensional traceless matrices. We note that there is a natural identification $\mathfrak{su}(n, \mathbb{C}) \simeq \mathfrak{sl}(n, \mathbb{C})$, where the latter is the Lie algebra of the complex special linear group. From now on we omit the $\mathbb{C}$ and write $\mathfrak{su}(n)$ and $\mathfrak{sl}(n)$ interchangeably.

We recall some facts on semisimple Lie algebras. A very useful tool is the root space decomposition of a semisimple Lie algebra $\mathfrak{g}$

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}_\alpha, \tag{2.1}$$

with $\mathfrak{h}$ the Cartan subalgebra, R the root system and $\mathfrak{g}_\alpha$ the root spaces, which are all one-dimensional.

A representation V is called irreducible if $\{0\}$ and V are the only $\mathfrak{g}$-invariant subspaces. For the adjoint representation, the invariant subspaces are precisely the ideals of the Lie algebra.

Therefore the adjoint representation is irreducible for simple Lie algebras. Schur's lemma states that any intertwining map between irreducible representations of a simple Lie algebra is a multiple of the identity. As a consequence, we have the following result [24].

Lemma 2.1. *Let $\mathfrak{g}$ be a simple Lie algebra. If B and $\tilde{B}$ are two nondegenerate symmetric bilinear forms on $\mathfrak{g}$ such that $B([x,y],z) = B(x,[y,z])$ and $\tilde{B}([x,y],z) = \tilde{B}(x,[y,z])$ for all x, y, $z \in \mathfrak{g}$. Then B and $\tilde{B}$ are proportional.*

Proof. By nondegeneracy of the bilinear form, B and $\tilde{B}$ induce isomorphisms φ_B, $\varphi_{\tilde{B}} : \mathfrak{g} \mapsto \mathfrak{g}^*$. We claim that the map $\varphi_{\tilde{B}}^{-1} \circ \varphi_B : \mathfrak{g} \mapsto \mathfrak{g}$ intertwines the adjoint representation. It is sufficient that the maps φ_B and $\varphi_{\tilde{B}}$ are intertwining for the adjoint representation on $\mathfrak{g}$ and the corresponding dual representation on $\mathfrak{g}^*$. To see this we note that

$$\varphi_B(\mathrm{ad}(x)y)(z) = -B([y,x],z) = -B(y,[x,z]) = \varphi_B(y)(-\mathrm{ad}(x)(z)) = \mathrm{ad}^*(x)\varphi_B(y)(z).$$

By Schur's lemma the map $\varphi_{\tilde{B}}^{-1} \circ \varphi_B$ is a multiple of the identity. Therefore we have $\varphi_B = \lambda \varphi_{\tilde{B}}$ for some $\lambda \in \mathbb{C}$, proving the lemma. □

In particular, this result implies that the trace form (defined by $\langle x, y \rangle = \mathrm{tr}(xy)$) and the Killing form (defined by $B(x,y) = \mathrm{tr}(\mathrm{ad}(x)\mathrm{ad}(y))$ are proportional to each other in the case of $\mathfrak{sl}(n)$. From now on we will use the term invariant inner product for a nondegenerate bilinear form B on a Lie algebra satisfying $B([x,y],z) = B(x,[y,z])$, although these are not inner products in a strict mathematical sense, since they lack the positivity criterion.

The following lemma is important in the representation theory of compact Lie algebras, in particular $\mathfrak{su}(n,\mathbb{R})$, as it ensures that for any irreducible finite-dimensional the operators can be chosen anti-Hermitian [4].

Lemma 2.2. *Let G be a compact Lie group and (ϕ, V) an irreducible finite-dimensional representation of G. Then there is an inner product on V such that $\phi(g)$ is a unitary operator for any $g \in G$.*

Proof. First, we fix an arbitrary inner product $\langle\ ,\ \rangle_1$ on V. We define a new inner product on V by

$$\langle v_1, v_2 \rangle = \int_G \langle \phi(x)v_1, \phi(x)v_2 \rangle_1 dx,$$

where $x \in G$, v_1, $v_2 \in V$ and dx is the normalized Haar measure on G. Due to the invariance of the Haar measure, the operators $\phi(y)$ are unitary with respect to this inner product for all element $y \in G$, as is shown by

$$\langle \phi(y)v_1, \phi(y)v_2 \rangle = \int_G \langle \phi(yx)v_1, \phi(yx)v_2 \rangle_1 dx = \int_G \langle \phi(yx)v_1, \phi(yx)v_2 \rangle_1 d(yx) = \langle v_1, v_2 \rangle.$$

Therefore, the inner product $\langle\ ,\ \rangle$ indeed makes the representation V unitary. □

For the corresponding real Lie algebra $\mathfrak{g}_\mathbb{R}$ of the compact Lie group G, the operators are all anti-Hermitian with respect to this inner product. The inner product for which the Lie group representation is unitary, is unique up to a scalar. This can be shown using a similar argument as the one used in the proof of Lemma 2.1. If we use an inner product on a finite-dimensional irreducible representation of a compact Lie algebra, which are in one to one correspondence to those representations of the Lie group, it will always be this inner product.

For calculations in Lie algebras it is often useful to use the index notation and the Einstein summation convention, meaning that repeated indices are summed over. We can fix an arbitrary basis $\{T^a\}$ for $\mathfrak{g}$. The structure constants f^{abc} are defined by $[T^a, T^b] = f^{abc}T^c$. These structure constants are antisymmetric in the first two indices for an arbitrary basis of $\mathfrak{g}$, since the Lie bracket is antisymmetric. The Jacobi identity in terms of the structure constants reads

$$f^{abe}f^{cde} + f^{bce}f^{ade} + f^{cae}f^{bde} = 0. \tag{2.2}$$

For $\mathfrak{su}(n)$ the trace form is nondegenerate. Therefore, it is possible to fix a basis $\{S^a\}$ which is orthonormal with respect to the trace form (and therefore orthogonal with respect to any invariant inner product). The following lemma is extremely useful for calculations concerning $\mathfrak{su}(n)$.

Lemma 2.3. *Let $\{S^a\}$ be a basis for $\mathfrak{su}(n)$ that is orthonormal with respect to the trace form. Then in this basis the structure constants are antisymmetric in all indices.*

Proof. This proof follows immediately from taking the trace of the relation

$$\left[S^a, S^b\right] S^c = f^{abd}S^dS^c$$

which leads to the expression $f^{abc} = \operatorname{tr}\left(\left[S^a, S^b\right] S^c\right)$, which is totally antisymmetric due to the cyclicity of the trace. □

It is immediate from the proof of the lemma that rescaling a basis which is orthonormal with respect to the trace form only amounts to a rescaling of all the structure constants. In the literature, no consistent normalization for the basis vectors is used. For the case $n = 2$ for example such a basis is given by the Pauli matrices, which are normalized such that $\operatorname{tr}\left(\sigma^a\sigma^b\right) = 2\delta^{ab}$. For $n = 3$, the Gell-Mann matrices suffice, which are given by

$$\begin{aligned}
\lambda^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda^2 &= \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda^3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda^4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda^5 &= \begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix} & \lambda^6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
\lambda^7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} & \lambda^8 &= \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{2.3}$$

The normalization is again such that $\operatorname{tr}\left(\lambda^a\lambda^b\right) = 2\delta^{ab}$. It is possible to generalize this construction to arbitrary n [8], but we will not present this here. If we choose a basis $\{S^a\}$ that is orthonormal with respect to the trace form, there is also a convenient expression for the anticommutator of two operators in the fundamental (or defining) representation of $\mathfrak{su}(n)$,

$$\{S^a, S^b\} = \frac{2}{n}\delta^{ab}\mathbf{1} + d^{abc}S^c. \tag{2.4}$$

The d^{abc} are the completely symmetric d-symbols, which are given by $d^{abc} = \operatorname{tr}\left(\left\{S^a, S^b\right\} S^c\right)$. In the case of $n = 2$ the d-symbols are all 0, since distinct Pauli matrices anticommute. We note that such an expression is only possible in the fundamental representation. For other representations of $\mathfrak{su}(n)$, the algebra of the operators (together with the identity) does not close under multiplication.

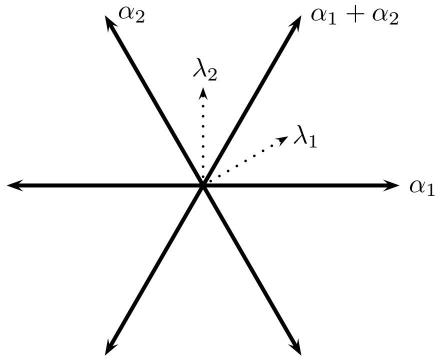

Figure 2.1: The root system of type A_2

2.3 Representation theory of $\mathfrak{su}(3)$

The representation theory of simple Lie algebras is well-known. In this section we will take a look at the case of the simple Lie algebra $\mathfrak{su}(3) = \{x \in \mathrm{Mat}(3, \mathbb{C}) | \mathrm{tr}(x) = 0\}$, which has dimension 8. The Cartan subalgebra has dimension 2 and consists of the diagonal matrices in $\mathfrak{su}(3)$. The root system R contains two simple roots which we label as α_1 and α_2. We can choose α_1 and α_2 to be unit vectors, which implies $(\alpha_1, \alpha_2) = -\frac{1}{2}$. The third positive root equals $\alpha_1 + \alpha_2$. This root system is of type A_2 and depicted in Fig. 2.1. The fundamental weights of $\mathfrak{su}(3)$ are $\lambda_1 = \frac{1}{3}(2\alpha_1 + \alpha_2)$ and $\lambda_2 = \frac{1}{3}(\alpha_1 + 2\alpha_2)$ and satisfy $(\lambda_i, \alpha_j) = \frac{1}{2}\delta_{ij}$. We denote the representation of $\mathfrak{su}(3)$ with highest weight vector $n_1\lambda_1 + n_2\lambda_2$ by $V(n_1, n_2)$. This representation is finite-dimensional if and only if n_1 and n_2 are nonnegative integers. We note the following result.

Lemma 2.4. *Let n_1, n_2 be nonnegative integers. Then the dimension of the representation $V(n_1, n_2)$ is given by* $\dim V(n_1, n_2) = \frac{1}{2}(n_1 + 1)(n_2 + 1)(n_1 + n_2 + 2)$.

Proof. By Weyl's dimension formula, we have

$$\dim V(n_1, n_2) = \prod_{\alpha \in R^+} \frac{(n_1\lambda_1 + n_2\lambda_2 + \delta, \alpha)}{(\delta, \alpha)},$$

where δ is half the sum of the positive roots, which equals $\alpha_1 + \alpha_2$ in this case. We note that

$$(n_1\lambda_1 + n_2\lambda_2 + \delta, \alpha_1) = \left(\frac{1}{2}n_1 + (\alpha_1 + \alpha_2, \alpha_1)\right) = \frac{1}{2}(n_1 + 1),$$

and similarly

$$(n_1\lambda_1 + n_2\lambda_2 + \delta, \alpha_2) = \frac{1}{2}(n_2 + 1),$$
$$(n_1\lambda_1 + n_2\lambda_2 + \delta, \alpha_1 + \alpha_2) = \frac{1}{2}(n_1 + n_2 + 2).$$

Therefore we have

$$\dim V(n_1, n_2) = \left(\frac{\frac{1}{2}(n_1 + 1)}{\frac{1}{2}}\right)\left(\frac{\frac{1}{2}(n_2 + 1)}{\frac{1}{2}}\right)\left(\frac{\frac{1}{2}(n_1 + n_2 + 2)}{1}\right) = \frac{1}{2}(n_1+1)(n_2+1)(n_1+n_2+2).$$

$\square$

The fundamental weight of the adjoint representation is the unique maximal root, which is $\alpha_1 + \alpha_2 = \lambda_1 + \lambda_2$ in this case, so the adjoint representation of $\mathfrak{su}(3)$ is (isomorphic to) $V(1,1)$, which has dimension 8, precisely the dimension of $\mathfrak{su}(3)$. To investigate the adjoint representation of $\mathfrak{su}(3)$, it will be useful to introduce another basis of $\mathfrak{su}(3)$ that preserves the root space decomposition of Eq. (2.1). Let e_{ij} be the (3×3)-matrix with only zeroes, except for one 1 on the $(i,j)^{\text{th}}$ entry. We define $x_1 = e_{12}$, $x_2 = e_{23}$, $x_{12} = [x_1, x_2] = e_{13}$, $y_1 = e_{21}$, $y_2 = e_{32}$, $y_{12} = [y_1, y_2] = e_{31}$, $h_1 = [x_1, y_1] = e_{11} - e_{22}$ and $h_2 = [x_2, y_2] = e_{22} - e_{33}$ and we note that these eight elements form a basis of $\mathfrak{su}(3)$ that indeed preserves the decomposition in Eq. (2.1). We label the simple roots such that $x_1 \in \mathfrak{g}_{\alpha_1}$. The basis B is ordered in the following way:

$$B = \{x_{12}, x_1, x_2, h_1, h_2, y_2, y_1, y_{12}\}.$$

We can now straightforwardly compute the matrices belonging to these elements of $\mathfrak{su}(3)$ when acting in the adjoint representation. For instance, x_{12} acts in the above basis as

$$\mathrm{ad}\ x_{12} = \begin{pmatrix} 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The action of the other elements of $\mathfrak{su}(3)$ can be computed similarly.

Finally, we note the relation between the Gell-Mann matrices T^a and the matrices in the above basis of $\mathfrak{su}(3)$. We have

$$T^1 = x_1 + y_1,\ T^2 = -I(x_1 - y_1),\ T^3 = h_1,\ T^4 = x_{12} + y_{12},$$
$$T^5 = -I(x_{12} - y_{12}),\ T^6 = x_2 + y_2,\ T^7 = -I(x_2 - y_2),\ T^8 = \frac{1}{\sqrt{3}}h_1 + \frac{2}{\sqrt{3}}h_2.$$

These relations will come in handy when investigating the representations of the Yangian of $\mathfrak{su}(3)$.

2.4 Summary

In this chapter we have reviewed some aspects on the theory of Lie algebras and its representation theory. We showed that there is a unique invariant inner product on a simple Lie algebra, up to normalization. Furthermore, we introduced the structure constants and d-symbols that will be useful for calculations in later chapters. In particular, we showed that in the right bases the structure constants are totally antisymmetric. In the next chapter we consider the Yangian, which is an algebra corresponding to a simple Lie algebra. There we will use the results that we have presented here. Later, we investigate the Yang-Baxter equation on different representations of $\mathfrak{su}(n)$, in particular the adjoint representation of $\mathfrak{su}(3)$ in much detail, and we will use the facts on the representation theory of $\mathfrak{su}(3)$ that we have presented here.

Chapter 3

The Yangian

3.1 Introduction

Symmetries play a large role in physics. For instance, in classical mechanics they can be used to simplify a problem by reducing the dimensions of the phase space. In quantum mechanics, a symmetry is related to a set of commuting observables. In such a case, this set can often be described by a Lie algebra. The Heisenberg model, for example, is invariant under the action of $\mathfrak{su}(2)$. In some cases, however, the Lie algebra symmetry is not the whole story. Then there is a larger algebra of operators commuting with the Hamiltonian of the system. In this chapter we investigate the Yangian, which can play the role of such a large symmetry algebra.

In the preceding chapter we have reviewed some properties of Lie algebras, which will helpful to us when investigating the Yangian. The Yangian was introduced by Drinfeld in 1985 and named in honor of C.N. Yang [15]. The Yangian is an associative algebra that, when viewed as a Lie algebra with the commutator bracket, contains a simple Lie algebra as a Lie subalgebra. It plays the role of a symmetry algebra in numerous areas in physics as an extension of a Lie algebra symmetry of a physical model. This is for instance the case in $N = 4$ supersymmetric Yang-Mills theory and multiple models in condensed matter theory [18, 20, 22]. In the next chapter, we will investigate the Haldane-Shastry model as an example of this.

In the first section we will define the Yangian, following Chari and Pressley [10]. We will state some of its properties, most of which are due to Drinfeld [15]. In the second section we look at the representation theory of the Yangian, in particular the connection with the representation theory of the underlying Lie algebra. This work is due to Drinfeld, Chari and Pressley [10, 11, 15]. The representations of the Yangian have numerous applications, e.g. determining the ground state and the first excited states of systems with Yangian symmetry [44] and the construction of solutions of the Yang-Baxter equation [10]. We will review the latter construction in Chapter 5.

3.2 Definition and properties

We fix a base $\{I^a\}$ for $\mathfrak{g}$ which is orthonormal with respect to the trace form. For instance, for $\mathfrak{su}(2)$ we could use $\{\frac{1}{\sqrt{2}}\sigma^a\}$, where σ^a are the Pauli matrices. The structure constants in this basis are defined by $[I^a, I^b] = c^{abc}I^c$. With this in mind, we can define the Yangian.

Definition 3.1. *(Yangian) The Yangian $Y(\mathfrak{g})$ is the associative algebra generated by the elements*

I^a and J^a with defining relations

$$[I^a, I^b] = c^{abc} I^c, \;\; [I^a, J^b] = c^{abc} J^c, \tag{3.1}$$

$$\Big[J^a, [J^b, I^c]\Big] - \Big[I^a, [J^b, J^c]\Big] = \mathfrak{c}^{abcdef}\{I^d, I^e, I^f\}, \tag{3.2}$$

$$\Big[[J^a, J^b], [I^r, J^s]\Big] + \Big[[J^r, J^s], [I^a, J^b]\Big] = (\mathfrak{c}^{abcdef} c^{rsc} + \mathfrak{c}^{rscdef} c^{abc})\{I^d, I^e, J^f\}, \tag{3.3}$$

where $\mathfrak{c}^{abcdef} = \frac{1}{24} c^{adp} c^{beq} c^{cfr} c^{pqr}$ *and* $\{X^a, X^b, X^c\} = \sum_{\sigma \in S_3} X^{\sigma(a)} X^{\sigma(b)} X^{\sigma(c)}$.

Drinfeld called the second and third relation 'terrific' [15], because of their right-hand sides which are extremely difficult to deal with as a result of the triple products. These remarkable relations follow from defining a map called the comultiplication $\Delta : \mathfrak{g} \to \mathfrak{g} \otimes \mathfrak{g}$ on the generators and demanding that the map is a homomorphism. This puts the structure of a Hopf algebra on $Y(\mathfrak{g})$ [15]. The action of this map on the generators is given by

$$\Delta(x) = x \otimes 1 + 1 \otimes x, \tag{3.4}$$

$$\Delta\left(J\left(x\right)\right) = J(x) \otimes 1 + 1 \otimes J(x) + \frac{1}{2}\left[x \otimes 1, C\right], \tag{3.5}$$

where $C = I^a \otimes I^a$ is the Casimir operator.

We now summarize some of the properties of the Yangian for further use. First of all, we note that the defining relations for the Yangian depend on the choice of invariant inner product. However, one can easily show using Lemma 2.1 that the resulting Yangian is always isomorphic [10, 15].

Drinfeld pointed out that for $\mathfrak{sl}(n)$ $(n > 2)$ the relation Eq. (3.3) follows from Eqs. (3.1) and (3.2) [15, 39]. For $\mathfrak{sl}(2)$ Eq. (3.2) follows from Eq. (3.1). This is in fact easy to establish. For $\mathfrak{sl}(2)$ the structure constants in the above basis are given by $c^{abc} = \mathrm{i}\sqrt{2}\varepsilon^{abc}$, where ε^{abc} is the three-dimensional Levi-Civita symbol. Therefore the left hand side of Eq. (3.2) is equal to

$$\mathrm{i}\sqrt{2}\left(\varepsilon^{bcd}[J^a, J^d] + \varepsilon^{cad}[J^b, J^d] + \varepsilon^{abd}[J^c, J^d]\right).$$

This expression is completely antisymmetric in the indices a, b and c. Therefore, it can only be nonzero if all indices are different, for instance if $(abc) = (123)$. In that case the equation reads

$$\mathrm{i}\sqrt{2}\left(\varepsilon^{23d}[J^1, J^d] + \varepsilon^{31d}[J^2, J^d] + \varepsilon^{12d}[J^3, J^d]\right) = \mathrm{i}\sqrt{2}\left([J^1, J^1] + [J^2, J^2] + [J^3, J^3]\right),$$

which certainly vanishes. Therefore the left-hand side is zero in all cases. The right-hand side is equal to

$$\begin{aligned} \frac{1}{6}\varepsilon^{adi}\varepsilon^{bej}\varepsilon^{cfk}\varepsilon^{ijk}\{I^d, I^e, I^f\} &= \frac{1}{6}\varepsilon^{adi}\varepsilon^{bej}\left(\delta^{ci}\delta^{fj} - \delta^{cj}\delta^{fi}\right)\{I^d, I^e, I^f\} \\ &= \frac{1}{6}\left(\varepsilon^{adc}\varepsilon^{bef} - \varepsilon^{adf}\varepsilon^{bec}\right)\{I^d, I^e, I^f\} = 0, \end{aligned}$$

because the triple product $\{I^d, I^e, I^f\}$ is symmetric in all of its indices. So both the left hand and the right hand side always vanish in the case of $\mathfrak{sl}(2)$.

For each $\lambda \in \mathbb{C}$ we can define an automorphism of the Yangian by [15]

$$\tau_\lambda : I^a \mapsto I^a, \;\; J^a \mapsto J^a + \lambda I^a. \tag{3.6}$$

This automorphism can be used to generate a whole family of representations of the Yangian out of one, as we will see in the next section. To see why this is an automorphism we replace J^a by $\tilde{J}^a = J^a + \lambda I^a$ in Eq. (3.2),

$$\begin{aligned}
\left[\tilde{J}^a, \left[\tilde{J}^b, I^c\right]\right] - \left[I^a, \left[\tilde{J}^b, \tilde{J}^c\right]\right] =& f^{bcd}\left[\tilde{J}^a, \tilde{J}^d\right] + f^{cad}\left[\tilde{J}^b, \tilde{J}^d\right] + f^{abd}\left[\tilde{J}^c, \tilde{J}^d\right] \\
=& \left[J^a, \left[J^b, I^c\right]\right] - \left[I^a, \left[J^b, J^c\right]\right] \\
&+ 2\lambda\left(f^{bcd}f^{ade} + f^{cad}f^{bde} + f^{abd}f^{cde}\right)\tilde{J}^e \\
&+ \lambda^2\left(f^{bcd}f^{ade} + f^{cad}f^{bde} + f^{abd}f^{cde}\right)I^e \\
=& \left[J^a, \left[J^b, I^c\right]\right] - \left[I^a, \left[J^b, J^c\right]\right]
\end{aligned}$$

as a result of the Jacobi identity. We see that Eq. (3.2) is indeed invariant under τ_λ. A similar calculation shows that I^a and $\tilde{J}^a$ satisfy Eq. (3.3) if I^a and J^a satisfy Eqs. (3.2) and (3.3).

Finally, we note that in the case of $\mathfrak{su}(n)$, there is a nontrivial homomorphism from the Yangian to the universal enveloping algebra $U(\mathfrak{su}(n))$, which is called the evaluation homomorphism. Its explicit action on the generators is given by

$$\varepsilon(x) = x, \tag{3.7}$$

$$\varepsilon(J(x)) = \frac{1}{4}\mathrm{tr}(x\{I^a, I^b\})I^a I^b. \tag{3.8}$$

In the case of $n = 2$, the expression $\mathrm{tr}(x\{I^a, I^b\})$ always vanishes. To see this, we can take the I^a to be proportional to the Pauli matrices. Then the anticommutator gives a nonzero expression if and only if $a = b$ in which case it is proportional to the identity matrix. Since the matrices in $\mathfrak{su}(2)$ are all traceless, the right hand side of (3.8) vanishes in all cases. For $n > 2$, one can write $x = v^a I^a$ to rewrite $\varepsilon(J(x))$ as

$$\varepsilon(J(x)) = \frac{1}{4}v^a d^{abc} I^b I^c,$$

which is a useful formula for computations.

3.3 Representations of the Yangian

In this section we will look at representations of the Yangian and in particular of tensor products of such representations. The comultiplication Δ will play a vital role in this construction of the tensor product representation. The representation theory of the Yangian will become important when studying the Yang-Baxter equation in later chapters.

Let A be an algebra over $\mathbb{C}$. A representation (ϕ, V) is a pair of a vector space V (over $\mathbb{C}$) and an algebra homomorphism $\phi : A \to GL(V)$. For physical applications it is often very useful to construct the tensor product of two representations, i.e. a representation $(\psi, V_1 \otimes V_2)$ that can be constructed out of two representations (ϕ_1, V_1) and (ϕ_2, V_2) in a natural way. In general, there is no way to do this as there are too many restrictions on ψ. For instance, the map $x \mapsto \psi(x) = \phi_1(x) \otimes \phi_2(x)$ is not linear, while $x \mapsto \psi(x) = \phi_1(x) \otimes 1 + 1 \otimes \phi_2(x)$ is not multiplicative. If A is a Hopf algebra there is a way to overcome this using the comultiplication Δ, which is an algebra homomorphism from A to $A \otimes A$. We first define the following homomorphism

$$(\phi_1, \phi_2) : A \otimes A \to GL(V_1) \otimes GL(V_2),$$

where ϕ_i acts on the i^{th} copy of A. Therefore, the map $(\phi_1, \phi_2) \circ \Delta$ is an algebra homomorphism from A to $GL(V_1) \otimes GL(V_2) \simeq GL(V_1 \otimes V_2)$. This is what we will call the tensor product of two representations of the Yangian. From now on we shortly write V for a representation (ϕ, V) if no confusion can arise.

Let $Y(\mathfrak{sl}(n))$ be the Yangian of a Lie algebra $\mathfrak{sl}(n)$. As we have seen, $Y(\mathfrak{sl}(n))$ carries the structure of a Hopf algebra and there is a well-defined notion of the representation $V_1 \otimes V_2$. Let V be an irreducible representation of the Lie algebra $\mathfrak{sl}(n)$. Using the evaluation homomorphism ε from Eq. (3.7), we can extend ϕ to a representation of the Yangian $\phi \circ \varepsilon$. This induced representation is certainly not unique. Using the one-parameter family of automorphisms τ_λ from Eq. (3.6), we can in fact construct a one-parameter family of irreducible representations of the Yangian corresponding to a single irreducible representation of the Lie algebra. Since τ_λ acts by the identity on $\mathfrak{sl}(n)$, all representations $\phi \circ \varepsilon \circ \tau_\lambda$ extend ϕ. This representation will be denoted by V_λ. It is clear that if V is irreducible, so is V_λ. The construction of this family of representations will be used a lot when finding solutions of the Yang-Baxter equation.

3.4 Summary

In this chapter we have studied the Yangian, which is an associative algebra corresponding to a simple Lie algebra. The definition of the Yangian that we presented was given in terms of a set of generators and relations, which extend the relations of the Lie algebra generators. After that we have reviewed some of its vital properties. These include the evaluation homomorphism in Eq. (3.7) in the case of $\mathfrak{sl}(n)$ and the one-parameter family of automorphims of the Yangian in Eq. (3.6). As we have seen, these maps play crucial roles in the representation theory of the Yangian, since they can be used to extend every Lie algebra representation to a one-parameter family of representations of the corresponding Yangian. In the next chapter, we will investigate the Haldane-Shastry model, which symmetry algebra is a representation of the Yangian.

Chapter 4

The Haldane-Shastry model

4.1 Introduction

In the preceding chapter we have seen the definition of the Yangian and investigated its properties. Here we will investigate the Haldane-Shastry (HS) model, which is a model for a $\mathfrak{su}(n)$ spin chain. It is well known that this model has a very large symmetry algebra, which is a representation of the Yangian, if all operators act in the fundamental representation [22]. We will explicitly show this for the case of $\mathfrak{su}(2)$. Using this large symmetry algebra, it is possible to give an exact description of the ground state and the first excited states of the model [44]. Furthermore, we will investigate a possible generalization of the Haldane-Shastry model to other representations of $\mathfrak{su}(2)$. We will present a general argument to exclude such a possibility of a spin chain with Yangian symmetry in the case of spin-1.

4.2 Definition of the model

In 1988 Haldane and Shastry independently introduced a model for a long range spin chain which later became known as the Haldane-Shastry model [21, 45]. It is also known as the $1/r^2$-model because the interactions are of this type. The model is one of the rare examples of a one-dimensional system in which anyon statistics and fractional quantization occur [44]. The one-dimensional model has periodic boundary conditions and can be conveniently embedded into the complex plane by mapping all sites to positions on the unit circle. To be precise, one has

$$\eta_j = \exp\left(\frac{2\pi \mathrm{i}}{N} j\right)$$

as the new locations for the N sites. Under this embedding, the Hamiltonian for the system is given by

$$H_{\mathrm{HS}} = \frac{1}{2}\left(\frac{2\pi}{N}\right)^2 \sum_{i \neq j} \frac{\mathbf{S}_i \cdot \mathbf{S}_j}{|\eta_i - \eta_j|^2}, \tag{4.1}$$

where $\mathbf{S}_j$ is the vector of spin operators S^a acting on the site η_j and can be expressed in terms of the Lie algebra generators that act on the sites. Originally, the model was introduced as a model for spin-$\frac{1}{2}$ particles, but when all $\mathfrak{su}(2)$ operators were replaced by $\mathfrak{su}(n)$ operators acting in the fundamental representation, a lot of properties of the model remain intact. In particular, the large symmetry algebra of the $\mathfrak{su}(n)$ model is a representation of the Yangian. In the case of $\mathfrak{su}(2)$ and spin-$\frac{1}{2}$ this fact can be used to find the ground state and elementary excitations of the model, using the representation theory of the Yangian of $\mathfrak{sl}(2)$ [44].

We now take a closer look at the symmetry algebra of the HS model. First of all, the Hamiltonian commutes with the total spin $\mathbf{S} = \sum_j \mathbf{S}_j$, i.e. we have

$$[H_{\mathrm{HS}}, \mathbf{S}] = 0. \tag{4.2}$$

The total spin satisfies $[S^a, S^b] = f^{abc} S^c$, where f^{abc} are the structure constants of $\mathfrak{su}(n)$. As a result of this, the above commutation relation implies that the Hamiltonian is $\mathfrak{su}(n)$-invariant. The Hamiltonian also commutes with the rapidity operator $\mathbf{\Lambda}$, which components are defined as

$$\Lambda^a = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}} \frac{\eta_i + \eta_j}{\eta_i - \eta_j} f^{abc} S_i^b S_j^c, \quad [H_{\mathrm{HS}}, \Lambda^a] = 0. \tag{4.3}$$

Under $\mathfrak{su}(n)$ transformations, Λ^a transforms as a vector

$$[S^a, \Lambda^b] = f^{abc} \Lambda^c. \tag{4.4}$$

One can check that S^a and Λ^a together satisfy the defining relations of the Yangian, as we will do in the case of $\mathfrak{su}(2)$ in Section 4.4. Therefore, the symmetry algebra of the HS model is a finite-dimensional representation of the Yangian.

4.3 Symmetry algebra of the Haldane-Shastry model

In this section we will explicitly show that the operators $\mathbf{S}$ and $\mathbf{\Lambda}$ commute with the Hamiltonian of the HS model. Before moving on to the calculations, we first list some preliminaries which we later refer to. We assume that all spin operators act in the fundamental representation of $\mathfrak{su}(n)$. We fix a basis $\{T^a\}$ of traceless Hermitian matrices of $\mathfrak{su}(n)$, such that $\mathrm{tr}(T^a T^b) = \frac{1}{n}\delta^{ab}$ and define our spin operators to be $S_i^a = \frac{1}{2} I \otimes I \cdots I \otimes T^a \otimes I \cdots \otimes I$, where T^a is in the i^{th} place and I is the identity operator. Therefore, we have the following relation for products of spin operators acting on the same site:

$$S_i^a S_i^b = \frac{1}{2} f^{abc} S_i^c + \frac{1}{2} d^{abc} S_i^c + \frac{1}{2n}\delta^{ab}. \tag{4.5}$$

Here, f^{abc} are the completely antisymmetric structure constants of the Lie algebra, while d^{abc} are the symmetric d-symbols. We will throughout this chapter always use the convention that the lower indices refer to the sites, while the upper indices label the different basis vectors of $\mathfrak{su}(n)$. The Einstein summation convention will be used, but only for the upper indices. Eq. (4.5) immediately leads to the following (anti-)commutation relations for the spin operators

$$[S_i^a, S_j^b] = \delta_{ij} f^{abc} S_i^c, \tag{4.6}$$

$$\{S_i^a, S_i^b\} = \frac{1}{n}\delta^{ab} + d^{abc} S_i^c. \tag{4.7}$$

Of course, the Jacobi identity for structure constants holds: $f^{abc} f^{cde} + f^{dbc} f^{ace} + f^{ebc} f^{adc} = 0$. Using the identity $[\{T^a, T^b\}, T^c] + [\{T^b, T^c\}, T^a] + [\{T^c, T^a\}, T^b] = 0$, which can easily be proved by expanding the left hand side, one can also derive a Jacobi-like relation between the structure constants and the d-symbols:

$$f^{abc}d^{cde} + f^{dbc}d^{ace} + f^{ebc}d^{adc} = 0. \tag{4.8}$$

Additionally, we prove the following relation: $f^{abc}f^{abd} = C\delta^{cd}$, where C is a constant. To prove it, we use Lemma 2.1. The Killing form satisfies the necessary conditions and in components it is equal to $B^{cd} = -f^{abc}f^{abd}$. Therefore it is proportional to the trace form, which is given by $\frac{1}{n}\delta^{cd}$, due to our choice of basis.

Furthermore, we add some results about complex numbers. In the following, we write $\eta_j = \exp(2\pi \mathrm{i} j/N)$ and abbreviate $\omega_{ij} = (\eta_i + \eta_j)/(\eta_i - \eta_j)$ and $t_{ij} = |\eta_i - \eta_j|^{-2}$. First of all, we have the following identity [30]

$$\begin{aligned}
t_{ij}(\omega_{jk} - \omega_{ik}) &= \frac{1}{|\eta_i - \eta_j|^2}\left(\frac{\eta_j + \eta_k}{\eta_j - \eta_k} - \frac{\eta_i + \eta_k}{\eta_i - \eta_k}\right) \\
&= -\frac{\eta_i\eta_j}{(\eta_i - \eta_j)^2}\frac{(\eta_j + \eta_k)(\eta_i - \eta_k) - (\eta_i + \eta_k)(\eta_j - \eta_k)}{(\eta_j - \eta_k)(\eta_i - \eta_k)} \\
&= -\frac{\eta_i\eta_j}{(\eta_i - \eta_j)^2}\frac{\eta_i\eta_k - \eta_k\eta_j + \eta_i\eta_k - \eta_k\eta_j}{(\eta_j - \eta_k)(\eta_i - \eta_k)} \\
&= \frac{2\eta_i\eta_j\eta_k}{(\eta_i - \eta_j)(\eta_j - \eta_k)(\eta_k - \eta_i)},
\end{aligned} \tag{4.9}$$

where we note that this expression is antisymmetric in all indices. The following identities are essential as well,

$$\sum_{\substack{i=1 \\ i\neq j}}^{N} t_{ij}\omega_{ji} = 0 \tag{4.10}$$

$$\sum_{\substack{i=1 \\ i\neq j}}^{N} \omega_{ij} = 0. \tag{4.11}$$

The proof of these two identities will be given in the appendix to this chapter.

Finally we note the following expression, which is true in any associative algebra:

$$[AB, CD] = A[B, C]D + AC[B, D] + [A, C]DB + C[A, D]B.$$

Since we have introduced all the relations necessary for the proofs of Eq. (4.2) and Eq. (4.3), we can start with the commutation relation of the Hamiltonian and the total spin operator [43],

$$\begin{aligned}
[H_{\mathrm{HS}}, S^a] &= \left[\frac{1}{2}\left(\frac{2\pi}{N}\right)^2 \sum_{i\neq j} \frac{S_i^b S_j^b}{|\eta_i - \eta_j|^2}, \sum_{k=1}^{N} S_k^a\right] \\
&= \frac{1}{2}\left(\frac{2\pi}{N}\right)^2 \sum_{\substack{i,j,k \\ i\neq j}} \frac{\delta_{ik} f^{bac} S_i^c S_j^b + \delta_{jk} f^{bac} S_i^b S_j^c}{|\eta_i - \eta_j|^2} \\
&= \frac{1}{2}\left(\frac{2\pi}{N}\right)^2 \sum_{i\neq j} \frac{f^{bac}(S_i^c S_j^b + S_i^b S_j^c)}{|\eta_i - \eta_j|^2} \\
&= 0
\end{aligned} \tag{4.12}$$

due to the antisymmetry of the structure constants.

The fact that the rapidity operator transforms as a vector under $\mathfrak{su}(n)$ transformations can also be checked easily, using the Jacobi identity [43]

$$\begin{aligned}
[S^a, \Lambda^b] &= \frac{1}{2} \sum_{\substack{i,j,k=1 \\ j\neq k}}^{N} \omega_{jk} f^{bcd} [S_i^a, S_j^c S_k^d] \\
&= \frac{1}{2} \sum_{\substack{i,j,k=1 \\ j\neq k}}^{N} \omega_{jk} f^{bcd} \left(f^{ace} \delta_{ij} S_i^e S_k^d + f^{ade} \delta_{ik} S_j^c S_i^e \right) \\
&= \frac{1}{2} \sum_{\substack{i,j,k=1 \\ j\neq k}}^{N} \omega_{jk} f^{bcd} \left(f^{ace} \delta_{ij} S_i^e S_k^d + f^{ace} \delta_{ij} S_k^d S_i^e \right) \\
&= \frac{1}{2} \sum_{\substack{i,k=1 \\ i\neq k}}^{N} \omega_{ik} f^{bcd} \left(f^{ace} S_i^e S_k^d + f^{ace} S_k^d S_i^e \right) \\
&= \frac{1}{2} \sum_{\substack{i,k=1 \\ i\neq k}}^{N} \omega_{ik} \left(f^{bcd} f^{ace} - f^{bce} f^{acd} \right) S_i^e S_k^d \\
&= \frac{1}{2} \sum_{\substack{i,k=1 \\ i\neq k}}^{N} \omega_{ik} f^{abc} f^{ced} S_i^e S_k^d \\
&= f^{abc} \Lambda^c. \qquad (4.13)
\end{aligned}$$

We note that these two properties only rely on the Lie algebra properties of $\mathfrak{su}(n)$ and not on the representation in which the spin operators act.

Now we proceed to calculate the commutation relation of the Hamiltonian and the rapidity operator. Using the definitions in Eqs. (4.1) and (4.3) we have

$$\begin{aligned}
[H_{\text{HS}}, \Lambda^b] &= \left(\frac{\pi}{N}\right)^2 \sum_{i\neq j, k\neq l} t_{ij} \omega_{kl} f^{bcd} [S_i^a S_j^a, S_k^c S_l^d] \\
&= \left(\frac{\pi}{N}\right)^2 \sum_{i\neq j, k\neq i} t_{ij} \omega_{kl} f^{bcd} \big(\delta_{jk} f^{ace} S_i^a S_j^e S_l^d + \delta_{jl} f^{ade} S_i^a S_k^c S_j^e \\
&\qquad\qquad + \delta_{ik} f^{ace} S_i^e S_l^d S_j^a + \delta_{jk} f^{ade} S_k^c S_i^e S_j^a \big) \\
&= \left(\frac{\pi}{N}\right)^2 \sum_{i\neq j, k\neq l} t_{ij} \omega_{kl} f^{bcd} \big(\delta_{jk} f^{ace} S_i^a S_j^e S_l^d + \delta_{jk} f^{ace} S_i^a S_l^d S_j^e \\
&\qquad\qquad + \delta_{jk} f^{ace} S_j^e S_l^d S_i^a + \delta_{jk} f^{ace} S_l^d S_j^e S_i^a \big) \\
&= \left(\frac{\pi}{N}\right)^2 \sum_{i\neq j\neq k\neq i} t_{ij} \omega_{jk} f^{bcd} f^{ace} S_j^e \{ S_i^a, S_k^d \} \\
&= \left(\frac{\pi}{N}\right)^2 \left(2 \sum_{i\neq j\neq k\neq i} t_{ij} \omega_{jk} f^{bcd} f^{ace} S_j^e S_i^a S_k^d + \sum_{i\neq j} t_{ij} \omega_{ji} f^{bcd} f^{ace} S_j^e \{ S_i^a, S_i^d \} \right). \qquad (4.14)
\end{aligned}$$

We investigate these two terms separately. For the first one we note that we can use the Jacobi identity and Eq. (4.9):

$$\begin{aligned}
2\sum_{i\neq j\neq k\neq i} t_{ij}\omega_{jk} f^{bcd} f^{ace} S_j^e S_i^a S_k^d &= 2\sum_{i\neq j\neq k\neq i} t_{ij}\omega_{jk}\left(f^{bac}f^{dec} - f^{bec}f^{dac}\right) S_j^e S_i^a S_k^d \\
&= 2\sum_{i\neq j\neq k\neq i} t_{ij}\left(\omega_{jk} - \omega_{ik}\right) f^{bac}f^{dec} S_j^e S_i^a S_k^d \\
&= 2\sum_{i\neq j\neq k\neq i} t_{ij}\left(\omega_{jk} - \omega_{ik}\right)\left(f^{aec}f^{bdc} - f^{adc}f^{bec}\right) S_j^e S_i^a S_k^d \\
&= -4\sum_{i\neq j\neq k\neq i} t_{ij}\left(\omega_{jk} - \omega_{ik}\right) f^{bac}f^{dec} S_j^e S_i^a S_k^d
\end{aligned}$$

and this expression is equal to a multiple of itself and therefore vanishes. For the second term in Eq. (4.14) we note

$$\sum_{i\neq j} t_{ij}\omega_{ji} f^{bcd} f^{ace} S_j^e \{S_i^a, S_i^d\} = \sum_{i\neq j} t_{ij}\omega_{ji} f^{bcd} f^{ace} S_j^e \left(\frac{1}{n}\delta^{ad} + d^{adf} S_i^f\right).$$

The first part of this expression vanishes as a result of Eq. (4.10), while for the second part it follows from Eq. (4.8) that

$$f^{bcd} f^{ace} d^{adf} = f^{ace}(f^{acd} d^{bdf} - f^{fcd} d^{bad}).$$

We claim that the right hand side is symmetric in e and f. For the first term this follows from $f^{abc} f^{abd} = C\delta^{ad}$. For the second term we note

$$f^{ace} f^{fcd} d^{bad} = f^{dce} f^{fca} d^{bda} = f^{acf} f^{ecd} d^{bad}.$$

Using the fact that $t_{ij}\omega_{ji} = -t_{ji}\omega_{ij}$, we see that

$$\sum_{i\neq j} t_{ij}\omega_{ji} f^{bcd} f^{ace} S_j^e \{S_i^a, S_i^d\} = 0.$$

We conclude that the commutator between the Hamiltonian and the rapidity operator vanishes. We note that this result, unlike Eq. (4.12) and Eq. (4.13), does depend on the fact that the $\mathfrak{su}(n)$ operators act in the fundamental representation. The relation Eq. (4.5) does not generalize to arbitrary representations. Using numerical calculations we have established that for other representations the commutator does not vanish.

4.4 Serre relation of the Yangian for the Haldane-Shastry model

In the preceding section we have shown that the total spin operator $\mathbf{S}$ and the rapidity operator $\mathbf{\Lambda}$ both commute with the Hamiltonian of the HS model. Therefore, these two operators generate a symmetry algebra of the model. We will explicitly show for the case of $\mathfrak{su}(2)$ that this algebra is a finite-dimensional representation of the Yangian. The necessary and sufficient condition is the existence of operators that satisfy the defining relations (Eqs. (3.1) –(3.3)) of the Yangian. We will show that this is the case for the operators $\mathbf{S}$ and $\mathbf{\Lambda}$, which satisfy Eq. (3.1) as we have seen in Eq. (4.13). In the case of $\mathfrak{su}(2)$, Eq. (3.2) is automatically satisfied. Therefore, we only have to check Eq. (3.3), which is also called the Serre relation of the Yangian. For $\mathfrak{su}(2)$ the structure constants in the basis of spin operators (which equal half the Pauli matrices) are equal to $f^{abc} = \mathrm{i}\varepsilon^{abc}$ where ε^{abc} is the three-dimensional Levi-Civita symbol. This will simplify the calculations enormously, as a result of the convenient formulas for the contraction of multiple Levi-Civita symbols. We start by computing the commutator of two rapidity operators, using that $\omega_{ij} = -\omega_{ji}$,

$$\begin{aligned}
[\Lambda^a, \Lambda^b] &= -\frac{\mathrm{i}}{4} \sum_{\substack{i\neq j\\ k\neq l}} \omega_{ij}\omega_{kl}\varepsilon^{acd}\varepsilon^{bef}\Big(\delta_{jk}\varepsilon^{deg} S_i^c S_j^g S_l^f + \delta_{jl}\varepsilon^{dfg} S_i^c S_k^e S_j^g \\
&\qquad\qquad + \delta_{ik}\varepsilon^{ceg} S_i^g S_l^f S_j^d + \delta_{il}\varepsilon^{cfg} S_k^e S_i^g S_j^d\Big) \\
&= -\frac{\mathrm{i}}{4} \sum_{\substack{i\neq j\\ k\neq l}} \omega_{ij}\omega_{kl}\varepsilon^{acd}\varepsilon^{bef}\varepsilon^{deg}\delta_{jk}\left(S_i^c S_j^g S_l^f + S_i^c S_l^f S_j^g + S_j^g S_l^f S_i^c + S_l^f S_j^g S_i^c\right) \\
&= -\frac{\mathrm{i}}{2} \sum_{i\neq j\neq k} \omega_{ij}\omega_{jk}\varepsilon^{acd}\varepsilon^{bef}\varepsilon^{deg} S_j^g \{S_i^c, S_k^f\}.
\end{aligned} \tag{4.15}$$

We split this into two terms: one term with three different sites involved (T^{ab}) and one with only two sites involved (R^{ab}). So T^{ab} consists of all the terms with $i \neq k$, while R^{ab} contains the terms with $i = k$. First we investigate R^{ab}. We will show that this term will not contribute to the Serre relation if and only if all the operators S_i^a act in the fundamental representation. We have

$$R^{ab} = -\frac{\mathrm{i}}{2}\sum_{i\neq j} \omega_{ij}\omega_{ji}\varepsilon^{acd}\varepsilon^{bef}\varepsilon^{deg} S_j^g \{S_i^c, S_i^f\}.$$

We insert the expression for the anticommutator and simplify the product of the Levi-Civita symbols

$$\begin{aligned}
R^{ab} &= -\frac{\mathrm{i}}{4}\sum_{i\neq j} \omega_{ij}\omega_{ji}\varepsilon^{acd}\varepsilon^{bef}\varepsilon^{deg}\delta^{cf} S_j^g \\
&= -\frac{\mathrm{i}}{4}\sum_{i\neq j} \omega_{ij}\omega_{ji}\varepsilon^{acd}\delta^{bd}\delta^{cg} S_j^g \\
&= \frac{\mathrm{i}}{4}\sum_{i\neq j} \omega_{ij}\omega_{ji}\varepsilon^{abc} S_j^c.
\end{aligned}$$

To proceed, we note that only the factor $\omega_{ij}\omega_{ji}$ depends on the label i. There is a simple expression for the remaining summation over i,

$$\sum_{\substack{i=1\\ i\neq j}}^{N} \omega_{ij}\omega_{ji} = \frac{(N-1)(N-2)}{3}, \tag{4.16}$$

which is proven in the appendix to this chapter. Conveniently, the right hand side does not depend on the index j. As a result, we see that R^{ab} is proportional to $\varepsilon^{abc}S^c$. Since we are ultimately interested in the quantity $[R^{ab}, \Lambda^c]$ because it enters the Serre relation in that way, we see that we can use the fact that the rapidity operator transforms as a vector under $\mathfrak{su}(2)$

transformation to circumvent the use of any indices for the sites. To be precise, R^{ab} appears in the Serre relation as

$$[R^{ab},[S^p,\Lambda^q]]+[R^{pq},[S^a,\Lambda^b]] = -\frac{(N-1)(N-2)}{12}\left(\varepsilon^{abc}\varepsilon^{pqr}[S^c,\Lambda^r]+\varepsilon^{abc}\varepsilon^{pqr}[S^r,\Lambda^c]\right). \quad (4.17)$$

This expression vanishes as a result of $[S^c,\Lambda^k]+[S^k,\Lambda^c]=0$. We conclude that the term R^{ab} in $[\Lambda^a,\Lambda^b]$ does not contribute to the Serre relation and can be ignored in the rest of the calculation. Here we would like to stress that this is only true when the spin operators in question have $s=\frac{1}{2}$. For other $\mathfrak{su}(2)$ representations, there is no convenient expression for the anticommutator. As a result, in that case Eq. (4.16) cannot be used and there is no simple form for R^{ab}. Numerical calculations imply that for higher spin representations, the above conclusion is not always true and R^{ab} does contribute to the Serre relation.

The remaining term that we have to calculate is $\left[T^{ab},[S^r,\Lambda^s]\right]+\left[T^{rs},[S^a,\Lambda^b]\right]$, where we recall that T^{ab} are the terms in Eq. (4.15) that act nontrivially on three different sites. We will not do this calculation in full detail but reduce the problem to two equations and check one of them. We start with the following observation.

Lemma 4.1. *Both the left hand side and right hand side of the Serre relation have the following symmetries:*

$$\begin{aligned}
&T^{abrs} = -T^{bars} = -T^{abrs},\\
&T^{abrs} = T^{rsab},\\
&T^{[abrs]} = 0.
\end{aligned}$$

where T^{abrs} is a shorthand notation for the complete left or right hand side and $T^{[abrs]}$ is the completely antisymmetrized version of T^{abij}. As a consequence, both sides have six independent components.

Proof. From $[I^a,J^b]=-[I^b,J^a]$ we see that the left hand side is antisymmetric in the first and last two indices. For the right-hand side this follows from

$$\begin{aligned}
24\mathfrak{c}^{abcdef}\{I^d,I^e,J^f\} &= c^{adp}c^{beq}c^{cfr}c^{pqr}\{I^d,I^e,J^f\}\\
&= c^{aeq}c^{bdp}c^{cfr}c^{qpr}\{I^e,I^d,J^f\}\\
&= -c^{aeq}c^{bdp}c^{cfr}c^{pqr}\{I^d,I^e,J^f\}\\
&= -24\mathfrak{c}^{bacdef}\{I^d,I^e,J^f\}.
\end{aligned}$$

The fact that both sides are symmetric under the exchange of the first and last pair of indices is obvious. Finally, we note that $T^{[abij]}$ is completely antisymmetric with four different indices, which can only take three values. Therefore, in each case at least two index values are equal and the expression vanishes due to antisymmetry. The statement about the number of independent components follows from the observation that the symmetries of both sides of the Serre relation are the same as that of the Riemann tensor (from the theory of Riemannian manifolds) which has six independent components in three dimensions [9]. □

The six independent components in question are T^{1212}, T^{2323}, T^{3131}, T^{1223}, T^{2331} and T^{3112}. We abbreviate T_l^{abij} and T_r^{abij} for the left and right hand side of the Serre relation respectively. To proceed, we plug in the values for the indices on both sides. We note that the structure constants acquire an extra factor of $\sqrt{2}$ in the orthonormal basis, i.e. $c^{abc} = \mathrm{i}\sqrt{2}\varepsilon^{abc}$. As a result, one finds

$$\begin{aligned} T_l^{1212} &= 2\Big[[J^1, J^2], [I^1, J^2]\Big] = 2\sqrt{2}\mathrm{i}\Big[[J^1, J^2], J^3\Big], \\ T_l^{1223} &= \Big[[J^1, J^2], [I^2, J^3]\Big] + \Big[[J^2, J^3], [I^1, J^2]\Big] = \sqrt{2}\mathrm{i}\Big[[J^1, J^2], J^1\Big] + \sqrt{2}\mathrm{i}\Big[[J^2, J^3], J^3\Big] \end{aligned}$$

and

$$\begin{aligned} T_r^{1212} &= \frac{1}{12}\left(c^{1di}c^{2ej}c^{cfk}c^{ijk}c^{12c}\right)\{I^d, I^e, J^f\} \\ &= \frac{\mathrm{i}\sqrt{2}}{3}\left(\varepsilon^{1di}\varepsilon^{2ej}\left(\delta^{3i}\delta^{fj} - \delta^{3j}\delta^{fi}\right)\varepsilon^{123}\right)\{I^d, I^e, J^f\} \\ &= \frac{\mathrm{i}\sqrt{2}}{3}\left(\varepsilon^{1d3}\varepsilon^{2ef} - \varepsilon^{1df}\varepsilon^{2e3}\right)\{I^d, I^e, J^f\} \\ &= \frac{\mathrm{i}\sqrt{2}}{3}\left(\{I^2, I^3, J^1\} - \{I^2, I^1, J^3\} + \{I^2, I^1, J^3\} - \{I^3, I^1, J^2\}\right) \\ &= \frac{\mathrm{i}\sqrt{2}}{3}\left(\{I^2, I^3, J^1\} - \{I^3, I^1, J^2\}\right). \end{aligned}$$

Similarly, one finds

$$T_r^{1223} = \frac{\mathrm{i}\sqrt{2}}{6}\left(\{I^3, I^3, J^2\} + \{I^2, I^3, J^3\} - \{I^1, I^1, J^2\} - \{I^2, I^1, J^1\}\right).$$

In the case of the Haldane-Shastry model, it suffices to check only the equations $T_l^{1212} = T_r^{1212}$ and $T_l^{1223} = T_r^{1223}$. The reason for this is that in the construction of the operators S^a and Λ^a all three spatial directions (x, y, z) are treated on equal footing. Furthermore, the structure constants are invariant under cyclic permutation. Therefore, cyclicly permuting the above two equations will lead to the other four.

We will only check the relation $T_l^{1212} = T_r^{1212}$ and leave the other equation for the reader. Under the identification $I^a = \sqrt{2}S^a$ and $J^a = \sqrt{2}\Lambda^a$, this equation becomes

$$\Big[[\Lambda^1, \Lambda^2], \Lambda^3\Big] = \frac{1}{6}\left(\{S^2, S^3, \Lambda^1\} - \{S^3, S^1, \Lambda^2\}\right). \tag{4.18}$$

We showed in Eq. (4.17) that the left hand side is equal to $[T^{12}, \Lambda^3]$ where T^{12} only contains the terms of $[J^1, J^2]$ with the three spin operators on three different sites. Using the definition of T^{ab} and the contraction of some Levi-Civita symbols, we find

$$T^{12} = \mathrm{i}\sum_{i\neq j\neq k\neq i}\omega_{ij}\omega_{jk}(S_i^1 S_j^1 S_k^3 + S_i^2 S_j^2 S_k^3 + S_i^3 S_j^3 S_k^3).$$

Therefore, we have

$$\begin{aligned} [T^{12}, \Lambda^3] &= -\sum_{\substack{i\neq j\neq k\neq i\\ l\neq m}}\omega_{ij}\omega_{jk}\omega_{lm}\left[S_i^1 S_j^1 S_k^3 + S_i^2 S_j^2 S_k^3 + S_i^3 S_j^3 S_k^3, S_l^1 S_m^2\right] \\ &= -\mathrm{i}\sum_{\substack{i\neq j\neq k\neq i\\ l\neq m}}\omega_{ij}\omega_{jk}\omega_{lm}\Big(\delta_{kl}S_i^1 S_j^1 S_k^2 S_m^2 - \delta_{km}S_i^1 S_j^1 S_k^1 S_l^1 + \delta_{jm}S_i^1 S_l^1 S_j^3 S_k^3 + \delta_{im}S_l^1 S_i^3 S_j^1 S_k^3 \\ &\quad + \delta_{kl}S_i^2 S_j^2 S_k^2 S_m^2 - \delta_{km}S_i^2 S_j^2 S_l^1 S_k^1 - \delta_{jl}S_i^2 S_j^3 S_m^2 S_k^3 - \delta_{il}S_i^3 S_m^2 S_j^2 S_k^3 + \delta_{kl}S_i^3 S_j^3 S_k^2 S_m^2 \end{aligned}$$

$$
\begin{aligned}
&- \delta_{km} S_i^3 S_j^3 S_l^1 S_k^1 - \delta_{jm} S_i^3 S_l^1 S_j^1 S_k^3 + \delta_{jl} S_i^3 S_j^2 S_m^2 S_k^3 + \delta_{il} S_i^2 S_m^2 S_j^3 S_k^3 - \delta_{im} S_l^1 S_a^1 S_j^3 S_k^3 \Big) \\
= -\mathrm{i} \sum_{\substack{i \neq j \neq k \neq i \\ l \neq m}} & \omega_{ij}\omega_{jk}\omega_{lm} \Big(\delta_{il} \left(-S_m^1 S_i^3 S_j^1 S_k^3 - S_i^3 S_m^2 S_j^2 S_k^3 + S_i^2 S_m^2 S_j^3 S_k^3 + S_m^1 S_a^1 S_j^3 S_k^3 \right) \\
&+ \delta_{jl} \left(-S_i^1 S_m^1 S_j^3 S_k^3 - S_i^2 S_j^3 S_m^2 S_k^3 + S_i^3 S_m^1 S_j^1 S_k^3 + S_i^3 S_j^2 S_m^2 S_k^3 \right) + \delta_{kl} (S_i^1 S_j^1 S_k^2 S_m^2 \\
&+ S_i^1 S_j^1 S_k^1 S_m^1 + S_i^2 S_j^2 S_k^2 S_m^2 + S_i^2 S_j^2 S_m^1 S_k^1 + S_i^3 S_j^3 S_k^2 S_m^2 + S_i^3 S_j^3 S_m^1 S_k^1) \Big).
\end{aligned}
$$

We note that there are two kinds of terms. Namely, the spin operators act either on three or four different sites. We start with the investigation of the terms were the spin operators act on four different sites, i.e. the terms were all indices are different. We read off that these terms are equal to

$$
\begin{aligned}
&- \mathrm{i} \sum_{i,j,k,m}{}' \Big\langle \omega_{ij}\omega_{jk}\omega_{im} \left(-S_m^1 S_i^3 S_j^1 S_k^3 - S_i^3 S_m^2 S_j^2 S_k^3 + S_i^2 S_m^2 S_j^3 S_k^3 + S_m^1 S_i^1 S_j^3 S_k^3 \right) + \\
&\omega_{ij}\omega_{jk}\omega_{jm} \left(-S_i^1 S_m^1 S_j^3 S_k^3 - S_i^2 S_j^3 S_m^2 S_k^3 + S_i^3 S_m^1 S_j^1 S_k^3 + S_i^3 S_j^2 S_m^2 S_k^3 \right) + \\
&\omega_{ij}\omega_{jk}\omega_{km} \big(S_i^1 S_j^1 S_k^2 S_m^2 + S_i^1 S_j^1 S_k^1 S_m^1 + S_i^2 S_j^2 S_k^2 S_m^2 + S_i^2 S_j^2 S_m^1 S_k^1 \\
&\qquad\qquad\qquad\qquad\qquad + S_i^3 S_j^3 S_k^2 S_m^2 + S_i^3 S_j^3 S_m^1 S_k^1 \big) \Big\rangle.
\end{aligned}
$$

The prime on the sum denotes that values of equal indices are omitted. Before we proceed, we note that some of these terms vanish.

$$
\begin{aligned}
\sum_{i,j,k,m}{}' \omega_{ij}\omega_{jk}\omega_{km} (S_i^2 S_j^2 S_k^1 S_m^1 + S_i^1 S_j^1 S_k^2 S_m^2) &= \sum_{i,j,k,m}{}' (\omega_{ij}\omega_{jk}\omega_{km} + \omega_{mk}\omega_{kj}\omega_{ji}) S_i^2 S_j^2 S_k^1 S_m^1 \\
&= \sum_{i,j,k,m}{}' (\omega_{ij}\omega_{jk}\omega_{km} - \omega_{ij}\omega_{jk}\omega_{km}) S_i^2 S_j^2 S_k^1 S_m^1 \\
&= 0.
\end{aligned}
$$

Similarly, the two terms for which all spin operators act in the same direction are 0. This leaves ten terms in total, which all have precisely two spin operators in the z-direction. We now collect the five terms that also have two spin operators acting in the x-direction and we simplify the result. Here the relation $\omega_{ij}\omega_{ik} + \omega_{ji}\omega_{jk} + \omega_{ki}\omega_{kj} = 1$ will often come in handy. Using this we can simplify the result to find

$$
\begin{aligned}
&- \mathrm{i} \sum_{i,j,k,m}{}' \Big\langle \omega_{ij}\omega_{jk}\omega_{im} \left(-S_m^1 S_i^3 S_j^1 S_k^3 + S_m^1 S_i^1 S_j^3 S_k^3 \right) + \\
&\qquad\qquad \omega_{ij}\omega_{jk}\omega_{jm} \left(-S_i^1 S_m^1 S_j^3 S_k^3 + S_i^3 S_m^1 S_j^1 S_k^3 \right) + \omega_{ij}\omega_{jk}\omega_{km} S_i^3 S_j^3 S_m^1 S_k^1 \Big\rangle \\
&= -\mathrm{i} \sum_{i,j,k,m}{}' \big(-\omega_{mj}\omega_{jk}\omega_{mi} + \omega_{im}\omega_{mk}\omega_{ij} - \omega_{im}\omega_{mk}\omega_{mj} + \omega_{mj}\omega_{jk}\omega_{ji} \\
&\qquad\qquad\qquad\qquad + \omega_{mk}\omega_{kj}\omega_{ji} \big) S_i^1 S_j^1 S_k^3 S_m^3 \\
&= -\mathrm{i} \sum_{i,j,k,m}{}' \left(\omega_{im}\omega_{mj}\omega_{jk} + \omega_{im}\omega_{mk}\omega_{ij} - \omega_{im}\omega_{mk}\omega_{mj} + \omega_{mj}\omega_{jk}\omega_{ji} + \omega_{mk}\omega_{kj}\omega_{ji} \right) S_i^1 S_j^1 S_k^3 S_m^3 \\
&= -\mathrm{i} \sum_{i,j,k,m}{}' \left(\omega_{ij} \left(\omega_{jm}\omega_{jk} + \omega_{km}\omega_{kj} \right) + \omega_{im}\omega_{mj}\omega_{jk} + \omega_{im}\omega_{mk}\omega_{ij} + \omega_{mi}\omega_{mk}\omega_{mj} \right) S_i^1 S_j^1 S_k^3 S_m^3
\end{aligned}
$$

$$
\begin{aligned}
&= -\mathrm{i} \sum_{i,j,k,m}{}' \left(\omega_{ij} - \omega_{ij}\omega_{mj}\omega_{mk} + \omega_{im}\omega_{mj}\omega_{jk} + \omega_{im}\omega_{mk}\omega_{ij} + \omega_{mi}\omega_{mk}\omega_{mj}\right) S_i^1 S_j^1 S_k^3 S_m^3 \\
&= -\mathrm{i} \sum_{i,j,k,m}{}' \big(- \omega_{ij}\omega_{mj}\omega_{mk} + \omega_{im}\omega_{mj}\omega_{jk} + \omega_{im}\omega_{mk}\omega_{ij} \\
&\qquad\qquad\qquad\qquad + \omega_{mj}\left(1 - \omega_{im}\omega_{ik} - \omega_{km}\omega_{ki}\right)\big) S_i^1 S_j^1 S_k^3 S_m^3 \\
&= -\mathrm{i} \sum_{i,j,k,m}{}' \left(\omega_{mj} + \left(\omega_{im}\omega_{mk}\omega_{ij} - \omega_{ij}\omega_{mj}\omega_{mk}\right) + \left(\omega_{im}\omega_{mj}\omega_{jk} - \omega_{mj}\omega_{im}\omega_{ik}\right)\right) S_i^1 S_j^1 S_k^3 S_m^3 \\
&= -\mathrm{i} \sum_{i,j,k,m}{}' \omega_{mj} S_i^1 S_j^1 S_k^3 S_m^3 .
\end{aligned}
$$

The terms between the inner brackets cancel each other after a permutation of indices. We conclude that the term with two spin operators in the x-direction and two in the z-direction is equal to

$$
i \sum_{i,j,k,m}{}' \omega_{ik} S_i^1 S_j^1 S_k^3 S_m^3 . \tag{4.19}
$$

Following similar steps one can show that the term containing two spin operators in the y-direction and two in the z-direction, which all act on different sites is equal to

$$
\mathrm{i} \sum_{i,j,k,m}{}' \omega_{ik} S_i^2 S_j^2 S_k^3 S_m^3 . \tag{4.20}
$$

Having dealt with the terms acting all on different sites, we can look for the terms with multiple spin operators on the same site. These are the following terms

$$
\begin{aligned}
&- \mathrm{i} \sum_{i \neq j \neq k \neq i} \Big\langle \omega_{ij}\omega_{jk}\omega_{ij} \left(-S_j^1 S_i^3 S_j^1 S_k^3 - S_i^3 S_j^2 S_j^2 S_k^3 + S_i^2 S_j^2 S_j^3 S_k^3 + S_j^1 S_i^1 S_j^3 S_k^3\right) \\
&+ \omega_{ij}\omega_{jk}\omega_{ik} \left(-S_k^1 S_i^3 S_j^1 S_k^3 - S_i^3 S_k^2 S_j^2 S_k^3 + S_i^2 S_k^2 S_j^3 S_k^3 + S_k^1 S_i^1 S_j^3 S_k^3\right) \\
&+ \omega_{ij}\omega_{jk}\omega_{ji} \left(-S_i^1 S_i^1 S_j^3 S_k^3 - S_i^2 S_j^3 S_i^2 S_k^3 + S_i^3 S_i^1 S_j^1 S_k^3 + S_i^3 S_j^2 S_i^2 S_k^3\right) \\
&+ \omega_{ij}\omega_{jk}\omega_{jk} \left(-S_i^1 S_k^1 S_j^3 S_k^3 - S_i^2 S_j^3 S_k^2 S_k^3 + S_i^3 S_k^1 S_j^1 S_k^3 + S_i^3 S_j^2 S_k^2 S_k^3\right) \\
&+ \omega_{ij}\omega_{jk}\omega_{ki} \left(S_i^1 S_j^1 S_k^2 S_i^2 + S_i^1 S_j^1 S_k^1 S_i^1 + S_i^2 S_j^2 S_k^2 S_i^2 + S_i^2 S_j^2 S_i^1 S_k^1 + S_i^3 S_j^3 S_k^2 S_i^2 + S_i^3 S_j^3 S_i^1 S_k^1\right) \\
&+ \omega_{ij}\omega_{jk}\omega_{kj} \left(S_i^1 S_j^1 S_k^2 S_j^2 + S_i^1 S_j^1 S_k^1 S_j^1 + S_i^2 S_j^2 S_k^2 S_j^2 + S_i^2 S_j^2 S_j^1 S_k^1 + S_i^3 S_j^3 S_k^2 S_j^2 + S_i^3 S_j^3 S_j^1 S_k^1\right) \Big\rangle .
\end{aligned}
$$

The Clifford algebra structure that the Pauli matrices satisfy allows for the following simplification

$$
\begin{aligned}
&- \mathrm{i} \sum_{i \neq j \neq k \neq i} \Big\langle \omega_{ij}\omega_{jk}\omega_{ij} \left(-\frac{1}{4} S_i^3 S_k^3 - \frac{1}{4} S_i^3 S_k^3 + \frac{\mathrm{i}}{2} S_i^2 S_j^1 S_k^3 - \frac{\mathrm{i}}{2} S_j^2 S_i^1 S_k^3\right) + \\
&\omega_{ij}\omega_{jk}\omega_{ik} \left(\frac{\mathrm{i}}{2} S_k^2 S_i^3 S_j^1 - \frac{\mathrm{i}}{2} S_i^3 S_k^1 S_j^2 + \frac{\mathrm{i}}{2} S_i^2 S_k^1 S_j^3 - \frac{\mathrm{i}}{2} S_k^2 S_i^1 S_j^3\right) + \\
&\omega_{ij}\omega_{jk}\omega_{ji} \left(-\frac{1}{4} S_j^3 S_k^3 - \frac{1}{4} S_j^3 S_k^3 + \frac{\mathrm{i}}{2} S_i^2 S_j^1 S_k^3 - \frac{\mathrm{i}}{2} S_i^1 S_j^2 S_k^3\right) + \\
&\omega_{ij}\omega_{jk}\omega_{jk} \left(\frac{i}{2} S_i^1 S_k^2 S_j^3 - \frac{\mathrm{i}}{2} S_i^2 S_j^3 S_k^1 - \frac{\mathrm{i}}{2} S_i^3 S_k^2 S_j^1 + \frac{i}{2} S_i^3 S_j^2 S_k^1\right) +
\end{aligned}
$$

$$\omega_{ij}\omega_{jk}\omega_{ki}\left(\frac{i}{2}S_i^3S_j^1S_k^2+\frac{1}{4}S_j^1S_k^1+\frac{1}{4}S_j^2S_k^2-\frac{\mathrm{i}}{2}S_i^3S_j^2S_k^1-\frac{\mathrm{i}}{2}S_i^1S_j^3S_k^2+\frac{i}{2}S_j^3S_i^2S_k^1\right)+$$
$$\omega_{ij}\omega_{jk}\omega_{kj}\left(\frac{i}{2}S_i^1S_j^3S_k^2+\frac{1}{4}S_i^1S_k^1+\frac{1}{4}S_i^2S_k^2-\frac{\mathrm{i}}{2}S_i^2S_j^3S_k^1-\frac{\mathrm{i}}{2}S_i^3S_j^1S_k^2+\frac{\mathrm{i}}{2}S_i^3S_j^2S_k^1\right)\Big\rangle.$$

Close inspection shows that there are only two kinds of terms. On one hand, there are terms with two spin operators pointing in the same direction and on the other hand, we have terms with three spin operators pointing in different directions. We start with investigating the terms with two spin operators in the z-direction. One finds

$$\begin{aligned}
&\frac{\mathrm{i}}{2}\sum_{i\neq j\neq k\neq i}\left(\omega_{ij}\omega_{jk}\omega_{ij}S_i^3S_k^3+\omega_{ij}\omega_{jk}\omega_{ji}S_j^3S_k^3\right)\\
=&\frac{\mathrm{i}}{2}\sum_{i\neq j\neq k\neq i}\left(\omega_{ij}\omega_{jk}\omega_{ij}S_i^3S_k^3+\omega_{ji}\omega_{ik}\omega_{ij}S_i^3S_k^3\right)\\
=&\frac{\mathrm{i}}{2}\sum_{i\neq j\neq k\neq i}\omega_{ij}\omega_{jk}\omega_{ij}S_i^3S_k^3+\frac{(N-1)(N-2)}{6}\sum_{i\neq k}\omega_{ik}S_i^3S_k^3+\frac{\mathrm{i}}{2}\sum_{i\neq k}\omega_{ik}^3S_i^3S_k^3\\
=&\frac{\mathrm{i}}{2}\sum_{i\neq j\neq k\neq i}\omega_{ji}\left(1-\omega_{kj}\omega_{ki}-\omega_{ij}\omega_{ik}\right)S_i^3S_k^3.
\end{aligned}$$

These three terms all vanish, due to Eq. (4.11) and Eq. (4.16). One has to be very careful when using these relations, since two values for the index j are omitted instead of one. However, the extra terms all vanish due to the symmetry of $S_i^3S_k^3$ in i and k. Similarly, one can show that the terms proportional to $S_i^2S_k^2$ and $S_i^1S_k^1$ vanish. This only leaves the terms with three spin operators in three different directions, which are (after relabeling indices)

$$\begin{aligned}
&\frac{1}{2}\sum_{i\neq j\neq k\neq i}(\omega_{ji}\omega_{ik}\omega_{ji}-\omega_{ij}\omega_{jk}\omega_{ij}+\omega_{ki}\omega_{ij}\omega_{kj}-\omega_{kj}\omega_{ji}\omega_{ki}+\omega_{jk}\omega_{ki}\omega_{ji}-\omega_{ik}\omega_{kj}\omega_{ij}\\
&+\omega_{ji}\omega_{ik}\omega_{ij}-\omega_{ij}\omega_{jk}\omega_{ji}+\omega_{ik}\omega_{kj}\omega_{kj}-\omega_{jk}\omega_{ki}\omega_{ki}-\omega_{ki}\omega_{ij}\omega_{ij}+\omega_{kj}\omega_{ji}\omega_{ji}\\
&+\omega_{ki}\omega_{ij}\omega_{jk}-\omega_{kj}\omega_{ji}\omega_{ik}-\omega_{ik}\omega_{kj}\omega_{ji}+\omega_{jk}\omega_{ki}\omega_{ij}+\omega_{ik}\omega_{kj}\omega_{jk}\\
&-\omega_{jk}\omega_{ki}\omega_{ik}-\omega_{ki}\omega_{ij}\omega_{ji}+\omega_{kj}\omega_{ji}\omega_{ij})S_i^1S_j^2S_k^3.
\end{aligned}$$

Close inspection reveals that all of these terms cancel each other. Therefore, the left hand side of the Serre relation only contains terms with four spin operators on four different sites, which is a remarkable fact. To be precise, using Eq. (4.19) and (4.20), we have

$$T_l^{1212}=\mathrm{i}\sum_{i,j,k,m}{}'\,\omega_{ik}S_i^1S_j^1S_k^3S_m^3+\mathrm{i}\sum_{i,j,k,m}{}'\,\omega_{ik}S_i^2S_j^2S_k^3S_m^3. \tag{4.21}$$

We look at the right hand side of the Serre equation Eq. (4.18), which is equal to $\frac{1}{6}(\{S^2,S^3,\Lambda^1\}-\{S^3,S^1,\Lambda^2\})$ in the particular case we are interested in. First, we investigate one of these triple products and expand all terms

$$\begin{aligned}
\{S^3,S^1,\Lambda^2\}=&\ \mathrm{i}\sum_{\substack{i,j\\k\neq m}}\omega_{km}\{S_i^1,S_j^3,S_k^3S_m^1\}\\
&=\mathrm{i}\sum_{i,j,k,m}{}'\,\omega_{km}\{S_i^1,S_j^3,S_k^3S_m^1\}+\mathrm{i}\sum_{i,k,m}{}'\,\omega_{km}\{S_i^1,S_i^3,S_k^3S_m^1\}+\mathrm{i}\sum_{i,m}{}'\omega_{im}\{S_i^1,S_i^3,S_i^3S_m^1\}
\end{aligned}$$

$$
\begin{aligned}
&+\mathrm{i}\sum_{i,m}{}' \omega_{mi}\{S_i^1, S_i^3, S_m^3 S_i^1\} + \mathrm{i}\sum_{i,j,m}{}' \omega_{im}\{S_i^1, S_j^3, S_i^3 S_m^1\} + \mathrm{i}\sum_{i,j,m}{}' \omega_{mi}\{S_i^1, S_j^3, S_m^3 S_i^1\} \\
&+\mathrm{i}\sum_{i,j,m}{}' \omega_{jm}\{S_i^1, S_j^3, S_j^3 S_m^1\} + \mathrm{i}\sum_{i,j,m}{}' \omega_{mj}\{S_i^1, S_j^3, S_m^3 S_j^1\} \\
=\ & 6\mathrm{i}\sum_{i,j,k,m}{}' \omega_{km} S_i^1 S_j^3 S_k^3 S_m^1 + 3\mathrm{i}\sum_{i,k,m}{}' \omega_{km}\{S_i^1, S_i^3\} S_k^3 S_m^1 + \mathrm{i}\sum_{i,m}{}' \omega_{im}\{S_i^1, S_i^3, S_i^3\} S_m^1 \\
&+\mathrm{i}\sum_{i,m}{}' \omega_{mi}\{S_i^1, S_i^3, S_i^1\} S_m^3 + 3\mathrm{i}\sum_{i,j,m}{}' \omega_{im}\{S_i^1, S_i^3\} S_j^3 S_m^1 + 3\mathrm{i}\sum_{i,j,m}{}' \omega_{mi}\{S_i^1, S_i^1\} S_j^3 S_m^3 \\
&+3\mathrm{i}\sum_{i,j,m}{}' \omega_{jm}\{S_j^3, S_j^3\} S_i^1 S_m^1 + 3\mathrm{i}\sum_{i,j,m}{}' \omega_{mj}\{S_j^3, S_j^1\} S_i^1 S_m^3.
\end{aligned}
$$

We can use the fact that $\{S_i^a, S_i^b\} = \frac{1}{2}\delta^{ab}$, which allows us to drop some terms and simplify some others. We note that this is only true for spin-$\frac{1}{2}$ operators and does not generalize to arbitrary spin. Similarly, we can rewrite

$$
\{S_i^1, S_i^3, S_i^1\} = 2S_i^1 S_i^3 S_i^1 + 2S_i^1 S_i^1 S_i^3 + 2S_i^3 S_i^1 S_i^1 = \frac{1}{2} S_i^3,
$$

which leads to

$$
\{S^3, S^1, \Lambda^2\} = 6\mathrm{i}\sum_{i,j,k,m}{}' \omega_{km} S_i^1 S_j^3 S_k^3 S_m^1 + \frac{3\mathrm{i}}{2}\sum_{i,j,m}{}' \omega_{im} S_j^1 S_m^1 + \frac{3\mathrm{i}}{2}\sum_{i,j,m}{}' \omega_{mj} S_i^3 S_m^3. \tag{4.22}
$$

This can be simplified further by noting the following relation, which follows from Eq. (4.11)

$$
\sum_{\substack{i=1 \\ j\neq i\neq m}}^{N} \omega_{im} = \sum_{\substack{i=1 \\ i\neq m}}^{N} \omega_{im} - \omega_{jm} = -\omega_{jm}
$$

and therefore the last two terms in Eq. (4.22) vanish. We find that

$$
\{S^3, S^1, \Lambda^2\} = 6\mathrm{i}\sum_{i,j,k,m}{}' \omega_{km} S_i^1 S_j^3 S_k^3 S_m^1 = -6\mathrm{i}\sum_{i,j,k,m}{}' \omega_{ik} S_i^1 S_j^1 S_k^3 S_m^3, \tag{4.23}
$$

which is very convenient expression and of course only holds for spin-$\frac{1}{2}$ operators. Similarly, one can show that

$$
\{S^2, S^3, \Lambda^1\} = 6\mathrm{i}\sum_{i,j,k,m}{}' \omega_{km} S_i^2 S_j^3 S_k^2 S_m^3 = 6\mathrm{i}\sum_{i,j,k,m}{}' \omega_{ik} S_i^2 S_j^2 S_k^3 S_m^3. \tag{4.24}
$$

Combining the results obtained in Eqs. (4.18), (4.21), (4.23) and (4.24) with each other, we see that the Serre relation is satisfied.

4.5 Existence of a spin-1 chain with Yangian symmetry

We have seen that S^a and Λ^a satisfy the Serre relation in the case of spin-$\frac{1}{2}$. We are interested in a spin-1 generalization of the Haldane-Shastry model, i.e. a long range spin chain with Yangian symmetry. Here, the Yangian symmetry should be generated by the total spin and some nontrivial operator to get an interesting result. Nontrivial means that the second operator, the analogue of the rapidity operator in the spin-$\frac{1}{2}$ case, should not be a multiple of the total spin. In that case, the defining relations of the Yangian are satisfied, but the symmetry algebra of the system is not necessarily large. We will give an argument why replacing the spin-$\frac{1}{2}$ operators by their spin-1 counterparts does not work, based on numerical calculations. In that case, the rapidity operator does not commute with the Hamiltonian and the Serre relation Eq. (3.3) is not satisfied. Of course, it could well be there is another spin-1 model for which there exists a Yangian symmetry. We can make a similar observation in the case of the Heisenberg spin chain. In the case of spin-$\frac{1}{2}$, the Hamiltonian reads [19]

$$H_{\mathrm{H}} = -J \sum_{i=1}^{N} \mathbf{S}_i \cdot \mathbf{S}_{i+1}$$

where we assume periodic boundary conditions. Here, the sign of J determines if the system describes ferromagnetic or antiferromagnetic behaviour. The model, as defined for spin-$\frac{1}{2}$ is integrable, which will be shown in the Chapter 6. If we naively replace the spin-$\frac{1}{2}$ operators by the corresponding operators for spin-1, we obtain a nonintegrable system. However, there is a spin-1 generalization of the Heisenberg model that is integrable which is known as the Babujian-Takhtajan model [3, 19, 47]

$$H_{\mathrm{BT}} = J \sum_{i=1}^{N} \left[\mathbf{S}_i \cdot \mathbf{S}_{i+1} - (\mathbf{S}_i \cdot \mathbf{S}_{i+1})^2 \right].$$

We see that an extra term is added to maintain the integrability. One could hope for a similar situation in the case of the HS model.

If there would be a spin-1 generalization of the HS model, then there would also be a nontrivial representation of the Yangian that is generated by the total spin and some other operator. We assume that this other operator is also constructed out of spin operators. In the case of the spin-$\frac{1}{2}$ HS model, this other operator is the rapidity which contains the product of two spin operators. A heuristic argument shows that this should always be the case. Suppose our second operator has (at most) a term with n spin operators. Then if we plug this in in the Serre relation, the left hand side contains product of at most $3n + 1 - 3 = 3n - 2$ spin operators. Indeed, we get $3n + 1$ spin operators from our products and lose three of them using the commutators. The right hand side contains terms with at most $n + 2$ spin operators. In principle, these numbers do not have to be equal because the number of spin operators could go down as a result of the following:

- Terms could cancel as a result of symmetries in the indices.
- The $\mathfrak{su}(2)$ commutation relations can be used to decrease the number of spin operators by one.
- The Casimir element (or the square of the spin operator on one site in physical language) could also lower the number of spin operators by two.

However, if we for instance assume $n = 3$ the maximum number of spin operators on the left hand side is 7 while on the right hand side it is 4. In that case all terms containing 5, 6 and 7 spin operators should cancel against each other. For higher numbers of n the difference only gets larger. Therefore, we assume that the maximum number of spin operators should be two and will not pursue the case of more spin operators. This allows terms with zero, one and two spin operators. Constant terms prevent the fact that the total operator should transform as a vector under $\mathfrak{su}(2)$ transformations. By translational invariance, the only term with one spin operator that is allowed is the total spin, which can be ignored. As a result, we only have to consider terms that have precisely two spin operators. The most general form for such an operator is

$$\Lambda^a = \sum_{i\neq j} \alpha_{ij} g^{abc} S_i^b S_j^c,$$

where α_{ij} and g^{abc} are arbitrary tensors. Translational and rotational invariance imply that we can write α_{ij} and g^{abc} as separate tensors. Of course, we may assume $\alpha_{ij} g^{abc} = \alpha_{ji} g^{acb}$. Furthermore, α_{ij} should either be antisymmetric or symmetric to avoid breaking translational invariance. We compute

$$[S^a, \Lambda^b] = 2\mathrm{i} \sum_{i,j} \alpha_{ij} g^{bcd} \varepsilon^{ace} S_i^d S_j^e.$$

It turns out that insisting on Eq. (4.4) leads to the requirement that α_{ij} is antisymmetric and that g^{abc} is proportional to ε^{abc}. To see this, one first takes the index values of a and b equal. Then one finds that α_{ij} is either antisymmetric or that g^{cde} vanishes, when the indices are all different. Then picking different choices of a and b one can show that the latter condition leads to $g^{abc} = 0$. Hence, α is antisymmetric. Again, plugging in all possible values for a and b gives relations between the different index values for g^{abc} which has, up to some constant, a unique solution. Therefore, the most general form for Λ^a is in fact equal to

$$\Lambda^a = \frac{\mathrm{i}}{2} \sum_{i\neq j} \alpha_{ij} \varepsilon^{abc} S_i^b S_j^c,$$

with $\alpha_{ij} = -\alpha_{ji}$, where we have rescaled α_{ij}. Numerical calculations show that the Serre relation Eq. (3.3) has no nonzero solutions if we plug in the above ansatz for Λ^a, in the case of spin-1, if the number of particles is six or below. This is already suggested by the calculation we presented in Section 4.4, where we showed that the Serre relation is satisfied for spin-$\frac{1}{2}$. There we found that both sides of Eq. (3.3) only contain terms with four spin operators that act on different sites. All the other terms cancelled. We already noted that this conclusion is based on the fact that in the case of spin-$\frac{1}{2}$ the spin operators also satisfy a Clifford algebra structure. For higher spin, it is still true that the terms with all spin operators acting on different sites on the left and right hand side of the Serre relation are the same. However, the terms with spin operators acting on the same site give problems. In the case of spin-$\frac{1}{2}$ one can use the Clifford algebra structure to simplify those. This no longer works for higher spin and, as the numerical calculations show, the extra terms do not cancel.

Therefore, there are no global spin-1 representations for the Yangian, which also excludes the possibility of a spin-1 chain with Yangian symmetry.

4.6 Summary

In this chapter we have investigated the Haldane-Shastry model, in particular the properties regarding the Yangian symmetry. We have explicitly checked the Yangian symmetry of the model in the case of the fundamental representation $\mathfrak{su}(2)$, which corresponds to spin-$\frac{1}{2}$. We provided a general argument to exclude the possibility of a representation of the Yangian extending a global spin-1 representation, if the components of the total spin should be part of the generators. As a consequence, there does not exist a higher spin generalization of the Haldane-Shastry model.

4.7 Appendix

In the preceding sections we looked at some properties of the Haldane-Shastry model. In particular, we calculated the commutator of the Haldane-Shastry Hamiltonian and the rapidity operator and we checked that the total spin and the rapidity satisfy the Serre relation for the Yangian. In these computations we used some identities regarding complex numbers, which we will prove here.

We will follow the notation of this chapter and set $\eta_j = \exp(2\pi \mathrm{i} j/N)$, $\omega_{ij} = (\eta_i + \eta_j)/(\eta_i - \eta_j)$ and $t_{ij} = |\eta_i - \eta_j|^{-2}$. We want to prove the following identities

$$\sum_{\substack{i=1\\i\neq j}}^{N} \omega_{ij} = 0, \tag{4.25}$$

$$\sum_{\substack{i=1\\i\neq j}}^{N} t_{ij}\omega_{ji} = 0, \tag{4.26}$$

$$\sum_{\substack{i=1\\i\neq j}}^{N} \omega_{ij}\omega_{ji} = \frac{(N-1)(N-2)}{3}. \tag{4.27}$$

We note that we have $\eta_j^N = 1$ and therefore we can write

$$z^N - 1 = \prod_{i=1}^{N}(z - \eta_i) \tag{4.28}$$

since both sides are polynomials of degree N with the same N roots. The coefficient in front of the term z^{N-1} is zero on the left hand side, so it should be zero on the right hand side as well. Hence, we find (for $N \neq 1$)

$$\sum_{i=1}^{N} \eta_i = 0.$$

We look at the reciprocal of Eq. (4.28) and use partial fraction decomposition to find

$$\frac{1}{z^N - 1} = \sum_{i=1}^{N} \frac{\eta_i}{N(z - \eta_i)}$$

and from this one easily solves for

$$\sum_{i=1}^{N} \frac{1}{z - \eta_i} = \frac{N z^{N-1}}{z^N - 1}. \tag{4.29}$$

We are interested in the following expression for $m = 1$, $m = 2$ and $m = 3$,

$$\sum_{i=1}^{N-1} \frac{1}{(\eta_i - 1)^m}. \tag{4.30}$$

Using Cauchy's theorem, we can rewrite the sum in terms of an integral [37]. We define a contour C that encloses all η_i for $1 \leq i \leq N-1$ with winding number 1, but does not enclose 1. The contour C' is a small circle around 1. Then we have

$$\begin{aligned}
\sum_{i=1}^{N-1} \frac{1}{(\eta_i - 1)^m} &= \frac{1}{2\pi i} \sum_{i=1}^{N-1} \oint_C \frac{1}{(z-1)^m} \frac{1}{z-\eta_i} dz \\
&\overset{(i)}{=} \frac{1}{2\pi i} \sum_{i=1}^{N} \oint_C \frac{1}{(z-1)^m} \frac{1}{z-\eta_i} dz \\
&\overset{(ii)}{=} \frac{1}{2\pi i} \oint_C \frac{1}{(z-1)^m} \frac{N z^{N-1}}{z^N - 1} dz \\
&\overset{(iii)}{=} -\frac{1}{2\pi i} \oint_{C'} \frac{1}{(z-1)^m} \frac{N z^{N-1}}{z^N - 1} dz \\
&= \operatorname{Res}(g_m(z), 1),
\end{aligned}$$

where we have defined $g_m(z) = -\frac{1}{(z-1)^m} \frac{N z^{N-1}}{z^N - 1}$ in the last line. At (i) we use that the contour does not contain the point $z = 1$ and therefore the extra term is 0. Then we use Eq. (4.29) at (ii). Finally, at (iii) we note that a circular contour at infinity will vanish since the integrand will go at least as $|z|^{-2}$ to zero. Therefore we integral over the contour $C + C'$ (which gives the same value as the integral over the circular contour at infinity) will vanish. Now we can use the following formula to calculate the residue of g_m at $z = 1$, where we note that the order of the pole at $z = 1$ is equal to $m + 1$.

$$\operatorname{Res}(g_m(z), 1) = \frac{1}{m!} \lim_{z \to 1} \frac{d^m}{dz^m} (z-1)^{m+1} g_m(z).$$

For Eq. (4.30), this leads to $-\frac{1}{2}(N-1)$ if $m = 1$, $-\frac{1}{12}(N^2 - 6N + 5)$ if $m = 2$ and $\frac{1}{8}(N^2 - 4N + 3)$ if $m = 3$. Now we have for Eq. (4.25)

$$\begin{aligned}
\sum_{\substack{i=1 \\ i \neq j}}^{N} \omega_{ij} &= \sum_{\substack{i=1 \\ i \neq j}}^{N} \frac{\eta_i + \eta_j}{\eta_i - \eta_j} \\
&= \sum_{i=1}^{N-1} \frac{\eta_i + 1}{\eta_i - 1} \\
&= \sum_{i=1}^{N-1} \frac{\eta_i - 1}{\eta_i - 1} + 2 \sum_{i=1}^{N-1} \frac{1}{\eta_i - 1} \\
&= N - 1 - 2\frac{N-1}{2} \\
&= 0
\end{aligned}$$

and for Eq. (4.26)

$$\begin{aligned}
\sum_{\substack{i=1 \\ i \neq j}}^{N} t_{ij} \omega_{ji} &= \sum_{\substack{i=1 \\ i \neq j}}^{N} \frac{\eta_i + \eta_j}{(\eta_i - \eta_j)^2} \frac{\eta_i + \eta_j}{\eta_i - \eta_j} \\
&= \sum_{i=1}^{N-1} \frac{\eta_i(\eta_i + 1)}{(\eta_i - 1)^3}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{N-1} \frac{(\eta_i - 1)^2 + 3(\eta_i - 1) + 2}{(\eta_i - 1)^3} \\
&= \sum_{i=1}^{N-1} \frac{1}{\eta_i - 1} + 3 \sum_{i=1}^{N-1} \frac{1}{(\eta_i - 1)^2} + 2 \sum_{i=1}^{N-1} \frac{1}{(\eta_i - 1)^3} \\
&= -\frac{1}{2}(N-1) - \frac{3}{12}(N^2 - 6N + 5) + \frac{2}{8}(N^2 - 4N + 3) \\
&= 0.
\end{aligned}
$$

For Eq. (4.27), one finds similarly

$$
\begin{aligned}
\sum_{\substack{i=1 \\ i \neq j}}^{N} \omega_{ij}\omega_{ji} &= -\sum_{\substack{i=1 \\ i \neq j}}^{N} \frac{(\eta_i + \eta_j)^2}{(\eta_i - \eta_j)^2} \\
&= -\sum_{i=1}^{N-1} \frac{(\eta_i + 1)^2}{(\eta_i - 1)^2} \\
&= -\sum_{i=1}^{N-1} \frac{(\eta_i - 1)^2 - 4(\eta_i - 1) - 4}{(\eta_i - 1)^2} \\
&= \sum_{i=1}^{N-1} 1 + 4 \sum_{i=1}^{N-1} \frac{1}{\eta_i - 1} + 4 \sum_{i=1}^{N-1} \frac{1}{(\eta_i - 1)^2} \\
&= -(N-1) + \frac{4}{2}(N-1) + \frac{4}{12}(N^2 - 6N + 5) \\
&= \frac{(N-1)(N-2)}{3},
\end{aligned}
$$

which completes the proof of the three identities.

Chapter 5

The Yang-Baxter equation

5.1 Introduction

In this chapter we will investigate the Yang-Baxter equation (YBE). This equation can be seen as a consistency equation in a scattering event of three particles. It turns up in various areas in physics, for instance in quantum computation [25] and condensed matter physics [3, 19, 47]. In the next chapter we will look into the latter application of the YBE, in particular by looking for an integrable system invariant under the adjoint representation of $\mathfrak{su}(3)$.

In the first section of this chapter we will define the Yang-Baxter equation and look at its physical meaning. After that we will look at a few different methods to find solutions of the YBE, which are known as R-matrices. In particular, we will be interested in a subclass of solutions, which are the rational R-matrices. The most useful method to find explicit expressions is by Chari and Pressley [10, 11] and makes use of the representation theory of the Yangian. We will review that method in detail.

We present the solutions of the YBE for two simple cases, which are the R-matrices corresponding to two representations of $\mathfrak{su}(2)$. These have been computed before by Babujian and Takhtajan [3, 47]. After that, we present the solution that we found for the R-matrix that is invariant under the adjoint representation of $\mathfrak{su}(3)$. Furthermore, we conjecture a result for the R-matrix corresponding to the adjoint representation of $\mathfrak{su}(n)$ for general n. In the final section, we give a result for the R-matrix on a reducible representation of $\mathfrak{su}(2)$.

5.2 Definition of the Yang-Baxter equation

We start with an abstract definition for the Yang-Baxter equation (YBE). Let A be a unital associative algebra. The YBE has the following form

$$R_{12}(\lambda)\, R_{13}(\lambda+\mu)\, R_{23}(\mu) = R_{23}(\mu)\, R_{13}(\lambda+\mu)\, R_{12}(\lambda)\,. \tag{5.1}$$

This equation holds in $A \otimes A \otimes A$. A solution R is a map from the complex numbers to $A \otimes A$, i.e. $R(\lambda) \in A \otimes A, \forall \lambda \in \mathbb{C}$. The subscripts in Eq. (5.1) imply in which of the spaces the objects are, for instance $R_{12}(\lambda) = R(\lambda) \otimes \mathbf{1}$, with $\mathbf{1}$ the unit of A. The solutions of Eq. (5.1) are called R-matrices. Eq. (5.1) has a physical interpretation in terms of the scattering of particles. It can be seen as a consistency equation in a three-particle scattering event. The R-matrix $R(\lambda)$ does play the role of a scattering matrix of two particles and the parameter λ can be related to the relative momentum. Eq. (5.1) now states that the scattering processes depicted in Fig. 5.1 are

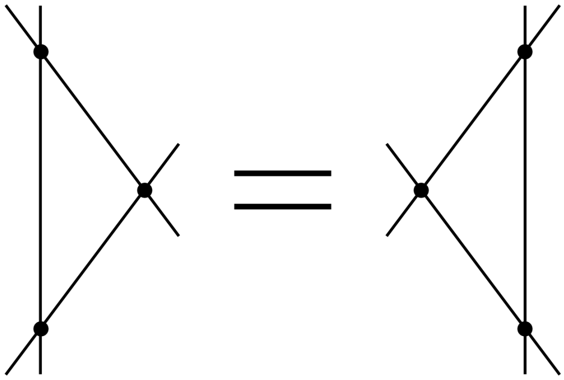

Figure 5.1: A schematic depiction of the Yang-Baxter equation. It tells us that the two scattering processes depicted here lead to the same result.

essentially the same, i.e. scattering of these three particles in reverse order leads to the same result.

We will be particularly interested in the case where $A = GL(V)$, where (ϕ, V) is a finite-dimensional irreducible representation of a simple Lie algebra $\mathfrak{g}$. In this case, we are looking for solutions of Eq. (5.1) that are invariant under the action of the Lie algebra, which means that

$$[\phi(x) \otimes 1 + 1 \otimes \phi(x), R(\lambda)] = 0 \tag{5.2}$$

is satisfied for all elements of the Lie algebra $x \in \mathfrak{g}$ and all $\lambda \in \mathbb{C}$. In the case of $\mathfrak{sl}(n)$, the existence of such solutions of the YBE is guaranteed by a fundamental property of the Yangian, which is the existence of a universal R-matrix [15]. This universal R-matrix is a solution of Eq. 5.1, in the algebra $Y(\mathfrak{g})$. It exists for any Lie algebra, but only in the case of $\mathfrak{sl}(n)$ one can conclude that there is an R-matrix for every irreducible representation. This is a consequence of the evaluation representation (Eq. (3.7)), which can be used to extend every representation of $\mathfrak{sl}(n)$ to a representation of $Y(\mathfrak{sl}(n))$. The R-matrix for an arbitrary irreducible representation can be found by applying this extended representation to the universal R-matrix. However, in practice this method is not very useful when looking for explicit expressions for these R-matrices since an expression for the universal R-matrix, if known, is very involved.

There is another general method to find an R-matrix for an arbitrary finite-dimensional irreducible representation (ϕ, V) of the Yangian $Y(\mathfrak{g})$. This has to do with so called intertwining operators. To define these, we recall from Section 3.3 the construction of representations V_λ for any $\lambda \in \mathbb{C}$, which have the same representation space V but have a different representation map $\phi \circ \tau_\lambda$. We have the following definition for the intertwining operator.

Definition 5.1. *(Intertwining operator) An intertwining operator $I_{V,W}(\lambda, \mu)$ corresponding to two representations V and W of $Y(\mathfrak{g})$ is a map $I_{V,W}(\lambda, \mu) : W_\mu \otimes V_\lambda \to V_\lambda \otimes W_\mu$ such that $I_{V,W}(\lambda, \mu)$ intertwines the action of $Y(\mathfrak{g})$.*

Chari and Pressley showed that there exists an essentially unique nonzero invertible intertwining operator except for a finite set of values $\lambda - \mu$ and that $I_{V,W}(\lambda, \mu)$ is a rational function of $\lambda - \mu$ with values in $\mathrm{Hom}\,(W \otimes V, V \otimes W)$ in the case of $\mathfrak{sl}(n)$ [10]. Here, essentially unique means that it is unique up to an arbitrary function that can be multiplied with the intertwining operator. Since, $I_{V,W}(\lambda, \mu)$ only depends on $\lambda - \mu$, we write $I_{V,W}(\lambda - \mu)$ instead. The following result is also due to Chari and Pressley.

Theorem 5.1. *Let $I_{V,V}(\lambda)$ be an intertwining operator of a representation V and σ the switch operator on $V \otimes V$, i.e. $\sigma(v_1 \otimes v_2) = (v_2 \otimes v_1)$. Then $R(\lambda) = I_{V,V}(\lambda)\sigma$ is an R-matrix corresponding to V.*

This theorem is a very useful method for finding explicit expressions for R-matrices corresponding to representations of the Yangian. We will later see an explicit example in the case of $\mathfrak{su}(3)$. The following lemma ensures that the R-matrices from this construction are indeed invariant under the action of the Lie algebra $\mathfrak{g}$.

Lemma 5.1. *If $I_{V,V}(\lambda)$ is an intertwining operator, then the R-matrix $R(\lambda) = I_{V,V}(\lambda)\sigma$ is invariant under the action of the Lie algebra.*

Proof. The proof consists of two observations. Firstly, all elements of the Lie algebra $\mathfrak{g}$, which is contained in $Y(\mathfrak{g})$, act the same on V_λ, for any $\lambda \in \mathbb{C}$. Therefore, $I_{V,V}(\lambda)$ commutes with the action of $\mathfrak{g}$ on V. Secondly, the action of $\mathfrak{g}$ commutes with σ as well. This is a consequence of the following computation, ($x \in \mathfrak{g}$, $v,w \in V$)

$$\sigma(x \cdot (v \otimes w)) = \sigma(x \cdot v \otimes w + v \otimes x \cdot w) = w \otimes x \cdot v + x \cdot w \otimes v = x \cdot (w \otimes v) = x \cdot \sigma(v \otimes w).$$

□

In the next section we will see two examples of R-matrices for two representations of $\mathfrak{su}(2)$. These can be obtained by less complicated techniques that do not involve the intertwining operator. Later on, we will see a more complicated example when we look at the adjoint representation of $\mathfrak{su}(3)$. In that case, we will use the construction of an intertwining operator to find the R-matrix.

5.3 R-matrices for spin representations

In this section we will see an example of an R-matrix on two representations of $\mathfrak{su}(2)$. First, we look at the spin-$\frac{1}{2}$ representation of $\mathfrak{su}(2)$. We will shortly write $\mathbf{s}$ for the spin-s representation, which is the $(2s+1)$-dimensional irreducible representation of $\mathfrak{su}(2)$. In this notation we have $\boldsymbol{\frac{1}{2}} \otimes \boldsymbol{\frac{1}{2}} = \mathbf{1} \oplus \mathbf{0}$. This is a restatement of the familiar fact from quantum mechanics that two spin-$\frac{1}{2}$ particles can either form a singlet ($\mathbf{0}$) or a triplet ($\mathbf{1}$). We claim that there exists an R-matrix $R(\lambda)$ on $\boldsymbol{\frac{1}{2}} \otimes \boldsymbol{\frac{1}{2}}$ of the following form

$$R(\lambda) = P_{\mathbf{1}} + f(\lambda)P_{\mathbf{0}}, \tag{5.3}$$

where $P_{\mathbf{s}}$ is the projection onto $\mathbf{s}$ and $f(\lambda)$ is an unknown function. The existence of the Lie algebra invariant R-matrix follows from the results in the previous section. If v is a highest weight vector of weight λ in $\boldsymbol{\frac{1}{2}} \otimes \boldsymbol{\frac{1}{2}}$, then $R(\lambda)(v)$ will be a highest weight vector of the same weight. So the module generated by $R(\lambda)(v)$ is isomorphic to the module generated by v. By Schur's lemma, $R(\lambda)$ should act by a scalar on these modules, since there are no multiplicities in the decomposition of $\boldsymbol{\frac{1}{2}} \otimes \boldsymbol{\frac{1}{2}}$. Since we can choose the normalization of $R(\lambda)$, we can put one of these scalars to 1 and the claim follows.

To find the function $f(\lambda)$ one can plug the ansatz in Eq. (5.3) into the Yang-Baxter equation from Eq. (5.1). One finds that

$$f(\lambda) = \frac{\lambda + \eta}{\lambda - \eta},$$

where $\eta \in \mathbb{C}$ is a free parameter. Here we could have made use of the intertwining operator from the previous section, but in this simple case using these more advanced techniques is not necessary. We note that with this normalization we have $R(0) = \sigma$. The R-matrix can be given in terms of the Pauli matrices by expressing the projectors $P_{\mathbf{s}}$ in the Casimir element

$$C = S^a \otimes S^a$$

with $S^a = \frac{1}{2}\sigma^a$ the spin operators which equal half the Pauli matrices.

Now we use the same method for the spin-1 representation of $\mathfrak{su}(2)$. In the end, we will find an R-matrix that is invariant under the spin-1 representation of $\mathfrak{su}(2)$. We note that this result can not be obtained by heuristically replacing the spin-$\frac{1}{2}$ in the R-matrix for the spin-$\frac{1}{2}$ representation by their spin-1 counterparts. The tensor product of two spin-1 representations decomposes as

$$\mathbf{1} \otimes \mathbf{1} = \mathbf{2} \oplus \mathbf{1} \oplus \mathbf{0} \tag{5.4}$$

and therefore, following similar considerations as in the previous case for the spin-$\frac{1}{2}$ representation, the R-matrix can be written as

$$R(\lambda) = P_{\mathbf{2}} + f(\lambda)P_{\mathbf{1}} + g(\lambda)P_{\mathbf{0}}, \tag{5.5}$$

with f and g unknown functions. Although it is considerably more difficult than in the spin-$\frac{1}{2}$, f and g can still be determined by plugging this ansatz into the Yang-Baxter equation. One finds that

$$f(\lambda) = \frac{\lambda + 2\eta}{\lambda - 2\eta}, \; g(\lambda) = \frac{(\lambda + \eta)(\lambda + 2\eta)}{(\lambda - \eta)(\lambda - 2\eta)}.$$

We note that

$$R(0) = P_{\mathbf{2}} - P_{\mathbf{1}} + P_{\mathbf{0}} = \sigma,$$

which was also the case for the spin-$\frac{1}{2}$ invariant R-matrix. This is an important fact that will be very useful in the construction of integrable models in the next chapter. Up to a possible normalization, it is always true for a rational R-matrix constructed from an intertwining operator for the Yangian representation. The reason for this is that the intertwining operator for $\lambda = 0$ intertwines the same two irreducible Yangian representations. Therefore, $I(0)$ is a multiple of the identity and $R(0)$ of the permutation operator. For higher-dimensional representations of $\mathfrak{su}(2)$ the results are similar. A general expression for the R-matrix corresponding to the spin-s representation is given in Refs. [3] and [10].

5.4 R-matrices for the adjoint representation of $\mathfrak{su}(3)$

In the previous sections we looked at the solutions of the Yang-Baxter equation that were invariant under $\mathfrak{su}(2)$, which were previously obtained by Babujian [3]. Now, we will look at the adjoint representation of $\mathfrak{su}(3)$ in more detail to find a new solution of the YBE. We will often refer to some facts on $\mathfrak{su}(3)$ that we already gave in Section 2.3 and we we will use the notation that was introduced there.

For notational convenience we will write only the dimension for the representations of $\mathfrak{su}(3)$. For instance, we write $\mathbf{8}$ instead of $V(1,1)$ for the adjoint representation of $\mathfrak{su}(3)$. Since the dimensions of the representations $V(m_1, m_2)$ and $V(m_2, m_1)$ are equal as a consequence of Lemma 2.4, we distinguish these by writing a bar on top of the latter if $m_2 > m_1$. For instance $\mathbf{3}$ denotes the representation $V(1,0)$ while $\bar{\mathbf{3}}$ corresponds to $V(0,1)$.

$\mathbf{27}$	$x_{12} \otimes x_{12}$
$\mathbf{10}$	$x_{12} \wedge x_1$
$\bar{\mathbf{10}}$	$x_{12} \wedge x_2$
$\mathbf{8}_s$	$\mathrm{Sym}\,(3x_1 \otimes x_2 + h_1 \otimes x_{12} - h_2 \otimes x_{12})$
$\mathbf{8}_a$	$\sqrt{5}\,(x_1 \wedge x_2 + h_1 \wedge x_{12} + h_2 \wedge x_{12})$
$\mathbf{1}$	C

Table 5.1: The highest weight vectors for the submodules in the decomposition of $\mathbf{8} \otimes \mathbf{8}$, where $v \wedge w = v \otimes w - w \otimes v$ and $\mathrm{Sym}(v \otimes w) = v \otimes w + w \otimes v$.

The representation $\mathbf{8} \otimes \mathbf{8}$ decomposes as follows,

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \mathbf{10} \oplus \bar{\mathbf{10}} \oplus \mathbf{8}_s \oplus \mathbf{8}_a \oplus \mathbf{1}. \tag{5.6}$$

Here, we distinguish the two copies of $\mathbf{8}$, since one is symmetric and the other is antisymmetric, meaning that the permutation operator acts by 1 and -1 respectively on these modules. Since $\mathfrak{su}(3)$ is a simple Lie algebra, we have $[\mathfrak{su}(3), \mathfrak{su}(3)] = \mathfrak{su}(3)$, so $\mathfrak{su}(3)$ acts by 0 on the one-dimensional representation $\mathbf{1}$. Therefore, this representation is spanned by the Casimir element, which is given by

$$C = \mathrm{Sym}\,(3\,(x_1 \otimes y_1 + x_2 \otimes y_2 + x_{12} \otimes y_{12}) + h_1 \otimes h_1 + h_1 \otimes h_2 + h_2 \otimes h_2)\,, \tag{5.7}$$

where Sym is the linear operator such that $\mathrm{Sym}(x \otimes y) = x \otimes y + y \otimes x$. The definitions of the elements of $\mathfrak{su}(3)$ are given in Sec. 2.3. The other highest weight vectors of all submodules in (5.6) are given in Table 5.1, where we have introduced the notation $x \wedge y = x \otimes y - y \otimes x$. These results can be checked by straightforward computations. We have normalized the antisymmetric highest weight vector for the adjoint representation with a factor of $\sqrt{5}$. As a result of this, the two highest weight vector for the two $\mathbf{8}$ modules have the same norm with respect to the trace inner product $(x, y) = \mathrm{tr}(x^\dagger y)$ which is the unique inner product (up to normalization) for which the corresponding Lie group representation is unitary.

To find the R-matrix, we follow the approach by Chari and Pressley [10, 11], based on the intertwining operator $I(\lambda)$ for the Yangian representation. The intertwining operator commutes with the action of the Lie algebra on $\mathfrak{su}(3) \otimes \mathfrak{su}(3)$ as a result of Lemma 5.1. This implies that if v is a highest weight vector, then $I(\lambda)(v)$ is a highest weight vector of the same weight. So, if we restrict $I(\lambda)$ to one of the modules $\mathbf{27}$, $\mathbf{10}$, $\bar{\mathbf{10}}$, $\mathbf{1}$, then this restricted operator is a multiple of the identity by Schur's lemma. For the two $\mathbf{8}$ representations, the situation is a little more complicated. If we denote the highest weight vectors of these modules by $v_{\mathbf{8}_s}$ and $v_{\mathbf{8}_a}$, then we know that $I(\lambda)$ acts as a (2×2)-matrix on these highest weight vectors, i.e.

$$I(\lambda) \begin{pmatrix} v_{\mathbf{8}_s} \\ v_{\mathbf{8}_a} \end{pmatrix} = \begin{pmatrix} m_{11}(\lambda) & m_{12}(\lambda) \\ m_{21}(\lambda) & m_{22}(\lambda) \end{pmatrix} \begin{pmatrix} v_{\mathbf{8}_s} \\ v_{\mathbf{8}_a} \end{pmatrix}.$$

Using the Lie algebra invariance of the intertwining operators, then the operator acts by this matrix on the complete $\mathbf{8} \oplus \mathbf{8}$ module, again by Schur's lemma. Therefore, $I(\lambda)$ can be written in the following form

$$\begin{aligned} I(\lambda) = {} & f_{27}(\lambda) P_{\mathbf{27}} + f_{10}(\lambda) P_{\mathbf{10}} + f_{\bar{10}}(\lambda) P_{\bar{\mathbf{10}}} \\ & + m_{11}(\lambda) P_{\mathbf{8}_s} + m_{12}(\lambda) M_{as} + m_{21}(\lambda) M_{sa} + m_{22}(\lambda) P_{\mathbf{8}_a} + f_1(\lambda) P_{\mathbf{1}}. \end{aligned} \tag{5.8}$$

Here $P_{\mathbf{s}}$ is the projection onto the submodule $\mathbf{s}$, while the operator M_{sa} (M_{as}) is the unique operator sending $v_{\mathbf{8}_s}$ ($v_{\mathbf{8}_a}$) to $v_{\mathbf{8}_a}$ ($v_{\mathbf{8}_s}$) that intertwines the Lie algebra representation. The objects $f_s(\lambda)$ are scalar functions that need to be determined. We set $f_{27}(\lambda) = 1$ as the normalization.

For this determination, we need the explicit action of all the Yangian generators on the representations $\mathbf{8}_\lambda \otimes \mathbf{8}$ and $\mathbf{8} \otimes \mathbf{8}_\lambda$. The action of an element x of the Yangian is given by

$$x \cdot (v \otimes w) = \Delta(x)(v \otimes w)$$

where $\Delta(x)$ acts by components on $v \otimes w$. The explicit expression for the comultiplication is given in Eq. (3.5). It follows from Eqs. (3.6) and (3.7) that the action of $J(x)$ on $\mathbf{8}_\lambda \otimes \mathbf{8}$ is given by

$$J(x)_{\lambda,0}\cdot(v\otimes w) = \frac{1}{4}(\mathrm{tr}(x\{I_\mu, I_\nu\})I_\mu I_\nu)+\lambda x)\cdot v\otimes w+\frac{1}{4}v\otimes((\mathrm{tr}(x\{I_\mu, I_\nu\})I_\mu I_\nu)\cdot w+\frac{1}{2}[x\otimes 1, C]\cdot(v\otimes w). \tag{5.9}$$

From Table 5.1 we see that the highest weight vector for the $\mathbf{10}$ submodule in $\mathbf{8} \otimes \mathbf{8}$ is given by $x_{12} \wedge x_1$. We compute the action of $J(x_2)$ on this highest weight vector in the representation $\mathbf{8}_\lambda \otimes \mathbf{8}$, using Eq. (5.9). The basis of $\mathfrak{su}(3)$ that is orthonormal with respect to the trace form is given by the Gell-Mann matrices in Eq. (2.3), scaled with a factor of $\frac{1}{\sqrt{2}}$. Using the expression for the anticommutator in Eq. 2.4, we see that we are interested in the expression $\mathrm{tr}(x_2 T^a)$, which is only nonzero if a equals 6 or 7 with values 1 and i respectively. This is a consequence of the relations between the Gell-Mann matrices and x_2 that are given at the end of Section 2.3. The nonzero d-symbols containing 6 and 7 are

$$d^{668} = d^{778} = -\frac{1}{\sqrt{3}}, \qquad d^{146} = d^{157} = -d^{247} = d^{256} = -d^{366} = -d^{377} = 1.$$

Using these expressions and the fact that the elements of $\mathfrak{su}(3)$ act in the adjoint representation by the commutator, one computes that

$$\begin{aligned} &\mathrm{tr}(x_2\{I_\mu, I_\nu\})I_\mu I_\nu \cdot x_{12} = 0 \\ &\mathrm{tr}(x_2\{I_\mu, I_\nu\})I_\mu I_\nu \cdot x_1 = \frac{3}{4}x_{12} \end{aligned}$$

and therefore the terms on the right hand side of Eq. (5.9) containing the trace give no contribution to $J(x_2) \cdot (x_{12} \otimes x_1 - x_1 \otimes x_{12})$ as a result of a cancellation.

The other terms in Eq. (5.9) can be computed rather easily. The Casimir element acts by 2 on $x_{12} \otimes x_{12}$, while it kills $x_{12} \wedge x_1$. Therefore

$$\begin{aligned} \frac{1}{2}[x_2 \otimes 1, C](x_{12} \wedge x_1) &= -\frac{1}{2}C(x_2 \otimes 1) \cdot (x_{12} \wedge x_1) \\ &= \frac{1}{2}C(x_2 \otimes 1) \cdot (x_1 \otimes x_{12}) \\ &= -\frac{1}{2}C \cdot (x_{12} \otimes x_{12}) = -x_{12} \otimes x_{12}. \end{aligned}$$

Using Eq. (5.9) and collecting all the contributions, we see that

$$J(x_2)_{\lambda,0} \cdot (x_{12} \wedge x_1) = (-1 + \lambda)x_{12} \otimes x_{12}. \tag{5.10}$$

If we would do the same calculation on the representation $\mathbf{8} \otimes \mathbf{8}_\lambda$ then the term proportional to λ changes sign. So in that case we have

$$J(x_2)_{0,\lambda} \cdot (x_{12} \wedge x_1) = (-1 - \lambda)x_{12} \otimes x_{12}. \tag{5.11}$$

We can compute one of the coefficients of the intertwining operator from this. By definition, it satisfies

$$J(x_2)_{\lambda,0}I(\lambda) = I(\lambda)J(x_2)_{0,\lambda}. \tag{5.12}$$

We let both sides of this equation act on $x_{12} \wedge x_1$. Using Eqs. (5.8), (5.10) and (5.11), it is clear that

$$(-1+\lambda)f_{10}(\lambda) = (-1-\lambda)f_{27}(\lambda). \tag{5.13}$$

We note that we had assumed $f_{27}(\lambda) = 1$ and therefore we have

$$f_{10}(\lambda) = -\frac{\lambda+1}{\lambda-1}. \tag{5.14}$$

For the other components we find similar results. We note that $J(x_1)_{0,\lambda}.v_{\bar{\mathbf{10}}} = (-1-\lambda)v_{\mathbf{27}}$ and $J(x_1)_{\lambda,0}.v_{\bar{\mathbf{10}}} = (-1+\lambda)v_{\mathbf{27}}$, leading to

$$f_{\bar{10}}(\lambda) = -\frac{\lambda+1}{\lambda-1}.$$

Furthermore, we have $x_{12}J(x_{12})_{0,\lambda}J(x_1)_{0,\lambda}.v_{\mathbf{1}} = 6(\lambda+1)(\lambda+3).v_{\mathbf{27}}$ and $x_{12}J(x_{12})_{\lambda,0}J(x_1)_{\lambda,0}.v_{\mathbf{1}} = 6(\lambda-1)(\lambda-3).v_{\mathbf{27}}$ and we find

$$f_1 = \frac{(\lambda+1)(\lambda+3)}{(\lambda-1)(\lambda-3)}.$$

Here, we have used the notation that $v_{\mathbf{s}}$ is the highest weight vector corresponding to the submodule $\mathbf{s}$. For the two copies of $\mathbf{8}$ the situation is somewhat more involved. The following identities can be obtained by explicit computation,

$$\begin{aligned}
J(x_{12})_{\lambda,\mu}v_{\mathbf{8}_s} &= 3v_{\mathbf{27}},\\
J(x_{12})_{\lambda,\mu}v_{\mathbf{8}_a} &= (5-2\lambda+2\mu)v_{\mathbf{27}},\\
J(x_2)_{0,\lambda}J(x_2)_{0,\lambda}v_{\mathbf{8}_s} &= \left(\frac{27}{4}+3\lambda\left(2+\lambda\right)\right)v_{\mathbf{27}},\\
J(x_2)_{\lambda,0}J(x_2)_{\lambda,0}v_{\mathbf{8}_s} &= \frac{3}{4}\left(3-2\lambda\right)^2 v_{\mathbf{27}},\\
J(x_2)_{0,\lambda}J(x_2)_{0,\lambda}v_{\mathbf{8}_a} &= \frac{\sqrt{5}}{4}\left(21+2\lambda-4\lambda^2\right)v_{\mathbf{27}},\\
J(x_2)_{\lambda,0}J(x_2)_{\lambda,0}v_{\mathbf{8}_a} &= \frac{\sqrt{5}}{4}\left(21-22\lambda+4\lambda^2\right)v_{\mathbf{27}}.
\end{aligned}$$

These eight expressions lead to four equations for m_{11}, m_{12}, m_{21} and m_{22}. The unique solution for those coefficients is given by

$$\begin{aligned}
m_{11}(\lambda) &= \frac{2\lambda^3-11\lambda-6}{2(\lambda-1)^2(\lambda-3)},\\
m_{12}(\lambda) &= \frac{3\sqrt{5}\lambda}{2(\lambda-1)^2(\lambda-3)},\\
m_{21}(\lambda) &= -\frac{3\sqrt{5}\lambda}{2(\lambda-1)^2(\lambda-3)},
\end{aligned}$$

$$m_{22}(\lambda) = \frac{-2\lambda^3 + 11\lambda - 6}{2(\lambda-1)^2(\lambda-3)}.$$

We note that this (2×2)-matrix is diagonalizable, but the resulting diagonal elements are not rational functions. Furthermore, the off-diagonal elements equal each other up to a minus sign.

The R-matrix can now easily be obtained by multiplying the intertwining operator with the permutation operator. From Table 5.1, we can easily see that the module **27**, $\mathbf{8}_s$ and **1** are symmetric and **10**, $\bar{\mathbf{10}}$ and $\mathbf{8}_a$ are antisymmetric. As a result, the R-matrix for the adjoint representation becomes

$$\begin{aligned} R(\lambda) = {} & P_{\mathbf{27}} - f_{10}(\lambda)P_{\mathbf{10}} - f_{\bar{10}}(\lambda)P_{\bar{\mathbf{10}}} \\ & + m_{11}(\lambda)P_{\mathbf{8}_s} - m_{12}(\lambda)M_{as} + m_{21}(\lambda)M_{sa} - m_{22}(\lambda)P_{\mathbf{8}_a} + f_1(\lambda)P_{\mathbf{1}}. \end{aligned} \tag{5.15}$$

We have explicitly checked that this R-matrix solves the YBE using a computer program. This result seems to have been obtained before by Alihauskas and Kulish, although it was written down in a different form [1]. In the form we have written down, the R-matrix is manifestly Hermitian if the spectral parameter λ is real. This is a result of the normalization of the highest weight vectors for the two **8** modules. If we had chosen a different normalization, the off-diagonal elements in the (2×2)-matrix acting on these modules would only equal each other up to a scalar factor. In the earlier examples for $\mathfrak{su}(2)$, the R-matrix was always Hermitian (since it is the sum of projections multiplied by real functions) and the explicit form did not depend on the normalization of the highest weight vectors.

We note that a more general rational R-matrix can be obtained by rescaling the parameter λ with a factor of $1/\eta$. Since the YBE is invariant under such a rescaling, the new R-matrix will still be a solution. Furthermore, a second independent solution of the YBE is given by $\tilde{R}(\lambda) = \sigma I(\lambda)$.

5.5 A conjecture on the adjoint representation of $\mathfrak{su}(n)$

In this section we try to generalize the result of the previous two sections by determining the rational R-matrix for the adjoint representation of $\mathfrak{su}(n)$ if $n > 3$. For $n = 2$ and $n = 3$ we have seen the result in the previous sections. We have obtained this R-matrix for general n, although it follows from a set of equations that we only have proven up to $n = 7$, using a computer program.

The highest weight of the adjoint representation of any simple Lie algebra is given by the maximal root. In the case of $\mathfrak{g} = \mathfrak{su}(n)$, this maximal root is given by $\beta = \alpha_1 + \cdots + \alpha_{n-1}$, which is the sum of all the simple roots. In terms of the fundamental weights, this can be written as $\beta = \lambda_1 + \lambda_n$ [24, 32]. We will use the notation $(m_1 m_2 \cdots m_{n-1})$ for the representation of highest weight $\sum_i^{n-1} m_i\lambda_i$. The tensor product representation of two copies of the adjoint representation decomposes as [32]

$$\begin{aligned} (10\cdots 01)\otimes(10\cdots 01) = {} & (20\cdots 02) \oplus (20\cdots 010) \oplus (010\cdots 02) \\ & \oplus (010\cdots 010) \oplus (10\cdots 01) \oplus (10\cdots 01) \oplus (0\cdots 0). \end{aligned} \tag{5.16}$$

In the case of $n = 3$, the submodule $(010\cdots 010)$ would be absent, while the modules $(20\cdots 010)$ and $(010\cdots 02)$ have to be interpreted as (30) and (03) respectively. When $n = 4$ the module $(010\cdots 010)$ should read (020). We have computed explicit expressions for the highest weight vectors of the submodules on the right hand side. To give these we introduce the following notation. We denote the matrix with 1 as the $(i,j)^{\text{th}}$ entry and 0 in all the other entries by

e_{ij}. The positive root spaces of $\mathfrak{su}(n)$ can be chosen in such a way that they correspond to the one-dimensional subspaces of $\mathfrak{su}(n)$ spanned by e_{ij} with $i < j$. With this choice of positive roots, the root space $\mathfrak{g}_{\alpha_1}$, with α_1 a root corresponding to one of the end nodes in the Dynkin diagram, is either spanned by e_{12} or $e_{(n-1)n}$ of which e_{12} is the most natural one. The highest weight vector of the adjoint representation is then given by e_{1n}.

The highest weight vectors of the subrepresentations on the right hand side of Eq. 5.16 are given in Table 5.1. Here, we recall the notations $v \wedge w = v \otimes w - w \otimes v$ and $\mathrm{Sym}(v \otimes w) = v \otimes w + w \otimes v$. The $(n \times n)$-identity matrix is denoted by $\mathbb{1}$. The Casimir element C is given by $C = I^a \otimes I^a$ where $\{I^a\}$ denotes a basis of $\mathfrak{g}$ that is orthonormal with respect the trace form. For instance, one could use the (normalized) generalized Gell-Mann matrices for this [8]. We recall that in such a basis, the commutator and anticommutator of two elements can be written as

$$[I^a, I^b] = f^{abc} I^c, \quad \{I^a, I^b\} = d^{abc} I^c + \frac{2}{n} \delta^{ab} \mathbb{1}, \tag{5.17}$$

where f^{abc} are the antisymmetric structure constants and d^{abc} are the symmetric d-symbols. From this, one sees that $\{C, \mathbb{1} \otimes e_{1n}\} - \frac{2}{n} e_{1n} \otimes \mathbb{1}$ indeed belongs to $\mathfrak{g} \otimes \mathfrak{g}$. The fact that each of these vector in the table below are highest weight vectors of the correct weights can easily be checked by straightforward calculations. For instance, $e_{1(n-1)} \wedge e_{1n}$ has weight

$$(\alpha_1 + \cdots + \alpha_{n-2}) + (\alpha_1 + \cdots + \alpha_{n-1}) = 2\lambda_1 + \lambda_{n-2}. \tag{5.18}$$

All positive root spaces kill both $e_{1(n-1)}$ and e_{1n}, except for $e_{(n-1)n}$ which only kills e_{1n}. However, we have

$$e_{(n-1)n} \cdot e_{1(n-1)} = [e_{(n-1)n}, e_{1(n-1)}] = -e_{1n},$$

so $e_{(n-1)n} \cdot \left(e_{1(n-1)} \wedge e_{1n}\right) = -e_{1n} \wedge e_{1n} = 0$. We conclude that $e_{1(n-1)} \wedge e_{1n}$ is indeed a highest weight vector of the correct weight. The other calculations are similar. The eigenvalues of the Casimir element C have been computed before [32] and are also given in Table 5.1.

Like in the previous section, the highest weight vectors for the two modules isomorphic to the adjoint representation are normalized such that they have the same norm with respect to the unique inner product on $\mathfrak{su}(n) \otimes \mathfrak{su}(n)$ such that the corresponding Lie group representation is unitary, which is given by $(v, w) = \mathrm{tr}(v^\dagger w)$. For this, the generalized Gell-Mann matrices are orthogonal and all have the same norm. By writing the two highest weight vectors as $v_s = n w^a d^{abc} I^b \otimes I^c$ and $v_a = \sqrt{n^2 - 4}\, w^a f^{abc} I^b \otimes I^c$, where w^a is the unique vector such that $e^{1n} = w^a I^a$, the norms can be computed easily using the following formulas that were obtained in [31]

$$d^{abc} d^{abd} = \frac{2n^2 - 8}{n} \delta^{cd}, \;\; f^{abc} f^{abd} = -2n \; \delta^{cd}.$$

Furthermore, one should note that the d-symbols are always real, while the structure constants are purely imaginary.

We compute the R-matrix using the construction of an intertwining operator as in the previous section in the case with $n = 3$. Using an argument involving Schur's lemma, the intertwining operator can be written as

$$\begin{aligned} I(\lambda) =& P_{2\lambda_1 + 2\lambda_{n-1}} + f_1(\lambda) P_{2\lambda_1 + \lambda_{n-2}} + f_2(\lambda) P_{\lambda_2 + 2\lambda_{n-1}} + f_3(\lambda) P_{\lambda_2 + \lambda_{n-2}} + f_4(\lambda) P_0 \\ &+ M_{\lambda_1 + \lambda_{n-1}}(\lambda). \end{aligned} \tag{5.19}$$

Submodule	Highest weight vector	Eigenvalue of C
$(20\cdots02)$	$e_{1n}\otimes e_{1n}$	2
$(20\cdots010)$	$e_{1(n-1)}\wedge e_{1n}$	0
$(010\cdots02)$	$e_{2n}\wedge e_{1n}$	0
$(010\cdots010)$	$\mathrm{Sym}\left(e_{2(n-1)}\otimes e_{1n}-e_{1(n-1)}\otimes e_{2n}\right)$	-2
$(10\cdots01)$	$n\left(\{C,\mathbb{1}\otimes e_{1n}\}-\frac{2}{n}e_{1n}\otimes\mathbb{1}\right)$	$-n$
$(10\cdots01)$	$\sqrt{n^2-4}[C,\mathbb{1}\otimes e_{1n}]$	$-n$
$(0\cdots0)$	C	$-2n$

Table 5.2: The highest weight vectors for the submodules in the decomposition of Eq. (5.16)

where P_Λ is the projection onto the submodule with highest weight Λ. Furthermore, $M_{\lambda_1+\lambda_{n-1}}(\lambda)$ is an operator on the direct sum of the two submodules that are isomorphic to the adjoint representation and can be written as a (2×2)-matrix, acting on the two highest weight vectors for these submodules, as given in Table 5.2. To compute the functions $f_i(\lambda)$ and the entries of $M_{\lambda_1+\lambda_{n-1}}(\lambda)$ we need to compute the action of some of the Yangian generators on the highest weight vectors. The Yangian generators act as follows on the elements of an arbitrary representation as

$$J(x)_{\lambda,\mu}=\left(\frac{1}{4}\mathrm{tr}(x\{I^a,I^b\})I^aI^b+\lambda x\right)\otimes\mathbb{1}+\mathbb{1}\otimes\left(\frac{1}{4}\mathrm{tr}(x\{I^a,I^b\})I^aI^b+\mu x\right)+\frac{1}{2}[x\otimes\mathbb{1},C], \tag{5.20}$$

which is the same as in the previous section. We conjecture the following expression for the intertwining operator

$$\begin{aligned}I(\lambda)=&P_{2\lambda_1+2\lambda_{n-1}}-\frac{\lambda+1}{\lambda-1}\left(P_{2\lambda_1+\lambda_{n-2}}+P_{\lambda_2+2\lambda_{n-1}}\right)+\left(\frac{\lambda+1}{\lambda-1}\right)^2P_{\lambda_2+\lambda_{n-2}}\\&+\frac{\lambda+1}{\lambda-1}\frac{\lambda+n}{\lambda-n}P_0+M_{\lambda_1+\lambda_{n-1}}(\lambda),\end{aligned} \tag{5.21}$$

where $M_{\lambda_1+\lambda_{n-1}}(\lambda)$, written as a matrix acting on the two copies of the adjoint representation, is given by

$$M_{\lambda_1+\lambda_{n-1}}(\lambda)=\begin{pmatrix}-\frac{2n+(2+n^2)\lambda-2\lambda^3}{2(\lambda-n)(\lambda-1)^2} & \frac{n\sqrt{(n^2-4)}\lambda}{2(\lambda-n)(\lambda-1)^2}\\ -\frac{n\sqrt{n^2-4}\lambda}{2(\lambda-n)(\lambda-1)^2} & -\frac{2n-(2+n^2)\lambda+2\lambda^3}{2(\lambda-n)(\lambda-1)^2}\end{pmatrix}. \tag{5.22}$$

We note that in the case of $n=3$ the term containing the projector $P_{\lambda_2+\lambda_{n-2}}$ is absent, since the corresponding submodule is not there. Plugging in $n=3$ leads to the result obtained in the previous section for the other coefficients. All these expressions are a consequence of the following identities, which we have explicitly checked up to $n=7$. Here, the highest weight vectors of weight Λ are denoted by v_Λ.

$$\left(\mathrm{tr}(e_{(n-1)n}\{I^a,I^b\})I^aI^b\otimes\mathbb{1}+\mathbb{1}\otimes\mathrm{tr}(e_{(n-1)n}\{I^a,I^b\})I^aI^b\right)\cdot v_{2\lambda_1+\lambda_{n-2}}=0, \tag{5.23}$$

$$\left(\mathrm{tr}(e_{12}\{I^a,I^b\})I^aI^b\otimes\mathbb{1}+\mathbb{1}\otimes\mathrm{tr}(e_{12}\{I^a,I^b\})I^aI^b\right)\cdot v_{\lambda_2+2\lambda_{n-1}}=0, \tag{5.24}$$

$$\left(\mathrm{tr}(e_{12}\{I^a,I^b\})I^aI^b\otimes\mathbb{1}+\mathbb{1}\otimes\mathrm{tr}(e_{12}\{I^a,I^b\})I^aI^b\right)\cdot v_{\lambda_2+\lambda_{n-2}}=0, \tag{5.25}$$

$$\left(J(e_{1n})_{\lambda,\mu}\right)^2\cdot v_0=-2(\lambda-\mu-1)(\lambda-\mu-n)v_{2\lambda_1+2\lambda_{n-1}}, \tag{5.26}$$

$$J(e_{1n})_{\lambda,\mu}\cdot v_{\lambda_1+\lambda_{n-1},s}=n(n-2)v_{2\lambda_1+2\lambda_{n-1}}, \tag{5.27}$$

$$J(e_{1n})_{\lambda,\mu} \cdot v_{\lambda_1+\lambda_{n-1},a} = \sqrt{n^2-4}(n+2-2\lambda+2\mu)v_{2\lambda_1+2\lambda_{n-1}}, \tag{5.28}$$

$$J\left(e_{(n-1)n}\right)_{\lambda,\mu} J\left(e_{1(n-1)}\right)_{\lambda,\mu} \cdot v_{\lambda_1+\lambda_{n-1},s} = n\left[(\lambda-\mu)^2+2\mu-n\lambda+\frac{1}{4}n^2\right] v_{2\lambda_1+2\lambda_{n-1}}, \tag{5.29}$$

$$J\left(e_{(n-1)n}\right)_{\lambda,\mu} J\left(e_{1(n-1)}\right)_{\lambda,\mu} \cdot v_{\lambda_1+\lambda_{n-1},a} = \frac{\sqrt{n^2-4}}{4}\left[n(n+4)-(6n+4)\lambda+4\lambda^2+(2n-4)\mu-4\mu^2\right] v_{2\lambda_1+2\lambda_{n-1}} \tag{5.30}$$

The intertwining operator satisfies the equation

$$I(\lambda)J(x)_{0,\lambda} = J(x)_{\lambda,0}I(\lambda) \tag{5.31}$$

Using Eq. (5.21), one can now calculate the functions f_1, f_2, f_3 and f_4 and the entries of $M_{\lambda_1+\lambda_{n-1}}(\lambda)$, by acting with Eq. (5.31) on all the highest weight vectors. For instance, to compute f_1 one first needs to compute the action of $J(e_{(n-1)n})$ on $v_{2\lambda_1+\lambda_{n-2}}$, using (5.23). Noting that

$$(e_{(n-1)n} \otimes \mathbb{1}) \cdot v_{2\lambda_1+\lambda_{n-2}} = -v_{2\lambda_1+2\lambda_{n-1}}$$

and the action of C from Table 5.2, one immediately finds that

$$J(e_{(n-1)n})_{\lambda,\mu} v_{2\lambda_1+\lambda_{n-2}} = (-\lambda+\mu+1)v_{2\lambda_1+2\lambda_{n-1}} \tag{5.32}$$

and therefore we find that

$$(\lambda+1)v_{2\lambda_1+2\lambda_{n-1}} = (-\lambda+1)f_1(\lambda)v_{2\lambda_1+2\lambda_{n-1}} \tag{5.33}$$

and therefore $f_1(\lambda) = -(\lambda+1)/(\lambda-1)$. The functions f_2, f_3 and f_4 can be computed similarly using Eqs. (5.24), (5.25) and (5.26). The entries of the matrix $M_{\lambda_1+\lambda_{n-1}}(\lambda)$ follow from Eqs. (5.27), (5.28), (5.29) and (5.30). These equations lead to four conditions on the entries, for which there is a unique solution.

The R-matrix is now obtained by applying the intertwining operator to the switch operator [10]. From Table 5.2, it is immediate which submodules are symmetric and which are antisymmetric. The resulting R-matrix for the adjoint representation of $\mathfrak{su}(n)$ is given by

$$\begin{aligned} R(\lambda) =& P_{2\lambda_1+2\lambda_{n-1}} + \frac{\lambda+1}{\lambda-1}\left(P_{2\lambda_1+\lambda_{n-2}} + P_{\lambda_2+2\lambda_{n-1}}\right) + \left(\frac{\lambda+1}{\lambda-1}\right)^2 P_{\lambda_2+\lambda_{n-2}} \\ &+ \frac{\lambda+1}{\lambda-1}\frac{\lambda+n}{\lambda-n}P_0 + N_{\lambda_1+\lambda_{n-1}}(\lambda), \end{aligned} \tag{5.34}$$

with $N_{\lambda_1+\lambda_{n-1}}(\lambda)$ given by

$$N_{\lambda_1+\lambda_{n-1}}(\lambda) = \frac{1}{2(\lambda-n)(\lambda-1)^2}\begin{pmatrix} -2n-(2+n^2)\lambda+2\lambda^3 & n\sqrt{n^2-4}\lambda \\ n\sqrt{n^2-4}\lambda & 2n-(2+n^2)\lambda+2\lambda^3 \end{pmatrix}. \tag{5.35}$$

For all other simple Lie algebras, the R-matrix for a minimal quantization of the adjoint representation was constructed by Chari and Pressley [11]. This minimal quantization means that the R-matrix is not considered on the adjoint representation, but on the direct sum of the adjoint representation and the trivial representation. The result Eq. (5.34) has a similar structure as their results.

Similar to the case $n = 3$, a more general for the rational R-matrix is obtained upon rescaling $\lambda \to \lambda/\eta$. Furthermore, a second independent solution of the Yang-Baxter equation is given by $\tilde{R}(\lambda) = \sigma I(\lambda)$.

5.6 R-matrices for reducible representations

In the previous sections we have looked at solutions of the YBE on irreducible representations of $\mathfrak{su}(n)$. A possible question is what happens if one considers reducible representations of simple Lie algebras. In this section we will take a look at a particular example, the direct sum of two $\mathbf{\frac{1}{2}}$ representations of $\mathfrak{su}(2)$. The R-matrix on the tensor product on two such representations has been considered before [5]. We will present the solution for the R-matrix on the direct sum representation in terms of the Pauli matrices.

For convenience, we write the representation on each site as $\mathbf{\frac{1}{2}}_A \oplus \mathbf{\frac{1}{2}}_B$, where we have labeled the representations as A and B. The R-matrix will be constructed on the 16-dimensional space $\left(\mathbf{\frac{1}{2}}_A \oplus \mathbf{\frac{1}{2}}_B\right) \otimes \left(\mathbf{\frac{1}{2}}_A \oplus \mathbf{\frac{1}{2}}_B\right)$. This representation decomposes as the direct sum of four copies of $\mathbf{1}$ and four copies of $\mathbf{0}$. Equivalently, one could write this as the direct sum of four copies of $\mathbf{\frac{1}{2}}_i \otimes \mathbf{\frac{1}{2}}_j$ where both i and j can take the values A and B. We label these four modules as 1,2,3 and 4. The R-matrix that we are looking for should be invariant under the action of $\mathfrak{su}(2)$, so it can be written in terms of operators that map complete irreducible submodules to each other. The R-matrix on this representation has a similar structure as the R-matrix we considered in Section 5.3. A solution on the direct sum representation is given by

$$\begin{aligned} R(\lambda) = {} & J_1(\lambda)\left(P_{\mathbf{1}} + \frac{\lambda+\eta}{\lambda-\eta}P_{\mathbf{0}}\right) + J_2(\lambda)\left(M_{12} + M_{34} + \frac{\lambda+\eta}{\lambda-\eta}\left(N_{12}+N_{34}\right)\right) \\ & + J_3(\lambda)\left(M_{13} + M_{24} + \frac{\lambda+\eta}{\lambda-\eta}\left(N_{13}+N_{24}\right)\right) + J_4(\lambda)\left(M_{14} + M_{23} + \frac{\lambda+\eta}{\lambda-\eta}\left(N_{14}+N_{23}\right)\right), \end{aligned} \tag{5.36}$$

where $P_{\mathbf{1}}$ and $P_{\mathbf{0}}$ are projectors on the direct sum of the four $\mathbf{1}$ and $\mathbf{0}$ modules respectively. The operator M_{ij} interchanges the highest weight vectors of the i^{th} and j^{th} $\mathbf{1}$ module while N_{ij} does the same for the $\mathbf{0}$ modules. The functions J_i are all unrestricted. This result can be interpreted as follows. Instead of the usual normalization where one can multiply the R-matrix with an arbitrary function, there are now four arbitrary functions which are undetermined. These four independent solutions all have a similar interpretation. The R-matrix for the irreducible spin-$\frac{1}{2}$ representation that we found in Section 5.3 is defined on $\mathbf{\frac{1}{2}} \otimes \mathbf{\frac{1}{2}}$. The first solution is the direct sum of four of the solutions in the irreducible case. The other three are the same, but they interchange the different copies of $\mathbf{\frac{1}{2}} \otimes \mathbf{\frac{1}{2}}$. Contrary to the R-matrices in the other cases, it is not automatically true that $R(0) = \sigma$. This is the case if and only if $J_1(0) = 1$ and $J_2(0) = J_3(0) = J_4(0) = 0$.

5.7 Summary

In this chapter we have looked at the Yang-Baxter equation and their solutions, known as R-matrices. First, we reviewed a general method to find those R-matrices corresponding to irreducible representation of simple Lie algebras. This method uses the representation theory of the Yangian. We gave two of those well-known solutions for the lowest-dimensional nontrivial irreducible representation of $\mathfrak{su}(2)$. After that, we presented the R-matrix for the adjoint representation of $\mathfrak{su}(3)$ using an intertwining operator for the corresponding family of Yangian representations. We also conjectured an expression for the rational R-matrix invariant under the adjoint representation of $\mathfrak{su}(n)$ $(n > 3)$, which was based on computations we checked up to $n = 7$. Finally, we have given an R-matrix for a reducible representation of $\mathfrak{su}(2)$, the direct sum of two lowest-dimensional non-trivial representations.

Chapter 6

Integrable spin chains

6.1 Introduction

In this chapter we will investigate a subclass of spin chains which are known as integrable. The kind of integrable systems we are interested in are particularly well-behaved because there is, in principle, a method to find all the eigenvalues and eigenstates of the system. This method is known as the Bethe Ansatz, introduced by Bethe in 1931. However, this method is often rather involved and numerical calculations are in most cases necessary to find all the states with their energies. The method is based on the fact that integrable systems have by definition a lot of commuting operators. This set of operators can be used to find the eigenstates of the system, using the well known fact that commuting diagonalizable operators have a common set of eigenstates. The eigenvalues of the Hamiltonian can then be found by acting with the Hamiltonian on the eigenstates.

The main goal of this chapter is to construct an integrable Hamiltonian that is invariant under the action of the adjoint representation of $\mathfrak{su}(3)$. Such a model could describe interesting physics, for instance in relation to the theory of Quantum Chromodynamics, that describes elementary particles. A particular class of these particles, which are called gluons, are vectors in the adjoint representation of $\mathfrak{su}(3)$. Therefore, a spin chain under this representation could be seen as a system that describes gluons. Although gluons have never been seen in chains in nature, it could well be that some of the properties of gluons are described well by such a spin chain. For instance, the Heisenberg chain is not a realistic models of electrons, but it gives a fairly accurate description of (anti-)ferromagnetic behaviour.

In the previous chapter we have seen the Yang-Baxter equation and some of their solutions, which are known as R-matrices. In this chapter we will show how one can construct an integrable Hamiltonian from an R-matrix, using a general procedure which is due to Faddeev [19]. An integrable Hamiltonian is characterized by the fact there is a large number of mutually commuting operators that contains the Hamiltonian itself. We will construct this family of commuting operators using a method which is proposed by Faddeev and is known as the Algebraic Bethe Ansatz. In particular, we will rederive the Hamiltonian for such a system in two simple cases, corresponding to the spin-$\frac{1}{2}$ and spin-1 representation of $\mathfrak{su}(2)$. Then we will review how the integrability of this systems leads to a description of the eigenvalues and eigenstates, involving the so called Bethe equations. In these sections we will follow the approach by Babujian [3], Faddeev [19], and Samaj and Bajnok [41]. Then we will use the same techniques to try to find an integrable spin chain that is invariant under the adjoint representation of $\mathfrak{su}(3)$. Furthermore, we will also give a general result for $\mathfrak{su}(n)$, based on the conjectured R-matrix from the previous chapter. As we will see, the constructed Hamiltonians will not be Hermitian. Therefore, we investigate a special subclass of operators, known as quasi-Hermitian, to check if we could still describe a physical system with our constructed Hamiltonian. In the final section, we look at an integrable Hamiltonian for which the vector space of states on each sites is a reducible representation of $\mathfrak{su}(2)$.

6.2 R-matrices and integrable systems

We start with an abstract construction for the Hamiltonian. Let $\mathfrak{g}$ be a Lie algebra, V a representation and $R(\lambda)$ a rational R-matrix on this representation, invariant under $\mathfrak{g}$. We will also assume that the set of points $\lambda \in \mathbb{C}$ where $R(\lambda)$ is invertible is dense in $\mathbb{C}$. The N-site Hamiltonian will be constructed on the space $W = \bigotimes_{i=1}^{N} V$, with one copy of V for every site. Our system is one-dimensional and we assume periodic boundary conditions, so the neighbours of the N^{th} site are sites $N-1$ and 1. For now we work on the space $V \otimes V \otimes W$, where the extra copies of V are known as the auxiliary spaces. To distinguish these $N+2$ copies of V, we label V_a and V_b for the auxiliary spaces and V_i for the i^{th} copy of V in W.

We define the monodromy matrix $T_a(\lambda)$ for the auxiliary space V_a as

$$T_a(\lambda) = R_{a1}(\lambda) R_{a2}(\lambda) \cdots R_{aN}(\lambda). \tag{6.1}$$

Here, obvious notations are used, so R_{ai} is the R-matrix acting acting nontrivially on the auxiliary space V_a and the i^{th} site, while being the identity on all other sites.

Lemma 6.1. *The monodromy matrix satisfies the following relation, which is known as the fundamental commutation relation,*

$$R_{ab}(\lambda - \mu) T_a(\lambda) T_b(\mu) = T_b(\mu) T_a(\lambda) R_{ab}(\lambda - \mu). \tag{6.2}$$

Proof. We first prove the case $N = 2$. We note that R-matrices that act on different spaces commute, i.e $[R_{ai}(\lambda), R_{bj}(\lambda)] = 0$ if $i \neq j$. Using the YBE, we see that

$$\begin{aligned}
R_{ab}(\lambda - \mu) T_a(\lambda) T_b(\mu) &= R_{ab}(\lambda - \mu) R_{a1}(\lambda) R_{a2}(\lambda) R_{b1}(\mu) R_{b2}(\mu) \\
&= R_{ab}(\lambda - \mu) R_{a1}(\lambda) R_{b1}(\mu) R_{a2}(\lambda) R_{b2}(\mu) \\
&= R_{b1}(\mu) R_{a1}(\lambda) R_{ab}(\lambda - \mu) R_{a2}(\lambda) R_{b2}(\mu) \\
&= R_{b1}(\mu) R_{a1}(\lambda) R_{b2}(\mu) R_{a2}(\lambda) R_{ab}(\lambda - \mu) \\
&= R_{b1}(\mu) R_{b2}(\mu) R_{a1}(\lambda) R_{a2}(\lambda) R_{ab}(\lambda - \mu) \\
&= T_b(\mu) T_a(\lambda) R_{ab}(\lambda - \mu).
\end{aligned}$$

Here, we have used the Yang-Baxter equation twice. The general case follows easily from the $N = 2$ case using the same method and an inductive argument. □

The next stap is to get rid of the auxiliary space, by taking the trace of the monodromy operator in the auxiliary space, i.e.

$$F(\lambda) = \text{tr}_a T_a(\lambda). \tag{6.3}$$

Here, we use the following definition for such a partial trace. If $M_1 \otimes M_2$ is an operator on $W_1 \otimes W_2$, then the trace over for instance W_1 is given by

$$\text{tr}_1 M_1 \otimes M_2 = \text{tr}\,(M_1)\, M_2, \tag{6.4}$$

which gives an operator on the space W_2.

The following lemma gives us a whole family of commuting operators.

Lemma 6.2. *All operators of the family $F(\lambda)$ commute with each other, i.e. we have*

$$[F(\lambda), F(\mu)] = 0. \tag{6.5}$$

Proof. Let $\lambda - \mu \in \{\rho \in \mathbb{C} | R(\rho) \text{ invertible}\}$. We write Eq. (6.2) as

$$T_a(\lambda)T_b(\mu) = (R_{ab}(\lambda-\mu))^{-1} T_b(\mu)T_a(\lambda)R_{ab}(\lambda-\mu).$$

If we take the trace over $V_a \otimes V_b$, we find that

$$\mathrm{tr}_{a,b} T_a(\lambda)T_b(\mu) = \mathrm{tr}_{a,b} T_b(\mu), T_a(\lambda) \tag{6.6}$$

using the cyclicity of the trace and the fact that $R_{ab}(\lambda-\mu)$ is an operator that acts nontrivially only on $V_a \otimes V_b$. Assume for a moment that $T_a(\lambda) = M_a \otimes 1 \otimes M_{N,a}$ and $T_b(\mu) = 1 \otimes M_b \otimes M_{N,b}$, with M_a, M_b operators on V_a and V_b respectively, while $M_{N,a}$ and $M_{N,b}$ both act on $\otimes_{i=1}^N V_i$. Then we can rewrite the left hand side of Eq. (6.6) as

$$\begin{aligned}
\mathrm{tr}_{a,b} T_a(\lambda)T_b(\mu) &= \mathrm{tr}_{a,b} M_a \otimes M_b \otimes M_{N,a}M_{N,b} \\
&= \mathrm{tr}\,(M_a)\,\mathrm{tr}\,(M_b)\,M_{N,a}M_{N,b} \\
&= \mathrm{tr}_a\,(M_a \otimes 1 \otimes M_{N,a})\,\mathrm{tr}_b\,(1 \otimes M_b \otimes M_{N,b}) \\
&= F(\lambda)F(\mu).
\end{aligned}$$

Since the trace is linear and since $T_a(\lambda)$ and $T_b(\mu)$ can be written as a linear combination of pure tensors, this result is valid in general. Doing the same calculation for the right hand side of Eq. (6.6) immediately leads to Eq. (6.5). For $\lambda - \mu \notin \{\rho \in \mathbb{C} | R(\rho) \text{ invertible}\}$, the result follows by analytic continuation. □

We will construct the Hamiltonian from this commuting family of operators. Since we do not want the Hamiltonian to be dependent on the chosen normalization of the R-matrix, the following choice is a natural one

$$H = \left(\frac{d}{d\lambda}\Big|_{\lambda=0} F(\lambda) \right) F(0)^{-1}, \tag{6.7}$$

since a normalization factor on $F(\lambda)$ will drop out. A Hamiltonian that is constructed in this way is integrable. The corresponding commuting operators is the family $F(\lambda)$, together with the Lie algebra operators. In the next section we will see why the above choice for H is the correct one and we will show some examples. In particular, we will see some useful properties of such integrable Hamiltonians that distinguish it from arbitrary systems.

6.3 A spin-$\frac{1}{2}$ integrable spin chain

In this section we will see an example of an integrable Hamiltonian in the case of $\mathfrak{su}(2)$. First, we look at the spin-$\frac{1}{2}$ representation of $\mathfrak{su}(2)$. We recall that we write $\mathbf{s}$ for the spin-s representation, which is the $(2s+1)$-dimensional irreducible representation of $\mathfrak{su}(2)$. The R-matrix that we found in the previous chapter is given by

$$R(\lambda) = P_{\mathbf{1}} + \frac{\lambda+1}{\lambda-1} P_{\mathbf{0}}.$$

Here we have put the parameter $\eta = 1$. This corresponds to replacing the argument of the R-matrix λ to $\frac{\lambda}{\eta}$ and the result in the end will not depend on this rescaling. At $\lambda = 0$, the R-matrix equals the permutation operator. We try to keep the discussion as general as possible and the only thing we will assume about the R-matrix is $R(0) = \sigma$. We first compute the monodromy at $\lambda = 0$,

$$\begin{aligned} T_a(0) &= R_{a1}(0)R_{a2}(0)\cdots R_{aN}(0) \\ &= \sigma_{a1}\sigma_{a2}\cdots\sigma_{aN} \\ &= \sigma_{aN}\sigma_{(N-1)N}\cdots\sigma_{23}\sigma_{12}, \end{aligned} \tag{6.8}$$

where the latter relation of the permutations is easily proven by an argument involving induction. We want to take the trace over auxiliary space. In the case of the spin-$\frac{1}{2}$, there is an easy way to do this. One can express the switch operator in terms of tensor products of the Pauli matrices and the identity matrix and find that $\mathrm{tr}_a\sigma_{Na} = \mathbb{1}_N$. However, we rather use a general argument.

Lemma 6.3. *Let V be a finite-dimensional vector space and σ the permutation operator on $V \otimes V$. Then the trace of σ over the first copy of V is the identity on the second copy, i.e. $tr_1\sigma = \mathbb{1}_2$.*

Proof. We pick a basis $B = \{e_1, e_2, \cdots, e_m\}$ of V. Let A_{ij} be the linear operator on V that sends e_i to e_j while it sends all other basis vectors to 0. Then we have $\mathrm{tr}A_{ij} = \delta_{ij}$. Indeed, if $i = j$, then A_{ij} acts diagonally along B, with eigenvalues 1 with multiplicity 1 and 0 with multiplicity $m-1$, so its trace equals 1. If $i \neq j$, then A_{ij} is nilpotent and therefore traceless. We claim that

$$\sigma = \sum_{i,j=1}^{m} A_{ij} \otimes A_{ji}. \tag{6.9}$$

To prove this, it suffices to calculate the action of the right hand side of Eq. (6.9) on an arbitrary basis element $e_k \otimes e_l$ of $V \otimes V$. Of all m^2 terms in the summation, only $A_{kl} \otimes A_{lk}$ will contribute, since all other terms kill at least one of both factors. Our claim is now proven by the following computation

$$(A_{kl} \otimes A_{lk})(e_k \otimes e_l) = A_{kl}(e_k) \otimes A_{lk}(e_l) = e_l \otimes e_k.$$

The lemma follows immediately from

$$\mathrm{tr}_1\sigma = \mathrm{tr}_1\left(\sum_{i,j=1}^{m} A_{ij} \otimes A_{ji}\right) = \sum_{i,j=1}^{m} \mathrm{tr}\,(A_{ij})\,A_{ji} = \sum_{i,j=1}^{m} \delta_{ij}A_{ji} = \sum_{i=1}^{m} A_{ii} = \mathbb{1},$$

since $\sum_{i=1}^{m} A_{ii}$ sends each of the basis vectors to itself. □

As a result of this we see that $F(0) = \sigma_{(N-1)N}\cdots\sigma_{23}\sigma_{12}$. We can easily compute the inverse of this, since all permutation operators square to the identity. Therefore, we have

$$F(0)^{-1} = \sigma_{12}\sigma_{23}\cdots\sigma_{(N-1)N}. \tag{6.10}$$

We also want to calculate $\frac{d}{d\lambda}\Big|_{\lambda=0} F(\lambda)$. For this, we introduce the notation

$$h_{ai} = \frac{d}{d\lambda}\Big|_{\lambda=0} R_{ai}(\lambda)\sigma_{ai}. \tag{6.11}$$

First we compute the derivative at $\lambda = 0$ of $T(\lambda)$,

$$\frac{d}{d\lambda}\Big|_{\lambda=0} T(\lambda) = \sum_{i=1}^{N} R_{a1}(0)R_{a2}(0)\cdots R_{a(i-1)}(0)R'_{ai}(0)R_{a(i+1)}(0)\cdots R_{aN}(0)$$
$$= \sum_{i=1}^{N} \sigma_{a1}\sigma_{a2}\cdots\sigma_{a(i-1)}h_{ai}\sigma_{ai}\sigma_{a(i+1)}\cdots\sigma_{aN}.$$

To simplify this, we note that $h_{ai}\sigma_{ai}\sigma_{a(i+1)} = \sigma_{ai}\sigma_{a(i+1)}h_{i(i+1)}$, which can be proven by acting with both sides on an arbitrary element. Therefore, we have

$$\frac{d}{d\lambda}\Big|_{\lambda=0} F(\lambda) = \sum_{i=1}^{N} \mathrm{tr}_a\ \sigma_{a1}\sigma_{a2}\cdots\sigma_{ai}\sigma_{a(i+1)}h_{(i+1)i}\sigma_{a(i+2)}\cdots\sigma_{aN}$$
$$= \sum_{i=1}^{N} \mathrm{tr}_a\ \sigma_{a1}\sigma_{a2}\cdots\sigma_{a(i-1)}\sigma_{ai}\sigma_{a(i+1)}\cdots\sigma_{aN}h_{i(i+1)}$$

using the fact that operators acting on two completely different pairs of spaces commute. Extra care is required for the last term, involving h_{aN}. There is no permutation in front, but one can still use that $h_{aN}\sigma_{aN}\sigma_{a1} = \sigma_{aN}\sigma_{a1}h_{N1}$ using the cyclicity of the trace. Computing the trace over the auxiliary space leads to

$$\frac{d}{d\lambda}\Big|_{\lambda=0} F(\lambda) = \sum_{i=1}^{N} \sigma_{(N-1)N}\cdots\sigma_{12}h_{i(i+1)} = F(0)\sum_{i=1}^{N} h_{i(i+1)}. \tag{6.12}$$

The Hamiltonian can be obtained by multiplying Eqs. (6.12) and (6.10). We find that

$$H = \sum_{i=1}^{N} h_{i(i+1)}. \tag{6.13}$$

We now see why Eq. (6.7) is the right construction for the Hamiltonian. In principle, any operator constructed from the family $F(\lambda)$ would be commuting with all the $F(\lambda)$, but this particular construction leads to a Hamiltonian which consists only of nearest-neighbour interactions, which makes it a physically viable model. We have barely used the fact that we are using the spin-$\frac{1}{2}$ R-matrix, except for the fact that $R(0) = \sigma$. In fact, this construction often works in general, since this is a weak property of an R-matrix as we noted in the previous chapter. Therefore, Eqs. (6.11) and (6.13) can be used to generate a lot of integrable Hamiltonians. We will see examples of this construction later on and also the problems that can arise. For now, we will give an expression for $h_{i(i+1)}$ in terms of the spin operators on the site. Note that from Eq. (5.3) we have that

$$h_{i(i+1)} = 2P_{\mathbf{0},i(i+1)}. \tag{6.14}$$

The sign comes from the fact that the permutation operator acts by minus the identity on $\mathbf{0}$. The Casimir element C of the $\mathbf{\frac{1}{2}} \otimes \mathbf{\frac{1}{2}}$ representation of $\mathfrak{su}(2)$ is given by

$$C = \sum_{a=1}^{3} S^a \otimes S^a,$$

with $S^a = \frac{1}{2}\sigma^a$ the spin operators. One can easily compute the scalars by which C acts on the modules $\mathbf{1}$ and $\mathbf{0}$ in the decomposition $\mathbf{\frac{1}{2}} \otimes \mathbf{\frac{1}{2}} = \mathbf{1} \oplus \mathbf{0}$. These are $\frac{1}{4}$ and $-\frac{3}{4}$ respectively. From these considerations, it is easily seen that

$$P_{\mathbf{0}} = -C + \frac{1}{4}. \tag{6.15}$$

Using Eqs. (6.13), (6.14) and (6.15), we find that

$$H = -\frac{1}{2}\sum_{i=1}^{N}\left(4\,\mathbf{S}_i \cdot \mathbf{S}_{i+1} - 1\right). \tag{6.16}$$

Here we have omitted the tensor product between the Pauli matrices, as is common in physics notation. Up to two constants, this is precisely the Heisenberg chain. The precise values for these constants are irrelevant, since any multiple of this Hamiltonian or the identity still commutes with the complete family $F(\lambda)$.

6.4 A spin-1 integrable spin chain

In the previous section, we have seen the construction of an integrable system. In the end, we precisely encovered the well-known Heisenberg spin chain. In this section, we use the same method for the spin-1 representation of $\mathfrak{su}(2)$. In the end, we will find an integrable spin-1 chain. We note that this spin chain can not be obtained by heuristically replacing the spin-$\frac{1}{2}$ operators in Eq. (6.16) by their spin-1 counterparts. We have computed the rational R-matrix in the previous chapter as

$$R(\lambda) = P_{\mathbf{2}} + f(\lambda)P_{\mathbf{1}} + g(\lambda)P_{\mathbf{0}},$$

with the functions f and g equal to

$$f(\lambda) = \frac{\lambda+2}{\lambda-2},\; g(\lambda) = \frac{(\lambda+1)(\lambda+2)}{(\lambda-1)(\lambda-2)},$$

where we again have put the parameter $\eta = 1$. We note again that

$$R(0) = P_{\mathbf{2}} - P_{\mathbf{1}} + P_{\mathbf{0}} = \sigma.$$

Therefore, the derivation of the previous section is applicable. We immediately find that the local Hamiltonian is given by

$$h_{i(i+1)} = R'(0)_{i(i+1)}\sigma_{i(i+1)} = P_{\mathbf{1},i(i+1)} + 3P_{\mathbf{0},i(i+1)} \tag{6.17}$$

using that $\mathbf{1}$ is an antisymmetric submodule and $\mathbf{0}$ is a symmetric submodule. The Casimir operator $C = \sum_{a=1}^{3} S^a \otimes S^a$, with S_i the spin-1 operators, acts by the scalars 1, -1 and -2 on the modules $\mathbf{2}$, $\mathbf{1}$ and $\mathbf{0}$ respectively. Therefore, we have

$$P_{\mathbf{1}} = -\frac{1}{2}(C+2)(C-1),$$
$$P_{\mathbf{0}} = \frac{1}{3}(C+1)(C-1),$$

and therefore the Hamiltonian becomes

$$H = -\frac{1}{2}\sum_{i=1}^{N}\left[\mathbf{S}_i \cdot \mathbf{S}_{i+1} - (\mathbf{S}_i \cdot \mathbf{S}_{i+1})^2\right]. \tag{6.18}$$

This is known as the Babujian-Takhtajan model as we already mentioned in Section 4.5. For higher-dimensional representations, one can also construct an integrable model [3, 47]. For spin s, the Hamiltonian is given by a polynomial in the Casimir element of degree $2s$. This is a consequence of the decomposition $\mathbf{s} \otimes \mathbf{s}$ which consists of $2s+1$ irreducible modules, on which the Casimir elements acts by different scalars. Therefore, each of these projections onto these modules is given by a polynomial of degree $2s$ in terms of the Casimir element. Since the Hamiltonian is constructed out of these projections it is given by such a polynomial as well.

The question that naturally arises is how the Hamiltonian depends on the normalization of the R-matrix. For the spin-$\frac{1}{2}$ case, this is immediately clear. The constructed Hamiltonian will always be a linear combination of the identity operator and the Casimir element and the system can only differ by a multiplicative and an additive constant. For the case of spin-1, the system is truely different if the relative prefactor between the Casimir element and the square of it in Eq. (6.18) changes. Suppose we multiply the R-matrix by a function $h(\lambda)$, such that $h(0) = 1$. Then the new R-matrix,

$$\tilde{R}(\lambda) = h(\lambda)R(\lambda) \tag{6.19}$$

satisfies $\tilde{R}(0) = \sigma$. The new Hamiltonian is constructed from the quantity

$$\frac{d}{d\lambda}\Big|_{\lambda=0}\tilde{R}(\lambda)\sigma = \frac{d}{d\lambda}\Big|_{\lambda=0}R(\lambda)\sigma + h'(0)\mathbb{1}, \tag{6.20}$$

which only differs by an additive constant. Therefore, the two Hamiltonians will give rise to the same physical system, which is of course the desired result. If the R-matrix is multiplied by a function such that $h(0) \neq 1$, then the derivation in the previous section is not immediately applicable, since the new R-matrix would not equal the permutation operator at $\lambda = 0$.

Finally, we want to make two remarks on the $\mathfrak{su}(2)$ integrable Hamiltonians. All the spin operators are Hermitian operators and therefore the Casimir operator as being the sum of squares of Hermitian operators is Hermitian. Since the Hamiltonian is a polynomial in the Casimir operator with real coefficients [3], the constructed Hamiltonian will always be Hermitian. From a physical point of view, this is desirable since it ensures that the energies of the system, given by the eigenvalues of the Hamiltonian, are real numbers. Furthermore, the Casimir element commutes with all the $\mathfrak{su}(2)$ operators. Therefore, all components of spin are conserved quantities for these systems.

6.5 The Bethe equations

In this section we will stick to the examples of the previous sections with the $\mathfrak{su}(2)$-invariant spin chains. In particular, we will use the integrability of the systems to construct all the eigenvectors. Furthermore, we will derive the so called Bethe equations from which the eigenvalues can be calculated in principle. The method that we review here is due to Faddeev [19].

First, we investigate the simplest model, which is the spin-$\frac{1}{2}$ model. We use the notation V for the representation space $\mathbf{\frac{1}{2}}$ and write the monodromy operator $T_a(\lambda)$ as a (2×2)-matrix in the auxiliary space V_a,

$$T_a(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix}. \tag{6.21}$$

Here, $A(\lambda)$ $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ are operators on $\bigotimes_{i=1}^{N} V_i$. In this form, it is clear that taking the trace over the auxiliary space V_a, gives

$$F(\lambda) = A(\lambda) + D(\lambda). \tag{6.22}$$

Since the Hamiltonian is constructed from the commuting family $F(\lambda)$, they share a set of eigenvectors. For this reason, we are particularly interested in the eigenvectors of $F(\lambda)$. To construct those, we will combine Eqs. (6.2) and (6.21) to get a set of relations between the operators $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$. For convenience, we multiply the R-matrix with $1-\lambda$, to get rid of the poles in $R(\lambda)$. This change does not lead to different relations between the operators, since Eq. (6.2) is invariant under multiplication of $R(\lambda)$ with a scalar function. The product of the two monodromy matrices in two different auxiliary spaces takes the following form.

$$T_a(\lambda)T_b(\mu) = \begin{pmatrix} A(\lambda)A(\mu) & A(\lambda)B(\mu) & B(\lambda)A(\mu) & B(\lambda)B(\mu) \\ A(\lambda)C(\mu) & A(\lambda)D(\mu) & B(\lambda)C(\mu) & B(\lambda)D(\mu) \\ C(\lambda)A(\mu) & C(\lambda)B(\mu) & D(\lambda)A(\mu) & D(\lambda)B(\mu) \\ C(\lambda)C(\mu) & C(\lambda)D(\mu) & D(\lambda)C(\mu) & D(\lambda)D(\mu) \end{pmatrix}. \tag{6.23}$$

An expression for $T_b(\mu)T_a(\lambda)$ can be obtained similarly. Here, we would like to stress that $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ are all operators and do not commute in general. In fact, Eq. (6.2) gives relations between the operators in different order. For instance the $(1,1)$ matrix element of Eq. (6.2) reads

$$A(\lambda)A(\mu) = A(\mu)A(\lambda),$$

which means that the family of operators $A(\lambda)$ is commuting. This is also the case for the families $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$. The $(1,3)$ entry reads

$$(\lambda-\mu-1)B(\lambda)A(\mu) = -B(\mu)A(\lambda) + (\lambda-\mu)A(\mu)B(\lambda), \tag{6.24}$$

which gives a relation between the families of operators $A(\lambda)$ and $B(\lambda)$. In particular, it tells us that if we interchange the operators $A(\lambda)$ and $B(\mu)$ we get an extra term and some multiplicative constants. An other relation that we need later on is the $(3,4)$ entry, which is

$$-B(\lambda)D(\mu) + (\lambda-\mu)D(\lambda)B(\mu) = (\lambda-\mu-1)B(\mu)D(\lambda). \tag{6.25}$$

There are also a few other relations, but those are not necessary to construct the eigenvectors of $F(\lambda)$. For the eigenvectors, we first need an expression for the R-matrix in terms of the spin operators and the identity. We have

$$P_{\mathbf{0}} = \frac{1}{4}\mathbb{1}\otimes\mathbb{1} - \frac{1}{4}\sigma^a\otimes S^a, \tag{6.26}$$

$$P_{\mathbf{1}} = \frac{3}{4}\ \mathbb{1}\otimes\mathbb{1} + \frac{1}{4}\sigma^a\otimes S^a, \tag{6.27}$$

where $S^a = \frac{1}{2}\sigma^a$. Therefore, if we write the R-matrix from Eq. (5.3) as a (2×2)-matrix in auxiliary space, we have

$$R(\lambda) = \begin{pmatrix} \frac{1}{2}-\lambda+S^3 & S^1-\mathrm{i}S^2 \\ S^1+\mathrm{i}S^2 & \frac{1}{2}-\lambda-S^3 \end{pmatrix} \tag{6.28}$$

where we recall that we have multiplied the R-matrix by a factor of $1-\lambda$. We consider the highest weight vector ω of the $\frac{\mathbf{1}}{\mathbf{2}}$-representation. The spin operators acts as follows on ω,

$$\left(S^1+\mathrm{i}S^2\right)\omega=0,\ S^3\omega=\frac{1}{2}\omega.$$

Therefore, applying the R-matrix to ω (or to be precise, the two-dimensional vector with ω as both entries) results in

$$R(\lambda)\omega=\begin{pmatrix}(1-\lambda)\,\omega & \xi\\ 0 & -\lambda\omega\end{pmatrix},$$

where ξ is another two-dimensional vector whose precise form is irrelevant. We consider the vector $\Omega=\omega\otimes\cdots\otimes\omega$ in the space $V^{\otimes N}$. If we let the monodromy matrix from Eq. (6.21) act on Ω, we get

$$T_a(\lambda)=\begin{pmatrix}(1-\lambda)^N\,\Omega & \Xi\\ 0 & (-\lambda)^N\,\Omega\end{pmatrix}, \tag{6.29}$$

where Ξ is now a vector in $V^{\otimes N}$. After taking the trace of this relation over the auxiliary space, we find that

$$F(\lambda)\Omega=\left((1-\lambda)^N+(-\lambda)^N\right)\Omega.$$

We see that Ω is an eigenvector of $F(\lambda)$. To find for extra eigenvectors, we assume them to be of the form

$$\Phi(\{\lambda\})=B(\lambda_1)B(\lambda_2)\cdots B(\lambda_m)\Omega, \tag{6.30}$$

where $\{\lambda\}=\{\lambda_1,\cdots,\lambda_m\}$ is a set of complex numbers. We want to know the action of $F(\lambda)=A(\lambda)+D(\lambda)$ on this vector. The strategy is to pull the $A(\lambda)$ through the m $B(\lambda_j)$ operators, using the relation in Eq. (6.24). First we rewrite it in the the form

$$A(\lambda)B(\mu)=f_1(\lambda-\mu)B(\mu)A(\lambda)+f_2(\lambda-\mu)B(\lambda)A(\mu). \tag{6.31}$$

with

$$f_1(\lambda)=\frac{\lambda+1}{\lambda},\ f_2(\lambda)=-\frac{1}{\lambda}.$$

If we let $A(\lambda)$ act on Eq. (6.30) the result should be of the form

$$\begin{aligned}A(\lambda)\Phi(\{\lambda\})\Omega=&\left(\prod_{j=1}^{m}f_1(\lambda-\lambda_j)\right)\alpha(\lambda)^N\Phi(\{\lambda\})\\&+\sum_{i=1}^{N}M_k(\lambda,\{\lambda\})B(\lambda_1)\cdots\hat{B}(\lambda_k)\cdots B(\lambda_m)B(\lambda)\Omega.\end{aligned} \tag{6.32}$$

Here, the hat on the operator means that is it omitted. The function α is given by $\alpha(\lambda)=1-\lambda$ and the coefficients M_k depend on λ and the set $\{\lambda\}=\{\lambda_1,\cdots,\lambda_m\}$. To see why Eq. (6.32) holds, one should use Eq. (6.31) m times. Each time, that one interchanges $A(\lambda)$ and $B(\lambda_j)$, it gives rise to a factor of $f_1(\lambda-\lambda_j)$ and, after one has pulled $A(\lambda)$ through all the $B(\lambda_i)$, $A(\lambda)$ will act by the scalar $\alpha(\lambda)^N$ on Ω. Furthermore, when $A(\lambda)$ and $B(\lambda_j)$, there is also an extra term that involves $B(\lambda)$ and $A(\lambda_j)$, giving rise to extra terms. Using again Eq. (6.31), one can pull the $A(\lambda_j)$ to the right and let it act on Ω. In any case, such terms are always contained in the second expression on the right hand side of Eq. (6.32). The coefficients $M_k(\lambda,\{\lambda\})$ can of course be very involved, since there are a lot of possibilities for such any of those terms to arise. However, there is a trick to compute them [19]. The coefficient $M_1(\lambda,\{\lambda\})$ can be computed rather easily, since there is only a single way that the term $B(\lambda_2)\cdots B(\lambda_m)B(\lambda)$ can arise. Therefore $M_1(\lambda,\{\lambda\})$ is given by

$$M_1(\lambda,\{\lambda\}) = f_2(\lambda-\lambda_1)\left(\prod_{j=2}^{m} f_1\left(\lambda_1-\lambda_j\right)\right)\alpha(\lambda_1)^N.$$

Since all the $B(\lambda_j)$ commute, the ordering of those operators is irrelevant. So all the other coefficients $M_k(\lambda,\{\lambda\})$ can be obtained by the substitution $\lambda_1 \to \lambda_k$ and therefore we find

$$M_k(\lambda,\{\lambda\}) = f_2(\lambda-\lambda_k)\left(\prod_{\substack{j=1\\ j\neq k}}^{m} f_1\left(\lambda_k-\lambda_j\right)\right)\alpha(\lambda_k)^N.$$

One can compute similar expression for the action of $D(\lambda)$ on $\Phi(\{\lambda\})$, resulting in

$$\begin{aligned} D(\lambda)\Phi(\{\lambda\})\Omega &= \left(\prod_{j=1}^{m} g_1(\lambda-\lambda_j)\right)\delta(\lambda)^N\Phi(\{\lambda\}) \\ &\quad + \sum_{i=1}^{N} N_k(\lambda,\{\lambda\})B(\lambda_1)\cdots\hat{B}(\lambda_k)\cdots B(\lambda_m)B(\lambda)\Omega, \end{aligned} \tag{6.33}$$

where $\delta(\lambda) = -\lambda$ and $g_1(\lambda) = (\lambda-1)/\lambda$. The coefficients $N_k(\lambda,\{\lambda\})$ are given by

$$N_k(\lambda,\{\lambda\}) = g_2(\lambda-\lambda_k)\left(\prod_{\substack{j=1\\ j\neq k}}^{m} g_1\left(\lambda_k-\lambda_j\right)\right)\delta(\lambda_k)^N. \tag{6.34}$$

where $g_2(\lambda) = 1/\lambda$. Therefore, we have found an explicit expression for the action of $F(\lambda)$ on the vector $\Phi(\{\lambda\})$. In general, we see that $\Phi(\{\lambda\})$ is not an eigenvector of $F(\lambda)$ as a result of the terms proportional $M_k(\lambda,\{\lambda\})$ and $N_k(\lambda,\{\lambda\})$. Those terms are called the unwanted terms for that reason. The goal is now to chose a particular condition on the elements of the set $\{\lambda\}$ such that those unwanted terms vanish. We note that $f_2(\lambda) = -g_2(\lambda)$. As a result, the unwanted terms vanish for all λ if and only if

$$\left(\prod_{\substack{j=1\\ j\neq k}}^{m} f_1\left(\lambda_k-\lambda_j\right)\right)\alpha(\lambda_k)^N = \left(\prod_{\substack{j=1\\ j\neq k}}^{m} g_1\left(\lambda_k-\lambda_j\right)\right)\delta(\lambda_k)^N$$

which can be rewritten as

$$\prod_{\substack{j=1\\ j\neq k}}^{m} \frac{\lambda_k-\lambda_j+1}{\lambda_k-\lambda_j-1} = \left(\frac{\lambda_k}{\lambda_k-1}\right)^N. \tag{6.35}$$

These relations are known as the Bethe equations. The eigenvalue $\Lambda(\lambda, \{\lambda\})$ of $F(\lambda)$ corresponding to the eigenvector $\Phi(\{\lambda\})$ can be found immediately from Eqs. (6.32) and (6.33) and equals

$$\Lambda(\lambda, \{\lambda\}) = \left(\prod_{j=1}^{m} f_1(\lambda - \lambda_j)\right) \alpha(\lambda)^N + \left(\prod_{j=1}^{m} g_1(\lambda - \lambda_j)\right) \delta(\lambda)^N \tag{6.36}$$

provided the conditions in Eq. (6.35) hold. To find the eigenvalues of the Hamiltonian in Eq. (6.16) we have to take the logarithmic derivative of Eq. (6.36) taking the Bethe equations into account. First, we have to reinstate the factor $1-\lambda$ that we divided out in the R-matrix. Then the eigenvalues for the transfer matrix become

$$\tilde{\Lambda}(\lambda, \{\lambda\}) = \left(\prod_{j=1}^{m} f_1(\lambda - \lambda_j)\right) + \left(\prod_{j=1}^{m} g_1(\lambda - \lambda_j)\right) \frac{\delta(\lambda)^N}{\alpha(\lambda)^N}. \tag{6.37}$$

As the second term in Eq. (6.37) as well as its derivative vanish for $\lambda = 0$, that term does not contribute to the eigenvalues of the Hamiltonian. The eigenvalues of the Hamiltonian $H = F(0)^{-1}F'(0)$ can be computed as

$$Z(\{\lambda\}) = \sum_{j=1}^{m} \frac{1}{\lambda_j(1-\lambda_j)}. \tag{6.38}$$

So, in principle we have found some of the eigenstates and the corresponding eigenvalues of the Hamiltonian. It turns out that $\Phi(\{\lambda\})$ are all highest weight states with total spin $N/2 - m$. A counting argument that we do not present here shows that the $\Phi(\{\lambda\})$ form all the independent highest weight states [19]. To show that $\Phi(\{\lambda\})$ are highest weight states, we first note that the monodromy matrix from Eq. (6.21) is a polynomial in λ of which the first terms of highest order are given by

$$T(\lambda) = (-1)^N \lambda^N + \left(\frac{N}{2} + \sigma^a \otimes S^a\right) \lambda^{N-1}$$

with σ^a the Pauli matrices and S^a the total spin operator on the spin chain. Matching the coefficients in the fundamental commutation relation, now leads to the following relation

$$\left[T(\lambda), \frac{1}{2}\sigma^a \otimes \mathbb{1} + \mathbb{1} \otimes S^a\right] = 0.$$

Plugging in the matrix form for the monodromy matrix gives us

$$[S^3, B(\lambda)] = 0$$
$$[S^1 + \mathrm{i}S^2, B(\lambda)] = A(\lambda) - D(\lambda)$$

From this it follows that

$$S^3\Phi(\{\lambda\}) = \left(\frac{N}{2} - m\right)\Phi(\{\lambda\})$$
$$(S^1 + \mathrm{i}S^2)\Phi(\{\lambda\}) = 0$$

where we note that to show the latter relation the Bethe equation are needed. We see that $\Phi(\{\lambda\})$ is indeed a highest weight vector of the correct weight. Since the spin chain is $\mathfrak{su}(2)$-invariant, the lowering operator $S^1 - \mathrm{i}S^2$ can be used to find all the other eigenstates of the Hamiltonian.

In principle, this method can also be used for the integrable models that are invariant under higher spin representations. However, since the dimension of the auxiliary space is also larger in such a case, the monodromy operator, when written as a matrix in auxiliary operator as in Eq. (6.21), will contain more than four different operators. As a result, the relations resulting from the fundamental commutation relation in Eq. (6.2) will be a lot more involved. There is a way to circumvent this problem, involving a so called Lax operator, which connects the R-matrix of the spin-$\frac{1}{2}$ representation with R-matrices for other representations and is defined on $\mathbf{\frac{1}{2}} \otimes \mathbf{s}$. We will denote the R-matrix of the spin-s representation by ${}_sR(\lambda)$. The Lax operator is denoted by ${}_{\frac{1}{2}s}R(\lambda)$ and satisfies the following two equations

$$
\begin{aligned}
{}_{\frac{1}{2}s}R_{12}(\lambda-\mu)\ {}_sR_{13}(\lambda)\ {}_{\frac{1}{2}s}R_{23}(\mu) &= {}_{\frac{1}{2}s}R_{23}(\mu)\ {}_sR_{13}(\lambda)\ {}_{\frac{1}{2}s}R_{12}(\lambda-\mu) && (6.39)\\
{}_{\frac{1}{2}s}R_{12}(\lambda-\mu)\ {}_{\frac{1}{2}}R_{13}(\lambda)\ {}_{\frac{1}{2}s}R_{23}(\mu) &= {}_{\frac{1}{2}s}R_{23}(\mu)\ {}_{\frac{1}{2}}R_{13}(\lambda)\ {}_{\frac{1}{2}s}R_{12}(\lambda-\mu) && (6.40)
\end{aligned}
$$

The first of these equations is an operator equation on the vector space $\mathbf{s}\otimes\mathbf{\frac{1}{2}}\otimes\mathbf{s}$ and the second on $\mathbf{\frac{1}{2}}\otimes\mathbf{s}\otimes\mathbf{\frac{1}{2}}$. The existence of such an operator ${}_{\frac{1}{2}s}R(\lambda)$ can be derived from the universal R-matrix of $Y(\mathfrak{su}(2))$. The proof is the same as before. One extends the representation $\mathbf{s}\otimes\mathbf{\frac{1}{2}}\otimes\mathbf{s}$ to a representation of the Yangian using the evaluation homomorphism in Eq. (3.7) and applies this representation to the fundamental Yang-Baxter equation in the Yangian, which is solved by the fundamental R-matrix. This leads to Eq. (6.39). The result in Eq. (6.40) is obtained similarly. The Lax operator ${}_{\frac{1}{2}s}R(\lambda)$ is the same for both equations, since it is the extended representation of $\mathbf{\frac{1}{2}}\otimes\mathbf{s}$ applied to the universal R-matrix. An explicit expression for the Lax operator is given by

$$
{}_{\frac{1}{2}s}R(\lambda) = \left(\frac{1}{2}-\lambda\right)\mathbb{1}_{\mathbf{\frac{1}{2}}}\otimes\mathbb{1}_{\mathbf{s}} + \sigma^a\otimes S^a. \tag{6.41}
$$

In case $\mathbf{s}=\mathbf{\frac{1}{2}}$, then this is the R-matrix we found before. The operators S^a are the spin operators in the representation $\mathbf{s}$.

Similar to the case of spin-$\frac{1}{2}$ the monodromy matrix is defined by

$$
{}_sT_a(\lambda) = {}_sR_{aN}(\lambda)\ {}_sR_{a(N-1)}(\lambda)\cdots\ {}_sR_{a1}(\lambda), \tag{6.42}
$$

where the auxiliary space is now a copy of the representation $\mathbf{s}$. The derivation of the fundamental commutation relation in Eq. (6.2) is the same. We also define a monodromy operator for the Lax operator in the space $\mathbf{\frac{1}{2}}\otimes\mathbf{s}^{\otimes N}$,

$$
{}_{\frac{1}{2}s}T_a(\lambda) = {}_{\frac{1}{2}s}R_{aN}(\lambda)\ {}_{\frac{1}{2}s}R_{a(N-1)}(\lambda)\cdots{}_{\frac{1}{2}s}R_{a1}(\lambda). \tag{6.43}
$$

We proceed by taking the trace over the auxiliary space for both monodromy operators.

$$
{}_{\frac{1}{2}s}F(\lambda) = \mathrm{tr}_a\ {}_{\frac{1}{2}s}T_a(\lambda),\ \ {}_sF(\lambda) = \mathrm{tr}_a\ {}_sT_a(\lambda). \tag{6.44}
$$

These are both operators on $\mathbf{s}^{\otimes N}$. The point is now to show that these families of operators commute as a consequence of Eqs. (6.39) and (6.40). To show this, we note the following equation

$$
{}_{\frac{1}{2}}R_{ab}(\lambda-\mu)\ {}_{\frac{1}{2}s}T_a(\lambda)\ {}_sT_b(\mu) = {}_sT_b(\mu)\ {}_{\frac{1}{2}s}T_a(\lambda)\ {}_{\frac{1}{2}}R_{ab}(\lambda-\mu), \tag{6.45}
$$

which proof is very similar to the one for Eq. (6.2). Taking the trace over the auxiliary spaces leads to the conclusion that ${}_{\frac{1}{2}s}F(\lambda)$ and ${}_sF(\mu)$ that commute. Therefore, these operators have a common set of eigenvectors. The eigenvectors of ${}_{\frac{1}{2}s}F(\lambda)$ can be constructed using the same method as for the spin-$\frac{1}{2}$ case. The first step is to write the monodromy operator ${}_{\frac{1}{2}s}T(\lambda)$ in the form of Eq. (6.21). The Lax operator as a matrix in auxiliary space is the same as Eq. (6.28), although the spin operators are now spin-s operators. These operators act slightly differently on the highest weight vector,

$$\left(S^1 + \mathrm{i}S^2\right)\omega = 0, \; S^3\omega = s\omega.$$

The rest of the derivation is now exactly the same. The eigenvectors of ${}_{\frac{1}{2}s}F(\lambda)$ are again of the form as in Eq. (6.30). The Bethe equations take a similar form and now read

$$\prod_{\substack{j=1 \\ j\neq k}}^{m} \frac{\lambda_k - \lambda_j + 1}{\lambda_k - \lambda_j - 1} = \left(\frac{\lambda_k - \frac{1}{2} + s}{\lambda_k - \frac{1}{2} - s}\right)^N. \tag{6.46}$$

Ultimately, the eigenvectors and eigenvalues of ${}_sF(\lambda)$ can be found by applying this operator to the vectors of the form $\Phi(\{\lambda\})$. Some unwanted terms arise, but when the Bethe equations in Eq. (6.46) hold, those terms vanish. The eigenvalues of ${}_sF(\lambda)$ are complicated and a simple form does not exist [3].

In summary, in this section we have seen the construction of the eigenvectors and eigenvalues of $\mathfrak{su}(2)$ integrable Hamiltonians. This is precisely why integrable systems are particularly interesting, since there is a way that in principle solves the system. Later, the same method will be used for the R-matrix for the adjoint representation of $\mathfrak{su}(3)$, to derive an integrable Hamiltonian for that representation.

6.6 An integrable Hamiltonian for the adjoint representation of SU(3)

In the previous chapter we have constructed the R-matrix for the adjoint representation of $\mathfrak{su}(3)$. Similarly to what we have done before for the fundamental and the adjoint representation of $\mathfrak{su}(2)$, we would also like to construct an integrable model using this R-matrix. However, this R-matrix in (5.15) is not solely given in terms of projectors, but also contains two terms that swap two highest weight vectors, which were not present in the $\mathfrak{su}(2)$ case. It is possible to choose different highest weight vectors for the two **8** submodules, such that the R-matrix is given only in terms of projectors. However, this would be unnatural for two reasons. On one hand the highest weight vectors would be dependent on λ and on the other hand the resulting highest weight vector would not be orthogonal with respect to the invariant inner product on $\mathfrak{su}(3) \otimes \mathfrak{su}(3)$. Therefore, we keep the R-matrix in this form.

In the derivation of the $\mathfrak{su}(2)$ integrable models, an important property of the R-matrix was used, namely that $R(0) = \sigma$. This is also the case for the R-matrix in (5.15). Therefore, we can straightforwardly compute the integrable Hamiltonian, which is of the form $H = \sum_i h_{i(i+1)}$. Using the formula

$$h = R'(0)\sigma, \tag{6.47}$$

which was derived in Section 6.3, we find that

$$h = 2P_{\mathbf{10}} + 2P_{\bar{\mathbf{10}}} + \frac{25}{6}P_{\mathbf{8}_s} - \frac{\sqrt{5}}{2}M_{as} + \frac{\sqrt{5}}{2}M_{sa} + \frac{1}{2}P_{\mathbf{8}_a} + \frac{8}{3}P_{\mathbf{1}}. \tag{6.48}$$

We would like to express this Hamiltonian in terms of tensor products of the Gell-Mann matrices. However, this is more complicated than in the case of the $\mathfrak{su}(2)$-invariant Hamiltonians. The projectors $P_{\mathbf{10}} + P_{\overline{\mathbf{10}}}$ and $P_{\mathbf{1}}$ can easily expressed as polynomials in the Casimir operator and the identity. However, the Casimir operator does not distinguish the two copies of $\mathbf{8}$ and therefore different operators are needed, which have been found by Roy and Quella [40]. Using these, the local Hamiltonian can be written as

$$h = 2 + \frac{3}{8}C - \frac{7}{24}C^2 - \frac{1}{12}C^3 - \frac{11}{216}K + \frac{1}{864}[K, Q], \tag{6.49}$$

where the operators C,K and Q can be expressed in the set of operators J^a, the Lie algebra generators of the adjoint representation. These operators can be obtained by fixing a basis I^a for $\mathfrak{su}(3)$ that is orthonormal with respect to the unique inner product that extends the trace form on $\mathfrak{su}(n, \mathbb{R})$. The operators J^a are defined by $J^a = \text{ad } I^a$, where ad is the Lie algebra homomorphism for the adjoint representation. The operators J^a are Hermitian with respect to the unique inner product that makes the adjoint representation of $SU(3)$ unitary. The explicit expressions for the operators C, K and Q are given by

$$C_{1,2} = J_1^a J_2^a, \; K_{1,2} = d^{abc} d^{def} J_1^a J_1^d J_1^e J_2^f J_2^b J_2^c, \; Q_{1,2} = d^{abc}\left(J_1^a J_1^b J_2^c - J_1^a J_1^b J_2^c\right).$$

We recall that the d-symbols are defined by $d^{abc} = \text{tr}\left(I^a\{I^b, I^c\}\right)$.

However, there is a big problem with the local Hamiltonian h from Eq. 6.49. It is not Hermitian, due to the term proportional to $[K, Q]$. To see this, one should note that the operators K and Q are both Hermitian, due to the symmetry of the d-symbols and the fact that the operators J^a are Hermitian. Therefore, the operator $[K, Q]$ is anti-Hermitian. Since we would like the Hamiltonian to describe a physical system, this is a major problem. The eigenvalues of the Hamiltonian correspond to the energy levels of the system and these should of course be real numbers. In Section 6.8 we will see if there is a solution of this problem, by investigating if the Hamiltonian is quasi-Hermitian. This is a weaker property than being Hermitian, but it still guarantees that the eigenvalues of the operator are real.

Ignoring this problem for now, we would like to diagonalize the object $H = \sum_i h_{i(i+1)}$ using the techniques we used for the $\mathfrak{su}(2)$ models. In particular, we would like to construct the Bethe equations. We would need to find the Lax operator such that both Eqs. (6.39) and (6.40) hold with the $\mathfrak{su}(3)$ R-matrices for the fundamental and the adjoint representation. The existence of this Lax operator follows from the universal R-matrix that also exists for the Yangian of $\mathfrak{su}(3)$. However, we have not been able to find the Lax operator. It cannot be constructed from the $\mathfrak{su}(3)$-invariant quantities $I^a \otimes J^a$, $(I^a \otimes J^a)^2$ and the identity. The operator $I^a \otimes J^a$ has three different distinct eigenvalues, corresponding to the three irreducible submodules in the decomposition of $\mathbf{3} \otimes \mathbf{8}$. Therefore, higher powers of the operator $I^a \otimes J^a$ can be expressed in terms of the above operators. It follows from Schur's lemma that the vector space of Lie algebra intertwining operators is three-dimensional in the case of $\mathbf{3} \otimes \mathbf{8}$. We conclude that the Lax operator is not intertwining for the Lie algebra representation, contrary to the case of $\mathfrak{su}(2)$.

Since there is no Lax operator available, it means we have to try a different method to find the eigenvalues Hamiltonian. One possibility would be to write the monodromy operator as a matrix in auxiliary space and apply the method we used in Section 6.5. However this monodromy operator would consist out of 64 different operators and the fundamental commutation relation in Eq. (6.2) leads to complicated relations between these operators. The Bethe equations, even in the case of $m = 1$ for the first excited states, will be very involved as well, although in principle one could find an expression for them.

6.7 An integrable Hamiltonian for the adjoint representation of SU(n)

In this section we will present an integrable Hamiltonian for the adjoint representation of $\mathfrak{su}(n)$, for $n > 3$, based on the conjectured result for the R-matrix in the previous chapter. This Hamiltonian can be obtained following the same considerations as in the previous section was done for $n = 3$. The expression one finds for the local Hamiltonian is

$$h = 2 + \frac{6+n^2}{8}C - \frac{n^2-2}{8n}C^2 - \frac{3}{8}C^3 - \frac{1}{8n}C^4 - \frac{2+n^2}{8n^3}K + \frac{1}{16n^3}\left[K, Q\right]. \tag{6.50}$$

The corresponding Hamiltonian is again not Hermitian, due to the anti-Hermitian term containing the commutator. We note that this result differs from the previous local Hamiltonian in the case $n = 3$, due to the appearance of the extra submodule in the decomposition of the tensor product of the two adjoint representations if $n > 3$.

Looking at the Hamiltonian in Eq. (6.50), it may be tempting to consider the limit in which n, which labels the Lie algebra $\mathfrak{su}(n)$, gets very large, since the anti-Hermitian term carries a prefactor n^{-3}. However, this is a bit misleading, since this prefactor depends on the choice of normalization of the operators and the d-symbols. In fact, this anti-Hermitian carries a lot of summations over $n^2 - 1$ elements, which are present in the operators K and Q. When the local Hamiltonian is written in terms of projectors and operators that interchange the highest weight vectors, then the anti-Hermitian carries a prefactor that scales with $\sqrt{n^2-4}$ and this factor is dominant for large n. Furthermore, the large n-limit could give a deceptive picture for another reason. The most important motivation for considering these models, is their integrability. If we would slightly approximate our local Hamiltonian by a Hermitian one, the integrability would probably be lost. Furthermore, it could well be that adding a non-Hermitian term could lead to very different behaviour with respect to the eigenstates and eigenvalues.

6.8 Quasi-Hermitian Hamiltonians

As we remarked above, the local Hamiltonian in Eq. (6.47) is not really a Hamiltonian, since it not Hermitian. Therefore, the eigenvalues are not necessarily real. However, when $N = 2$, the eigenvalues are real. To see this, one has to look at the eigenvalues of the (2×2)-matrix that acts on the highest weight vectors which are real. Therefore, the Hamiltonian could lead to physical results after all. This is for instance the case if the Hamiltonian is quasi-Hermitian [14, 35]. We have the following definition for such a Hamiltonian.

Definition 6.1. *A diagonalizable linear operator H on a separable Hilbert space is called quasi-Hermitian if there exists an invertible positive Hermitian linear operator ρ such that $H^\dagger = \rho H \rho^{-1}$.*

In particular, every Hermitian operator is quasi-Hermitian, by taking $\rho = \mathbb{1}$. The following lemma justifies this definition.

Lemma 6.4. *The spectrum of a quasi-Hermitian operator is real.*

Proof. Let H be a quasi-Hermitian operator and ρ be a positive operator such that $H^\dagger = \rho H \rho^{-1}$. Since ρ is positive Hermitian, there exists η such that $\eta^\dagger \eta = \rho$. This η is invertible. Indeed, if not, then also $\eta^\dagger$ would not be invertible. Since the product of two non-invertible operators cannot be invertible, ρ would not be invertible in that case, which is a contradiction. As a result, the following equation holds

$$\left(\eta^\dagger\right)^{-1} H^\dagger \eta^\dagger = \eta H \eta^{-1},$$

which can be rewritten as

$$\left(\eta H \eta^{-1}\right)^\dagger = \eta H \eta^{-1}. \tag{6.51}$$

This implies that H equals a Hermitian operator up to conjugation. Since conjugate operators have the same spectrum, the result follows. □

What about the integrable Hamiltonian for the adjoint representation of $\mathfrak{su}(3)$? For $N = 2$ it is indeed quasi-Hermitian. As pointed out earlier, it is diagonalizable such that the resulting diagonal operator has real entries. This is clearly equivalent to Eq (6.51).

For larger N numerical calculations show that H is not necessarily quasi-Hermitian. Even for $N = 3$, some of the eigenvalues of the constructed Hamiltonian are imaginary. The reason is that the sum of quasi-Hermitian operators is not necessarily quasi-Hermitian. Consider the (2×2)-matrices A and B, defined by

$$A = \begin{pmatrix} 4\mathrm{i} & 5 \\ 5 & -4\mathrm{i} \end{pmatrix}, \; B = \begin{pmatrix} 0 & -5 \\ -5 & 0 \end{pmatrix}.$$

Then B is Hermitian and therefore quasi-Hermitian. The matrix A has eigenvalues -3, 3 and is therefore quasi-Hermitian. However, the sum of these operators equals a diagonal operator with imaginary entries and is certainly not quasi-Hermitian.

We conclude that the Hamiltonian constructed in the previous section is not really a Hamiltonian. It is an operator that has some properties similar to integrable Hamiltonians, in particular the local conserved quantities, but it can never describe a real physical system.

6.9 An integrable Hamiltonian for a reducible representation

In the previous sections we have looked at integrable spin chains where the states on every site spanned an irreducible representation of a Lie algebra. In this section, we will study an integrable spin chain on a reducible representation of $\mathfrak{su}(2)$, namely the direct sum of two spin-$\frac{1}{2}$ representation. Integrable spin chain on reducible representation have been constructed before [5, 6], in particular in the case of a tensor product of two irreducible representations. In such a case, one gets a system that could be interpreted as a spin ladder instead of a spin chain, where the sites are replaced by a bond with two sites. The interpretation of our construction is not immediately clear. The constructed spin chain will be based on the result from Section 5.6, where we already computed an R-matrix on this representation. Contrary to all the other R-matrices we encountered before, the R-matrix in Eq. (5.36) does not necessarily satisfy $R(0) = \sigma$, where σ is the permutation operator. This is an important property in the construction of an integrable spin chain and therefore it should be imposed by setting $J_1(0) = 1$ and $J_2(0) = J_3(0) = J_4(0) = 0$). We introduce the notation $J_i'(0) = J_i$ and put $J_1 = 0$. The local Hamiltonian can be computed as $R'(0)\sigma$ and is given by

$$\begin{aligned} h =& 2P_{\mathbf{0}} + J_2\left(M_{12} + M_{34} + N_{12} + N_{34}\right) \\ &+ J_3\left(M_{13} + M_{24} + N_{13} + N_{24}\right) + J_4\left(M_{14} + M_{23} + N_{14} + N_{23}\right), \end{aligned}$$

where the notation is as in Section 5.6. Contrary to the integrable Hamiltonians we encountered so far, there are now four parameters we can choose from instead of an overall normalization. Every operator on $\mathbf{\frac{1}{2}} \oplus \mathbf{\frac{1}{2}}$ can be written as a (2×2)-matrix which entries are operators on $\mathbf{\frac{1}{2}}$. We can also do this for the local Hamiltonian which gives (putting $J_3 = J_4 = 0$ for simplicity)

$$\begin{aligned} h = &- \begin{pmatrix} S^a & 0 \\ 0 & S^a \end{pmatrix} \otimes \begin{pmatrix} S^a & 0 \\ 0 & S^a \end{pmatrix} + \frac{1}{4} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \otimes \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \\ &+ J_2 \left(\begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \otimes \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix} \right) \end{aligned} \tag{6.52}$$

in terms of the spin operators $S^a = \frac{1}{2}\sigma^a$, with σ^a the Pauli matrices. Naively, the first term in the local Hamiltonian in Eq. (6.52) could be seen as the sum of two local Hamiltonians of the Heisenberg model, acting on both copies of $\mathbf{\frac{1}{2}}$ on each site. The term proportional to J_2 could be seen as an interaction between the two states. However, this interpretation is a bit misleading. One could have made a different choice of highest weight vectors for the two $\mathbf{\frac{1}{2}}$-modules in the beginning, such that the inner product is preserved. Such a transformation can be identified with conjugation by this unitary (2×2)-matrix that acts on the (2×2)-matrices in the local Hamiltonian in Eq. (6.52). For arbitrary unitary transformations, the eigenvalues of the Hamiltonian are not preserved. This is a major problem for the Hamiltonian that severely limits the physical implications of this model. For a genuine physical model, the choice of highest weight vectors should not amount to different eigenvectors with different eigenvalues, since this implies that the states and corresponding energies can not be determined objectively. Therefore, we conclude that the Hamiltonian given in Eq. (6.52) is not a physical model.

6.10 Summary

In this chapter, we have studied integrable spin chains. We have reviewed how solutions of the Yang-Baxter equation lead to the construction of such integrable models. After that, we constructed all the integrable models out of the R-matrices from the previous chapter, starting with the lowest-dimensional non-trivial irreducible representations of $\mathfrak{su}(2)$ to retrieve the well-known Heisenberg and Babujian-Takhtajan models. Furthermore, we reviewed the Bethe Ansatz, which is a method to obtain an abstract description of the eigenstates and eigenvalues of the Hamiltonian. After that, we tried to mimic these methods for the adjoint representation of $\mathfrak{su}(3)$. However, the constructed Hamiltonian is not Hermitian. This is a consequence of the decomposition of the tensor product of two adjoint representations of $\mathfrak{su}(n)$, which is not free of multiplicities if $n > 2$. Since the eigenvalues of the Hamiltonian are also not real in general, it seems to be impossible to describe a physical system with this Hamiltonian. We also tried to find the Bethe equations for the Hamiltonian. However, a direct approach leads to very complicated equations, while an indirect approach involving the Lax operator did not succeed because we were not able to find an expression for the Lax operator in the case of $\mathfrak{su}(3)$. We also presented the Hamiltonian that follows for the conjectured expression for the R-matrix corresponding to the adjoint representation of $\mathfrak{su}(n)$, which we presented in the previous chapter for $n > 3$. This Hamiltonian has the same problems as in the case $n = 3$. Finally, we presented the Hamiltonian which follows from the R-matrix on the direct sum of two spin-$\frac{1}{2}$ representations of $\mathfrak{su}(2)$. However, such a Hamiltonian seems unphysical as the eigenvalues of it depend on the choice of highest weight vectors for the modules.

Chapter 7

Integrable spin chains with defects

7.1 Introduction

In the previous chapter we have constructed various integrable spin chains. We started with the well known Heisenberg and Babujian-Takhtajan models and then constructed an integrable spin chain for the adjoint representation of $\mathfrak{su}(3)$. Although the latter differed from the first two models in the sense that it was non-Hermitian, all these systems were homogeneous as all sites were considered to be essentially the same. In this chapter, we will consider integrable systems which are not translationally invariant, by introducing defects on one or more of the sites. Such systems are interesting for several reasons. First of all, different kinds of spatial dependence in the system become possible as a result of the broken translational symmetry. Furthermore, such a defect could be seen as a link between two different systems. Finally, if one introduces many defects on the sites, it may be possible to study integrable systems with disorder.

The idea of implementing a defect in the system is very simple. One replaces one of the R-matrices in the monodromy operator from Eq. (6.21) by a so called Lax operator $L(\lambda)$ that satisfies the Lax equation

$$R_{12}(\lambda)L_{13}(\lambda+\mu)L_{23}(\mu) = L_{23}(\mu)L_{13}(\lambda+\mu)R_{12}(\lambda). \tag{7.1}$$

We encountered this equation before in the derivation of the Bethe equations for higher spin $\mathfrak{su}(2)$-invariant models. It is an operator equation on $V \otimes V \otimes W$ (with V and W vector spaces), where $R(\lambda)$ is an R-matrix on $V \otimes V$. The Lax operator is an operator acting on $V \otimes W$. This method has been used to construct various kinds of integrable systems, for example an anisotropic spin-$\frac{1}{2}$ XXZ model with defects [42] and a model with spin-$\frac{1}{2}$ as well as spin-s sites [2, 49]. In this chapter, we will only investigate the case where $W = V$, so where the Lax operator is defined on the same vector space as the R-matrix. For the $\mathfrak{su}(2)$-invariant models we consider, this means that we will equip all lattice sites with the same representation. Furthermore, the defects will not break the Lie algebra invariance in our models.

This chapter is organized as follows. First, we develop a general formula for a Hamiltonian with a defect on one of the sites. Two examples, a spin-$\frac{1}{2}$ and a spin-1 spin chain with a defect are presented. After, that we will consider spin chains with more defects. Finally, we will show how the Bethe equations need to be modified if defects are present.

7.2 Integrable Hamiltonians with one defect

As noted in the introduction, the idea of creating a defect is by replacing one of the R-matrices in the monodromy matrix, by a solution of the Lax equation (7.1). One can place as many defects as desired, but the calculations will get more cumbersome in the case of more defects. For that reason, we first investigate the case of a single defect. Since all sites are essentially the same in the original model (due to the cyclicity of the partial trace in the transfer matrix), it does not matter which site we pick for the defect. For that reason, we can assume the defect to be on site with label 1. The new monodromy operator becomes

$$T_a(\lambda) = L_{a1}(\lambda)R_{a2}(\lambda)\cdots R_{aN}(\lambda)$$

and the fundamental commutation relation from Eq. (6.2) still holds, with the same proof. The transfer matrix $F(\lambda) = \mathrm{tr}_a T(\lambda)$ is still commuting for different values of the spectral parameter, i.e. $[F(\lambda), F(\mu)] = 0$. The Hamiltonian is constructed as before

$$H = F(0)^{-1}F'(0). \tag{7.2}$$

We assume that the R-matrix satisfies $R(0) = \sigma$, but the Lax operator does not necessarily equal the permutation operator at $\lambda = 0$ (if $V \neq W$ this is even impossible). As a result, the derivation of the Hamiltonian is more involved than before. We note that we have assumed $V = W$, which allows us to use various identities with permutation operators. We start with the computation of $F(0)$,

$$\begin{aligned} F(0) &= \mathrm{tr}_a\left(L_{a1}(0)\sigma_{a2}\cdots\sigma_{aN}\right) \\ &= \mathrm{tr}_a\left(L_{a1}(0)\sigma_{a1}\sigma_{a1}\cdots\sigma_{aN}\right) \\ &= \mathrm{tr}_a\left(\sigma_{aN}L_{N1}(0)\sigma_{N1}\sigma_{N(N-1)}\cdots\sigma_{12}\right) \\ &= L_{N1}(0)\sigma_{N1}\sigma_{N(N-1)}\cdots\sigma_{12}. \end{aligned} \tag{7.3}$$

For later use, we note that we could also first use the cyclicity of the trace to get $L_{a1}(0)$ to the right and pull σ_{a2} to the right,

$$\begin{aligned} F(0) &= \mathrm{tr}_a\left(\sigma_{a2}\cdots\sigma_{aN}L_{a1}(0)\right) \\ &= \mathrm{tr}_a\left(\sigma_{23}\cdots\sigma_{2(N-1)}\sigma_{2N}L_{21}(0)\sigma_{a2}\right) \\ &= \sigma_{23}\cdots\sigma_{2(N-1)}\sigma_{2N}L_{21}(0) \end{aligned} \tag{7.4}$$

The derivative $F'(0)$ can be written as

$$F'(0) = \mathrm{tr}_a\left(L'_{a1}(0)\sigma_{a2}\cdots\sigma_{aN}\right) + \sum_{i=2}^{N}\mathrm{tr}_a\left(L_{a1}(0)\sigma_{a2}\cdots\sigma_{a(i-1)}h_{ai}\sigma_{ai}\cdots\sigma_{aN}\right) \tag{7.5}$$

where we have used the notation

$$h_{ab} = R'_{ab}(0)\sigma_{ab}.$$

The first term in Eq. (7.5) can be computed similarly to Eq. (7.3) and is given by

$$L'_{N1}(0)\sigma_{N1}\sigma_{N(N-1)}\cdots\sigma_{12}. \tag{7.6}$$

The first $N-2$ terms in the summation in Eq. (7.5) can be found by using $h_{ai}\sigma_{ai}\sigma_{a(i+1)} = \sigma_{ai}\sigma_{a(i+1)}h_{i,i+1}$ and by pulling σ_{aN} to the left,

$$\sum_{i=2}^{N-1}\mathrm{tr}_a\left(L_{a1}(0)\sigma_{a2}\cdots\sigma_{a(i-1)}h_{ai}\sigma_{ai}\cdots\sigma_{aN}\right)$$

$$
\begin{aligned}
&= \sum_{i=2}^{N-1} \mathrm{tr}_a \left(L_{a1}(0) \sigma_{a2} \cdots \sigma_{a(i-1)} \sigma_{ai} \cdots \sigma_{aN} h_{i(i+1)} \right) \\
&= L_{N1}(0) \sigma_{N1} \sigma_{N(N-1)} \cdots \sigma_{12} \sum_{i=2}^{N-1} h_{i(i+1)} \\
&= F(0) \sum_{i=2}^{N-1} h_{i(i+1)}. \qquad (7.7)
\end{aligned}
$$

The final term in the summation in Eq. (7.5) equals

$$
\begin{aligned}
\mathrm{tr}_a \left(L_{a1}(0) \sigma_{a2} \cdots \sigma_{a(N-1)} h_{aN} \sigma_{aN} \right) &= \mathrm{tr}_a \left(\sigma_{a2} \cdots \sigma_{a(N-1)} \sigma_{aN} h_{Na} L_{a1}(0) \right) \\
&= \mathrm{tr}_a \left(\sigma_{23} \cdots \sigma_{2(N-1)} \sigma_{2N} h_{N2} L_{21}(0) \sigma_{a2} \right) \\
&= \sigma_{23} \cdots \sigma_{2(N-1)} \sigma_{2N} h_{N2} L_{21}(0). \qquad (7.8)
\end{aligned}
$$

Multiplying the right hand sides of Eqs. (7.6), (7.7) and (7.8) with $F(0)^{-1}$ from Eqs. (7.3) and (7.4) on the left, leads to the Hamiltonian

$$
H = L_{21}(0)^{-1} L'_{21}(0) + \sum_{i=2}^{N-1} R'_{i(i+1)}(0) \sigma_{i(i+1)} + L_{21}(0)^{-1} R'_{N2}(0) \sigma_{N2} L_{21}(0), \qquad (7.9)
$$

where we used the identity

$$
\sigma_{12} \cdots \sigma_{N(N-1)} \sigma_{N1} L_{N1}(0)^{-1} L'_{N1}(0) \sigma_{N1} \sigma_{N(N-1)} \cdots \sigma_{12} = L_{21}(0)^{-1} L'_{21}(0).
$$

A priori, the meaning of the Hamiltonian in Eq. (7.9) is unclear. For instance, there is no guarantee that the Hamiltonian is Hermitian or that it physical properties will truely differ from a Hamiltonian without a defect. However, some small differences are immediately apparent. First of all, the first term in the Hamiltonian from Eq. (7.9) is an interaction term involving three sites. Furthermore, the roles of the neighbouring sites of the defect is no longer symmetric, i.e. the labels N and 2 appear asymmetrically in the expression for the Hamiltonian. In the next section, we will see the simplest example of a Hamiltonian, in the case of the spin-$\frac{1}{2}$ Heisenberg model with a single defect.

7.3 The Heisenberg model with one defect

In this section, we will study the spin-$\frac{1}{2}$ Heisenberg model with a defect. We recall that the R-matrix that leads to the Heisenberg Hamiltonian equals

$$
R(\lambda) = P_{\mathbf{1}} + \frac{\lambda + \eta}{\lambda - \eta} P_{\mathbf{0}} \qquad (7.10)
$$

with $P_{\mathbf{s}}$ the projection onto the submodule $\mathbf{s}$ in the decomposition $\frac{\mathbf{1}}{\mathbf{2}} \otimes \frac{\mathbf{1}}{\mathbf{2}}$. Here we recall that the notation $\mathbf{s}$ is used for the $(2s+1)$-dimensional irreducible representation of $\mathfrak{su}(2)$. We keep the extra parameter η in place, to keep the discussion as general as possible. The Lax operator that we choose is $L(\lambda) = R(\lambda + \mathrm{i}\gamma)$, with γ the (real) defect parameter. The appearance of the factor i will become clear later on. The fact that this Lax operator satisfies the Lax equation follows by taking a shift of parameters in the Yang-Baxter equation. The local Hamiltonian is

$$
h = \frac{2}{\eta} P_{\mathbf{0}} = \frac{2}{\eta} \left(\frac{1}{4} - \mathbf{S} \otimes \mathbf{S} \right). \qquad (7.11)
$$

The term $L_{12}(0)^{-1}L'_{12}(0)$ in the Hamiltonian can straightforwardly be computed as

$$L_{12}(0)^{-1}L'_{12}(0) = \frac{2\eta}{\eta^2+\gamma^2}\left(\frac{1}{4} - \mathbf{S}_1\cdot\mathbf{S}_2\right). \tag{7.12}$$

The final term in Eq. (7.9) is more difficult to compute, as it contains the interactions between multiple sites. We note the following identities which can be shown by explicit computation,

$$\begin{aligned}
(\mathbf{S}_1\cdot\mathbf{S}_2)(\mathbf{S}_2\cdot\mathbf{S}_N) &= \frac{1}{4}(\mathbf{S}_1\cdot\mathbf{S}_N) - \frac{i}{2}\mathbf{S}_1\cdot(\mathbf{S}_2\times\mathbf{S}_N)\\
(\mathbf{S}_2\cdot\mathbf{S}_N)(\mathbf{S}_1\cdot\mathbf{S}_2) &= \frac{1}{4}(\mathbf{S}_1\cdot\mathbf{S}_N) + \frac{i}{2}\mathbf{S}_1\cdot(\mathbf{S}_2\times\mathbf{S}_N)\\
(\mathbf{S}_1\cdot\mathbf{S}_2)(\mathbf{S}_2\cdot\mathbf{S}_N)(\mathbf{S}_1\cdot\mathbf{S}_2) &= -\frac{1}{16}(\mathbf{S}_2\cdot\mathbf{S}_N) + \frac{1}{8}(\mathbf{S}_1\cdot\mathbf{S}_N).
\end{aligned}$$

With the help of these identities, one finds that

$$L_{21}(0)^{-1}h_{N2}L_{21}(0) = \frac{2}{\eta} - \frac{2}{\eta}\frac{\gamma^2}{\eta^2+\gamma^2}(\mathbf{S}_2\cdot\mathbf{S}_N) - \frac{2\eta}{\eta^2+\gamma^2}(\mathbf{S}_1\cdot\mathbf{S}_N) + \frac{4\gamma}{\eta^2+\gamma^2}\mathbf{S}_1\cdot(\mathbf{S}_2\times\mathbf{S}_N). \tag{7.13}$$

Adding Eqs. (7.12) and (7.13) with the local Hamiltonians for the non-defected sites, leads up to a multiplicative factor $2/\eta$ and an additive constant to the following result

$$H = -\sum_{i=1}^{N}\mathbf{S}_i\cdot\mathbf{S}_{i+1} + H_{\text{defect}} \tag{7.14}$$

where the first term is precisely the Heisenberg Hamiltonian. The defect Hamiltonian H_{defect} equals

$$H_{\text{defect}} = \sin^2(\Delta)(\mathbf{S}_N\cdot\mathbf{S}_1 + \mathbf{S}_1\cdot\mathbf{S}_2 - \mathbf{S}_N\cdot\mathbf{S}_2) + \sin(2\Delta)\mathbf{S}_N\cdot(\mathbf{S}_1\times\mathbf{S}_2). \tag{7.15}$$

where we have introduced the parameter $\Delta = \arctan(\gamma/\eta)$. We note that the Heisenberg model is exactly retrieved in the case $\Delta = 0$ (corresponding to $\gamma = 0$) where the defect is absent. The case $\Delta = \pi/2$ (or $\gamma = \infty$) also corresponds to the Heisenberg Hamiltonian, but only on the $N-1$ non-defected sites. We also note the defect Hamiltonian is Hermitian as a result of the factor i in the Lax operator $L(\lambda) = R(\lambda + \mathrm{i}\gamma)$. Physically, the first term in Eq. (7.15) corresponds to a decrease of the strength of the interaction between de defect site and the neighbouring sites, while a next-nearest neighbour interaction between the two neighbours of the defect is created. The second term in Eq. (7.15) is a three-site interaction term on the defect site and the two neighbouring sites. The strength of the defect Hamiltonian is bounded, as the relative prefactors of the terms in the defect Hamiltonian with the Heisenberg are only sines, which are limited to take values between -1 and 1.

The three site interaction given by $\mathbf{S}_N\cdot(\mathbf{S}_1\times\mathbf{S}_2)$ is in fact the only $\mathfrak{su}(2)$-invariant interaction that acts non-trivially on three sites, i.e. any such interaction can be rewritten as a linear combination of this interaction, two-site interactions and constants. Because, the number of spin operators in this interaction is odd, this term breaks the time-reversal symmetry of the system, since angular momentum is odd under this symmetry. Contrary to the Heisenberg model, this interaction is also not invariant under a reversal of the spin chain (with all the sites in the opposite order) as interchanging sites N and 2 leads to a minus sign. If we would consider the operator $\mathbf{S}_N\cdot(\mathbf{S}_1\times\mathbf{S}_2)$ on just the three sites, we can easily find its eigenvalues. The $\mathfrak{su}(2)$ on these three sites decomposes as

$$\mathbf{\frac{1}{2}}\otimes\mathbf{\frac{1}{2}}\otimes\mathbf{\frac{1}{2}} = \mathbf{\frac{3}{2}}\oplus\mathbf{\frac{1}{2}}\oplus\mathbf{\frac{1}{2}}.$$

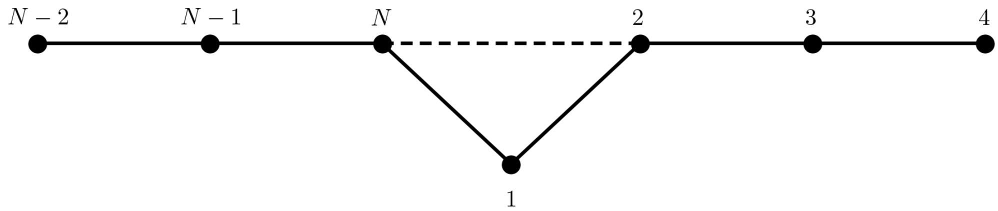

Figure 7.1: A possible interpretation of a defect site in a spin chain, where the defect site is slightly moved from the rest of the chain.

We choose the highest weight vectors for the two $s = \frac{1}{2}$ modules in these decomposition as

$$v_1 = \frac{1}{\sqrt{3}} \left(|\uparrow\rangle |\uparrow\rangle |\downarrow\rangle + \alpha |\uparrow\rangle |\downarrow\rangle |\uparrow\rangle + \alpha^2 |\downarrow\rangle |\uparrow\rangle |\uparrow\rangle \right)$$
$$v_2 = \frac{1}{\sqrt{3}} \left(|\uparrow\rangle |\uparrow\rangle |\downarrow\rangle + \alpha^2 |\uparrow\rangle |\downarrow\rangle |\uparrow\rangle + \alpha |\downarrow\rangle |\uparrow\rangle |\uparrow\rangle \right)$$

where $|\uparrow\rangle$ ($|\downarrow\rangle$) is the highest (lowest) weight state of the $\mathbf{\frac{1}{2}}$-module and the tensor product between the states is omitted. The complex number α is the solution of $1 + \alpha + \alpha^2 = 0$ with positive imaginary part. In this basis, the operator $\mathbf{S}_N \cdot (\mathbf{S}_1 \times \mathbf{S}_2)$ can be written as [50]

$$\mathbf{S}_N \cdot (\mathbf{S}_1 \times \mathbf{S}_2) = 2\sqrt{3}(P_{\left(\frac{1}{2}\right)_2} - P_{\left(\frac{1}{2}\right)_1}), \tag{7.16}$$

with $\left(\mathbf{\frac{1}{2}}\right)_i$ is the module spanned by v_i. In particular, this operator kills the module with highest weight vector the ferromagnetic state $|\uparrow\rangle |\uparrow\rangle |\uparrow\rangle$, while it distinguishes the other two 'chiral' states with $s = \frac{1}{2}$.

Furthermore, we note that the Hamiltonian in Eq. (7.14) is still Lie algebra invariant, since the Lax operator that we considered was also Lie algebra invariant. It may be possible that there are other Lax operators that solve the Lax equation on $\mathbf{\frac{1}{2}} \otimes \mathbf{\frac{1}{2}} \otimes \mathbf{\frac{1}{2}}$ as well (leading to different Hamiltonians), but the Lax operator that we used is the unique $\mathfrak{su}(2)$-invariant solution. Picking another Lax operator would therefore probably break the Lie algebra invariance of the system.

A possible interpretation of the origin of such a defect is depicted in Fig. 7.1. The defect site can be considered as a 'kink in the cable', where it is taken out of the spin chain, ensuring that the two neighbors of the defect will get nearer, creating an interaction on all the three sites. Finally, we note that the extra parameter η only amounts to an overal multiplicative constant and a scaling in the defect parameter γ. This is again a consequence of the fact that a rescaling of the spectral parameter λ in the R-matrix can be used to set $\eta = 1$.

We also state the Hamiltonian for the Babujian-Takhtajan model with a single defect on the first site. Up to an additive constant and a multiplicative factor $(2\eta)^{-1}$ it is given by

$$H = \sum_{i=1}^{N} (C_{i(i+1)}^2 - C_{i(i+1)}) + H_{\text{defect}}$$

with $C_{ij} = \mathbf{S}_i \cdot \mathbf{S}_j$. The defect Hamiltonian is

$$H_{\text{defect}} = \frac{\delta^2}{\delta^2 + 4} \left[\left(C_{N2}^2 - C_{N2} \right) - \left(C_{N1}^2 - C_{N1} \right) - \left(C_{12}^2 - C_{12} \right) \right]$$

$$+\frac{2\delta}{\delta^2+4}D_{N12}+\frac{3\delta^2}{(\delta^2+4)(\delta^2+1)}\left(2D_{N12}^2+\{C_{12}+C_{N1},C_{N2}-D_{N12}\}\right)$$
$$+\frac{\delta\left(1-2\delta^2\right)}{(\delta^2+4)(\delta^2+1)}\{D_{N12},C_{N2}\}.$$

Here, the operator D_{ijk} is defined by $D_{ijk}=\mathbf{S}_i\cdot(\mathbf{S}_j\times\mathbf{S}_k)$ and $\delta=\gamma/\eta$. Although the spin-1 defect Hamiltonian is much more involved than its spin-$\frac{1}{2}$ counterpart, there are also some similarities. First of all, the Hamiltonian is again Hermitian. This is a consequence of two facts about the $\mathfrak{su}(2)$ R-matrix. First, the R-matrix for any irreducible representation of $\mathfrak{su}(2)$ (when normalized such that is preserves the highest weight state) is unitary for purely imaginary spectral parameter, i.e. $R(\mathrm{i}\gamma)^\dagger R(\mathrm{i}\gamma)=\mathbb{1}$. Secondly, the logarithmic derivative at $\mathrm{i}\gamma$ of the coefficients in the R-matrix is also real if we choose that particular normalization. Furthermore, the structure of the defect Hamiltonian is very similar to the one in Eq. (7.15). The two-site interactions decrease the strength of the interaction between the defect site and its neighbouring sites, while creating an interaction between the neighbouring sites. The other terms are all interactions involving the defect site and both its neighbouring sites. Finally, setting the defect parameter γ equal to 0 results in the Babujian-Takhtajan model, while in the limit $\gamma\to\infty$ the defect site disappears from the system.

7.4 Integrable Hamiltonians with more defects

In the previous sections we have seen the construction and examples of integrable Hamiltonians with one defect. In principle the same method can be used to introduce more than one defect in the system. In this section we will explore such systems, starting with two defects. There are two possibilities for the second defect, either next to the first defect site or further away. We will start with the first case and we put two defects on the first two sites, which leads to the monodromy operator

$$T(\lambda)=L_{a1}(\lambda)\tilde{L}_{a2}(\lambda)R_{a3}(\lambda)\cdots R_{aN}(\lambda)$$

with two possibly different Lax operators $L(\lambda)$ and $\tilde{L}(\lambda)$. The corresponding Hamiltonian can be derived as in Section 7.2 as

$$H=\tilde{L}_{32}(0)^{-1}L_{31}(0)^{-1}L_{31}'(0)\tilde{L}_{32}(0)+\tilde{L}_{32}(0)^{-1}\tilde{L}_{32}'(0)+\sum_{i=3}^{N-1}h_{i(i+1)}$$
$$+\tilde{L}_{32}(0)^{-1}L_{31}(0)^{-1}h_{N3}L_{31}(0)\tilde{L}_{32}(0)^{-1},$$

where we recall that $h=R'(0)\sigma$. The final term contains a nontrivial action between four sites. This distinguishes this Hamiltonian from the one where only one defect was present. So, defects on neighbouring sites create an even larger defect with interaction terms on all of the defect sites and the neighbouring sites.

The second case corresponds to defects that are not next to each other. In principle, it could matter how many sites there are between defects, but we restrict ourselves to the case of minimum distance of one site. This corresponds to the monodromy operator

$$T(\lambda)=L_{a1}(\lambda)R_{a2}(\lambda)\tilde{L}_{a3}(\lambda)R_{a4}(\lambda)\cdots R_{aN}(\lambda).$$

The derivation of the Hamiltonian is again similar to the one in Section 7.2 and leads to

$$H=L_{21}(0)^{-1}h_{N2}L_{21}(0)+L_{21}(0)^{-1}L_{21}'(0)+\tilde{L}_{43}(0)^{-1}h_{24}\tilde{L}_{43}(0)+\tilde{L}_{43}(0)^{-1}\tilde{L}_{43}'(0)+\sum_{i=4}^{N-1}h_{i(i+1)}.$$

Interestingly, the two defects do not interact with each other, but are independent. Therefore, there are only two-site and three-site interactions, but no higher order interactions. It is clear that this situation generalizes to more defects, as long as no defects are placed next to each other. With this in mind, we can construct integrable models with $N = 2M$ sites, with M defects and at most three-site interactions. For example, for the spin-$\frac{1}{2}$ Heisenberg model this approach leads to

$$H = -\sum_{i=1}^{N} \mathbf{S}_i \cdot \mathbf{S}_{i+1} + H_{\text{defect}}$$

with

$$H_{\text{defect}} = \sum_{i=1}^{M} \sin^2(\Delta_i) \left(\mathbf{S}_{2i-1} \cdot \mathbf{S}_{2i} + \mathbf{S}_{2i} \cdot \mathbf{S}_{2i+1} - \mathbf{S}_{2i-1} \cdot \mathbf{S}_{2i+1}\right) + \sin(2\Delta_i)\mathbf{S}_{2i-1} \cdot \left(\mathbf{S}_{2i} \times \mathbf{S}_{2i+1}\right).$$

7.5 Bethe equations for models with defects

The most important reason for studying the integrable models constructed using the quantum inverse scattering method, is the fact that one can find a description for all the eigenvalues and eigenstates of the model. In this section, we will mimic the construction from Section 6.5 for the Heisenberg spin chain with defects. We put the extra parameter $\eta = 1$ for simplicity. Furthermore, we will assume that on each site there is a defect corresponding to the Lax operator $L^{\gamma}(\lambda) = R(\lambda + \mathrm{i}\gamma)$, which leads to the monodromy matrix

$$T(\lambda) = L_{a1}^{\gamma_1}(\lambda) \cdots L_{aN}^{\gamma_N}(\lambda). \tag{7.17}$$

First, one writes the monodromy operator as a matrix in the auxiliary space,

$$T(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix} \tag{7.18}$$

The entries of this matrix are operators acting on the spin chain and they differ from the operators on the chain without defects. However, the fundamental commutation relation from Eq. (6.2) will lead to the same set of commutation-like relations between these operators since the R-matrix is the same. So, the families of operators $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ $D(\lambda)$ are all commuting, i.e.

$$[A(\lambda), A(\mu)] = 0, \;\; [B(\lambda), B(\mu)] = 0, \text{etc.}$$

Furthermore, we have the relations

$$\begin{aligned} A(\lambda)B(\mu) &= f_1(\lambda-\mu)B(\mu)A(\lambda) + f_2(\lambda-\mu)B(\lambda)A(\mu), \\ D(\lambda)B(\mu) &= g_1(\lambda-\mu)B(\mu)D(\lambda) + g_2(\lambda-\mu)B(\lambda)D(\mu), \end{aligned}$$

with $f_1(\lambda) = (\lambda+1)/\lambda$, $g_1(\lambda) = (\lambda-1)/\lambda$ and $f_2(\lambda) = -g_2(\lambda) = -1/\lambda$. For convenience, we multiply the R-matrix with a factor $1-\lambda$, such that it is given in matrix form as

$$R(\lambda) = \begin{pmatrix} \frac{1}{2} - \lambda + S^3 & S^1 - \mathrm{i}S^2 \\ S^1 + \mathrm{i}S^2 & \frac{1}{2} - \lambda - S^3 \end{pmatrix}. \tag{7.19}$$

The Lax operator can be written in similar form as

$$L^{\gamma}(\lambda) = \begin{pmatrix} \frac{1}{2} - \lambda - \mathrm{i}\gamma + S^3 & S^1 - \mathrm{i}S^2 \\ S^1 + \mathrm{i}S^2 & \frac{1}{2} - \lambda - \mathrm{i}\gamma - S^3 \end{pmatrix}. \tag{7.20}$$

Let ω be the highest weight vector of the $\frac{\mathbf{1}}{\mathbf{2}}$-representation and $\Omega = \omega \otimes \cdots \otimes \omega$. Then, the transfer matrix $F(\lambda) = A(\lambda) + D(\lambda)$ acts on Ω as

$$F(\lambda)\Omega = (h_1(\lambda) + h_2(\lambda))\,\Omega, \tag{7.21}$$

with

$$h_1(\lambda) = \prod_{i=1}^{N}(1 - \lambda - \mathrm{i}\gamma_i),\ \ h_2(\lambda) = \prod_{i=1}^{N}(-\lambda - \mathrm{i}\gamma_i).$$

The rest of the derivation is analogous to the one in Section 6.5. The other eigenstates of $F(\lambda)$ are searched in the form

$$\Phi(\{\lambda\}) = B(\lambda_1)\cdots B(\lambda_m)\Omega.$$

and the transfer matrix acts on this state as

$$F(\lambda)\Phi(\{\lambda\}) = \Lambda(\{\lambda\},\lambda)\Phi(\{\lambda\})$$

with

$$\Lambda(\{\lambda\},\lambda) = h_1(\lambda)\prod_{j=1}^{m} f_1(\lambda-\lambda_j) + h_2(\lambda)\prod_{j=1}^{m} g_1(\lambda-\lambda_j) \tag{7.22}$$

provided the Bethe equations hold

$$\prod_{\substack{j=1\\ j\neq k}}^{m} \frac{\lambda_k - \lambda_j + 1}{\lambda_k - \lambda_j - 1} = \prod_{i=1}^{N} \frac{\lambda_k + \mathrm{i}\gamma_i}{\lambda_k + \mathrm{i}\gamma_i - 1}. \tag{7.23}$$

These Bethe equations have the same structure as the equations without defects in Eq. (6.35). However, due to the defects, not all sites are treated equal anymore and therefore the right hand side in Eq. (7.23) has a different term in the product for each site. To compute the eigenvalues of the corresponding Hamiltonian, we should take the logarithmic derivative of Eq. (7.22). First, we should reinstate the factor $1-\lambda$ that we divided out in the R-matrix. Then the eigenvalues of the transfer matrix become

$$\tilde{\Lambda}(\{\lambda\},\lambda) = \prod_{j=1}^{m} f_1(\lambda-\lambda_j) + \frac{h_2(\lambda)}{h_1(\lambda)}\prod_{j=1}^{m} g_1(\lambda-\lambda_j). \tag{7.24}$$

The logarithmic derivative of this quantity is extremely complicated, due to the defects. However, for systems that do not have defects on neighbouring sites, the second term in Eq. (7.24) and its derivative vanish at $\lambda = 0$ and the eigenvalues of the Hamiltonian equal

$$Z(\{\lambda\}) = \sum_{j=1}^{m} \frac{1}{\lambda_j(1-\lambda_j)} \tag{7.25}$$

which is the same expression as for the Hamiltonian without defects. However, the conclusion that the eigenvalues of the Hamiltonian are the same would be false, as the Bethe equations are different. We note that if there are no defects on neighbouring sites, then $\gamma_i = 0$ for at least half of the sites. Similar tp the case without defects, one case show that $\Phi(\{\lambda\})$ is a spin highest weight state of spin $\frac{N}{2} - m$, provided the Bethe equations hold.

We also state the Bethe equations for $\mathfrak{su}(2)$-invariant Hamiltonians with defects corresponding to higher-dimensional representations. These can be obtained similarly as in the previous example and read

$$\prod_{\substack{j=1 \\ j\neq k}}^{m} \frac{\lambda_k - \lambda_j + 1}{\lambda_k - \lambda_j - 1} = \prod_{i=1}^{N} \frac{\lambda_k + \mathrm{i}\gamma_i - \frac{1}{2} + s}{\lambda_k + \mathrm{i}\gamma_i - \frac{3}{2} + s}$$

for the spin-s representation.

7.6 Summary

In this chapter we have investigated integrable spin chain with defects. These systems are no longer translationally invariant, as the local Hamiltonians for some of the sites are different than for others. We have looked at systems where the defect is implemented by a shift in the R-matrix on one of the sites. Systems with one defect will contain also next-nearest neighbour interactions as well three-site interactions around the defect site. Defects next to each other will generate even higher order interactions involving more sites. However, if defects are not on neighbouring sites, then they are essentially independent and the defect Hamiltonians can be added separately to the original Hamiltonian. We have computed the defect Hamiltonian for a single defect on the spin-$\frac{1}{2}$ Heisenberg and the spin-1 Babujian-Takhtajan model. The structure of both defect Hamiltonians is very similar. There is a decrease of the interaction strength between the defect site and its neighbouring sites, while an interaction between the two neighbours of the defect is created. Furthermore, three-site interactions appear in the system around the defect. Finally, we have computed the Bethe equations for the Heisenberg Hamiltonian. The structure of the Bethe equations is the same as in the case without defects, although there is a correction for the defects. In the case of the Heisenberg model with non-neighbouring defects, the abstract expression for the eigenvalues in terms of a set of parameters λ_i is the same as in the case without defects, although the eigenvalues itself will be different due to the correction in the Bethe equations.

Chapter 8

Conclusion and discussion

In this thesis we have analyzed various manifestations of the Yangian in the theory of spin chains. In particular, we have looked at the Haldane-Shastry model and integrable spin chains constructed by the Quantum Inverse Scattering Method (QISM). In the case of the Haldane-Shastry model the Yangian appears in the symmetry algebra, which provides a way to find explicit expressions for the ground state, the elementary excitations and the corresponding energies. [43, 44]. In the QISM, the significance of the Yangian resides at a much deeper level. The representation theory of the Yangian is used to find so called R-matrices, which are solutions of the Yang-Baxter equation (YBE) [10, 11]. In the QISM, every R-matrix leads to an integrable model, which can be solved in principle using the Algebraic Bethe Ansatz (ABA), which gives an abstract description of all the eigenstates and eigenvalues [19, 41].

In the first two chapters after the introduction, we reviewed some mathematical background concerning Lie algebras and Yangian. Starting with Lie algebras, we studied its root space decomposition and representation theory. In particular, we reviewed some special facts on $\mathfrak{su}(3)$ which we needed later on. The chapter on the Yangian focuses mainly on results obtained by Drinfeld in 1985 [15]. The Yangian is an algebra that extends a Lie algebra, similar to the universal enveloping algebra, although the Yangian has extra generators and its defining relations are much more involved. One of its key properties is the existence of a comultiplication on the Yangian, which gives the Yangian the structure of a Hopf algebra. Among other things, it allows for the construction of a tensor product representation out of two Yangian representations. For $\mathfrak{su}(n)$ there exists a nontrivial homomorphism from the Yangian to the universal enveloping algebra, for $\lambda \in \mathbb{C}$. This construction is useful to extend a representation of $\mathfrak{su}(n)$, to infinitely many representations of the Yangian.

In Chapter 4 we investigated the Haldane-Shastry model. It is a $\mathfrak{su}(n)$-invariant spin chain, with all sites carrying the fundamental representation of $\mathfrak{su}(n)$. In particular, we looked at the symmetry algebra of the model, which is a representation of the Yangian, as found out by Haldane et al. [22]. The algebra is generated by the total spin and another operator, which is called the rapidity. For $n = 2$, we showed in full detail that these operators commute with the Hamiltonian and satisfy the defining relations of the Yangian. Finally, we gave an argument why a $\mathfrak{su}(2)$ spin-1 chain with Yangian symmetry cannot exist. The reason for this is that only in the fundamental representation the $\mathfrak{su}(2)$ operators satisfy a Clifford algebra structure that ensures the generating relations of the Yangians hold. For $\mathfrak{su}(n)$ $(n > 2)$ the Lie algebra operators also satisfy a similar structure in the fundamental representation which allows the construction of the Haldane-Shastry model with Yangian symmetry, although we did not show this.

In the final chapters of this thesis, we dealt with spin chains that were constructed using the

QISM. The main objective was the construction of a spin chain with adjoint $\mathfrak{su}(n)$ symmetry. To get acquired with the notations and constructions, we reviewed in detail the theory of integrable spin chains giving some well-known examples along the way. The starting point is the Yang-Baxter equation, whose solutions are called R-matrices. The solutions of the Yang-Baxter equation can be found using the representation theory of the Yangian. In principle, there exists a solution of the Yang-Baxter equation for every irreducible representation of $\mathfrak{su}(n)$. The most straightforward way to find explicit expressions for those solutions is due to Chari and Pressley [10, 11] and involves the construction of an intertwining operator for the corresponding Yangian representations.

Both the R-matrix and the intertwining operator act on the tensor product of two representations, that both equaled the adjoint representation of $\mathfrak{su}(n)$ in our case. The tensor product decomposition of the adjoint representation with itself consists out of seven irreducible representations for $\mathfrak{su}(n)$ (if $n > 3$) and six for $\mathfrak{su}(3)$. In both cases, the adjoint representation itself appears twice in the decomposition, while all other submodules are different. The fact that this decomposition of the adjoint representation is not multiplicity-free, makes the structure of the R-matrix more complicated. On all the irreducible modules that appear once in the decomposition the R-matrix acts by a scalar, while the R-matrix has a structure on a (2×2)-matrix if it acts on the two copies of the adjoint representation. This is different than for the $\mathfrak{su}(2)$-invariant R-matrices, where such a multiplicity does not appear and thus the R-matrix can be written as a linear combination of mutually orthogonal projections.

To find the explicit coefficients in the R-matrices, we first determined the intertwining operator for the Yangian representations. We let the Yangian generators act on the highest weight vectors in the submodules and used the intertwining property to find all the coefficients. For $n = 3$ we explicitly showed the derivation of one of the coefficients. Using a computer program, we explicitly verified that the corresponding R-matrix satisfies the YBE. For $n > 3$ we determined these coefficients based on a set of equations that we checked numerically up to $n = 7$. We note that the solution of the Yang-Baxter equation that we found is Hermitian, with respect to the inner product on $\mathfrak{su}(n) \otimes \mathfrak{su}(n)$ for which the corresponding $SU(n)$ representation is unitary.

The QISM allows one to construct an integrable spin chain for every solution of the YBE. We reviewed how this leads to the Heisenberg and the Babujian-Takhtajan models for the spin-$\frac{1}{2}$ and spin-1 representation of $\mathfrak{su}(2)$. The R-matrix that we found for $\mathfrak{su}(n)$ leads to a non-Hermitian Hamiltonian. The non-Hermitian term arises due to the matrix structure in the R-matrix, which is a consequence of the fact that the decomposition of the adjoint representation with itself is not free of multiplicities. For more than two sites, the Hamiltonian has imaginary eigenvalues. This makes it hard to give a physical interpretation of the model, as it is impossible to identify the eigenvalues with real energies.

In the final chapter, we have studied inhomogeneous integrable spin chains. This inhomogeneiety can be achieved by putting defects on some of the sites. In the construction of the models, we replaced the R-matrix on some of the sites by a so called Lax operator that solves the Lax equation. The defects give rise to next-nearest neighbour interactions and three-site interactions around the defect. We presented two simple models with defects, with the Heisenberg model and the Babujian-Takhtajan model. The structure of both defect Hamiltonians is very similar, although the expression for the Babujian-Takhtajan model is very involved. For neighbouring defects, higher order interactions will appear, with interactions on more sites. Non-neighbouring interactions are essentially independent and one can seperately add the defect Hamiltonians to the system. We also found the Bethe equations for the $\mathfrak{su}(2)$-invariant models with defects and the eigenvalues for the Heisenberg model with defects on non-neighbouring sites.

As discussed above, the QISM does not lead to a Hermitian integrable spin chain for the adjoint representation of $\mathfrak{su}(n)$. This poses the question for which representations of $\mathfrak{su}(n)$ a

Hermitian spin chain can be constructed. In any case, the intertwining operator can always be used to find an R-matrix that is invariant under a specific representation of $\mathfrak{su}(n)$. If the decomposition of the tensor product of this representation with itself is free of any multiplicities, then the R-matrix will be a real linear combination of mutually orthogonal projectors. Therefore, the local Hamiltonian, given by the product of the derivative of the R-matrix and a permutation operator, will therefore be such a linear combination as well. Since projectors are Hermitian, the Hamiltonian will be as well. So, for those representations of $\mathfrak{su}(n)$ that have a multiplicity free decomposition a Hermitian integrable spin chain can be constructed. In particular, representations whose Young tableau is rectangular, fall into this category [23, 26]. If this is the case, then there is a shortcut to the R-matrix, which is due to Mackay [33]. This method does not involve the Yangian and the coefficients in the R-matrix can be found straight away from the Lie algebra properties of $\mathfrak{su}(n)$, although Mackay showed how his construction is related to the Yangian.

Finally, we want to note an opening for further research. In the final chapter we considered integrable spin chains with defects, in particular the Heisenberg model and the Babujian-Takthajan model with defects. It is still unclear how these systems physically differ from their homogeneous counterparts. For instance, it is not known if and how the ground state is affected by the defect. This question, along with questions on many more physical aspects of the model, still needs to be answered. A possible starting point would be a numerical study of the eigenvalues and eigenstates of the model, using the Bethe equations. Furthermore, it could be interesting to look for models with disorder, by creating many defects on the sites.

Bibliography

[1] S.I. Alihauskas and P.P. Kulish, *Spectral resolution of $SU(3)$-invariant solutions of the Yang-Baxter equation*, J. Sov. Math. **35**, 2563 (1986).

[2] N. Andrei and H. Johannesson, *Heisenberg chain with impurities (an integrable model)*, Phys. Lett. A. **100**, 108 (1984).

[3] H.M. Babujian, *Exact solution of the isotropoic Heisenberg chain with arbitrary spins: Thermodynamics of the model*, Nucl. Phys. B **215**, 317 (1982).

[4] E.P. van den Ban, *Lie groups*, Lecture Notes, Universiteit Utrecht (2010).

[5] T. Barnes, E. Dagotto, J. Riera and E.S. Swanson, *Excitation spectrum of Heisenberg spin ladders*, Phys. Rev. B, **47**, 3196 (1993).

[6] M. T. Batchelor, X.-W. Guan, N. Oelkers, Z. Tsuboi, *Integrable models and quantum spin ladders: comparison between theory and experiment for the strong coupling ladder compounds*, Advances in Physics, **56**, 465 (2007).

[7] H. Bethe, *Zur Theorie der Metalle*, Zeitschrift für Physik, **31**, 205 (1931).

[8] R.A. Bertlmann and P. Krammer, *Bloch vectors for qudits*, J. Phys. A. **41**, 235303 (2008).

[9] S.M. Carroll, *Spacetime and geometry. An introduction to general relativity*, Addison Wesley (2004).

[10] V. Chari and A. Pressley, *Yangians and R-matrices*, L'Enseign. Math. **36**, 267 (1990).

[11] —, *Fundamental representations of Yangians and singularities of R-matrices*, J. reine angew. Math. **417**, 87 (1991).

[12] —, *A Guide to Quantum Groups*, Cambridge: Cambridge Unitversity Express (1998).

[13] S. Capponi, P. Lecheminant and K. Totsuka, *Phases of one-dimensional SU(N) cold atomic Fermi gasesFrom molecular Luttinger liquids to topological phases*, Annals of Physics, **367**, 50 (2016).

[14] O.A. Castro-Alvaredo, A. Fring, *A spin chain model with non-Hermitian interaction: The Ising quantum spin chain in an imaginary field*, J. Phys A. **42**, 465211 (2009).

[15] V.G. Drinfeld, *Hopf algebras and the quantum Yang-Baxter equation*, Sov. Math. Dokl. **31**, 254 (1985).

[16] V.G. Drinfeld, *A new realization of Yangians and quantized affine algebras*, Sov. Math. Dokl. **36**, 212 (1988).

[17] H.A. Dye, *Unitary Solutions to the Yang-Baxter Equation in Dimension Four*, Quantum Information Processing **2**, 117 (2004).

[18] F.H.L. Essler, H. Frahm, F. Göhmann, A. Klümper, V.E. Korepin, *The One-Dimensional Hubbard Model*, Cambridge (2003).

[19] L.D. Faddeev, *How Algebraic Bethe Ansatz works for integrable model*, arXiv:hep-th/9605187, (1996).

[20] L. Ferro, *Yangian Symmetry in N=4 super Yang-Mills*, arXiv:hep-th/1107.1176, (2011).

[21] F.D.M. Haldane, *Exact Jastrov-Gutzwiller resonant-valance-bond ground state of the spin-$\frac{1}{2}$ antiferromagnetic Heisenberg chain with $1/r^2$ exchange*, Phys. Rev. Lett. **60**, 635 (1988).

[22] F.D.M. Haldane, Z.N.C. Ha, J.V. Talstra, D. Bernard and V. Pasquier, *Yangian symmetry of integrable quantum chains with long-range interactions and a new description of states in conformal field theory*, Phys. Rev. Lett. **66**, 1529 (1991).

[23] M. Hamermesh, *Group Theory and Its Application to Physical Problems*, Dover Books on Physics (1962).

[24] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics, Springer (1972).

[25] L.H. Kauffmann, S.J. Lomonaco Jr, *Braiding Operators are Universal Quantum Gates*, New J. Phys. **6**, 134 (2004).

[26] T. Kobayashi, *Multiplicity-free theorems of the Restrictions of Unitary Highest Weight Modules with respect to Reductive Symmetric Pairs*, Progress in Mathematics, Springer (2006).

[27] P.P. Kulish, N.Y. Reshetikin and E.K. Sklyanin, *Yang-Baxter equation and representation theory: I*, Lett. Math. Phys. **5** 393 (1981).

[28] P.P. Kulish and E.K. Sklyanin, *Solutions of the Yang-Baxter equation*, J. Sov. Math. **19**, 1596 (1982).

[29] P.P. Kulish and N.Y. Reshetikin, *Gl_3-invariant solutions of the Yang-Baxter equation and associated quantum systems*, J. Sov. Math. **34**, 1948 (1986).

[30] R.B. Laughlin, D. Giuliano, R. Caracciolo and L. White, *Quantum Number Fractionalization in Antiferromagnets*, Springer Series in Solid-State Sciences, (2008).

[31] A.J. MacFarlane, A. Sudbery and P.H. Weisz, *On Gell-Mann's λ-Matrices, d- and f-tensors, Octets, and Parametrizations of $SU(3)$*, Commun. Math. Phys. **11**, 77 (1968).

[32] A.J. MacFarlane, H. Pfeiffer, *On characteristic equations, trace identities and Casimir operators of simple Lie algebras*, J. Math. Phys **41**, 3192 (2000).

[33] N.J. Mackay, *Rational R-matrices in irreducible representations*, Jour. Phys. A., **24**, 4017 (1991).

[34] M. Mourigal, M. Enderle, A. Klöpperpieper, J.-S. Caux, A. Stunault and H.M. Rønnow, *Fractional spinon excitations in the quantum Heisenberg antiferromagnetic chain*, Nature Physics **9**, 435 (2013).

[35] A. Mostafazadeh, *Is Pseudo-Hermitian Quantum Mechanics an Indefinite-Metric Quantum Theory?*, Czech. J. Phys. **53**, 1079 (2003).

[36] G. Mussardo, *Statistical Field Theory*, Oxford: Oxford University Press (2010).

[37] S. Lang, *Complex Analysis*, Springer, New York (1985).

[38] N. Reshetikin, *Integrable models of quantum one-dimensional magnets with $O(N)$ and $Sp(2K)$ symmetry*, Theor. Math. Phys. **63**, 1596 (1985).

[39] A. Rocén, *Yangians and their representations*, Thesis, Department of Mathematics, University of York (2010).

[40] A. Roy and T. Quella, *Chiral Haldane phases of SU(N) quantum spin chains in the adjoint representation*, arXiv:cond:mat/1512.05229 [cond-mat.str-el], (2015).

[41] L. Samaj, Z. Bajnok. *Introduction to the Statistical Physics of Integrable Many-body Systems*, Cambridge (2013).

[42] P. Schmitteckert, P. Schwab and U. Eckern, *Quantum coherence in an exactly solvable one-dimensional model with defects*, Europhys. Lett. **30**, 543 (1995).

[43] D. Schuricht, *The Haldane-Shastry model, its Yangian, and the dynamical spin correlations*, Diploma thesis, Fakultät für Physik, Universität Karlsruhe (2003).

[44] —, *Fraktionale Quantiserung und Yangian Symmetrie in SU(n) Spinketten*, PhD thesis, Fakultät für Physik, Universität Karlsruhe (2006).

[45] B.S. Shastry, *Exact solution of an $S = \frac{1}{2}$ Heisenberg antiferromagnetic chain with long-ranged interactions*, Phys. Rev. Lett. **60**, 639 (1988).

[46] L. Stronks, J. van de Leur, D. Schuricht, *On rational R-matrices with adjoint SU(n) symmetry*, arXiv:1606:02516, (2016).

[47] L.A. Takhtajan, sl The picture of low-lying excitations in the isotropic Heisenberg chain of arbitrary spins, Phys. Lett. A. **87**, 479 (1982).

[48] D.A. Tennant, R.A. Cowley, S.E. Nagler and A.M. Tsvelik, *Measurement of the spin-excitation continuum in one-dimensional KCuF3 using neutron scattering*, Phys. Rev. B. **52**, 13368 (1995).

[49] H.J. de Vega and F Woynarovich, *New integrable quantum chains combining different kinds of spins*, J. Phys. A. Math. Gen. **25**, 4499 (1992).

[50] X.G. Wen, F. Wilczek, A. Zee, *Chiral spin states and superconductivity*, Phys. Rev. B **39**, 11413 (1989).

[51] A. Zamolodchikov and V. Fateev, sl A model factorized S-matrix and an integrable spin-1 Heisenberg chain, Sov. J. Nucl. Phys. **32**, 298 (1980).