Bethe Ansatz Techniques and the AdS/CFT Correspondence

Johnson Leow

Abstract

In the present thesis, we will review various Bethe ansatz techniques such as the (nested) algebraic Bethe ansatz and the (nested) coordinate Bethe ansatz using well known models like the XXX model and the Hubbard model. Subsequently, the Bethe ansatz techniques are applied to obtain Bethe ansatz equations and Bethe vectors of the $\mathfrak{su}(2|2)$-invariant S-matrix. Finally, these results are used to obtain the full S-matrix and Bethe equations of the string sigma model.

Acknowledgements

Contents

Chapter 1

Introduction

In 1931, Hans Bethe constructed a method, which is nowadays called the "Coordinate Bethe ansatz" (CBA), for calculating the exact eigenvalues and eigenvectors of the one-dimensional spin-1/2 XXX Heisenberg model [7, 23]. The term "Bethe ansatz", refers to the wave function Bethe had used for the eigenvectors. Since then, many other quantum many body systems have been solved by some variant of the Bethe ansatz. One problem especially worth mentioning is the repulsive δ interaction problem which has been solved by C. N. Yang in the late sixties [27, 16]. For this he used the Bethe ansatz twice; the second time in a generalized form which is often called the "Bethe-Yang" ansatz. Following this result, the repulsive δ interaction problem with an arbitrary irreducible representation of the permutation group was solved by B. Sutherland by repeated use of the Bethe-Yang ansatz [25]. This method is nowadays called the "nested (coordinate) Bethe ansatz".

Around the same time when Yang published his article on the δ interaction problem, the classical inverse scattering method (CISM) was invented by Gardner, Green, Kruskal and Miura after studying the Korteweg-deVries equation [17]. In the following years, this method matured [15] and people tried to quantize the theory. The quantum version of the CISM was obtained by the Leningrad school around the beginning of the eighties and it resulted in what we nowadays call the "Quantum Inverse Scattering Method" (QISM) [21]. The QISM was very powerful and it was applied to all kinds of models. One of those models was the XXX model and the results it obtained was exactly the same as with Bethe's CBA. Therefore, the QISM is also called the "Algebraic Bethe Ansatz" (ABA). Because of the strong algebraic framework, the QISM was also used to prove the integrability of the XXX model. Furthermore, the development of the QISM also led to the introduction of mathematical objects like quantum groups [10] which has intimate connections with areas of mathematics like topology, non-commutative geometry, representation theory, differential geometry and with areas of physics like conformal field theory, quantum field theory and string theory. So even though the focus in this thesis will be on the application of the Bethe ansatz, it is important to note that the underlying mathematical structure of the Bethe ansatz is a fascinating subject itself which intertwines areas of mathematics and physics.

Although the Bethe ansatz technique arose as a method to solve problems in condensed matter physics, it has been steadily gaining attention from string theorists as a tool to study

the AdS/CFT correspondence.

In the late nineties, Maldacena proposed a duality, which is nowadays called the "AdS/CFT correspondence". One of its conjectures is that the type IIB superstring theory on the curved background $AdS_5 \times S^5$ should be equivalent to $\mathcal{N} = 4$ SYM gauge theory [1] and this is supported by the fact that the superconformal algebra of both theories is $\mathfrak{psu}(2, 2|4)$.

To prove this conjecture however, is not so easy. For example, the AdS/CFT correspondence predicts that the spectrum of scaling dimensions of the gauge theory should coincide with the spectrum of the energies of the string states. But the calculation of both of these quantities has proven to be problematic. Furthermore, the AdS/CFT correspondence is a duality of the strong/weak type: the weak coupling regime of the gauge theory maps to the strong coupling regime of the string theory, and vice versa. Now, since it is not known how to access the strong coupling regime in either theory, a test of the AdS/CFT correspondence is rather troublesome.

However, progress was made, when Minahan and Zarembo [24] discovered that the gauge invariant operators in the scalar field sector of the gauge theory were, at the planar one-loop level, isomorphic with translation invariant eigenvectors of an integrable $\mathfrak{so}(6)$ quantum spin chain. There, the spin chain Hamiltonian corresponded to the gauge theory one-loop dilatation operator, and it was diagonalized using the Bethe ansatz technique. Inspired by the work of Minahan and Zarembo, integrable structures has subsequently been found in various sectors and various loops of the planar gauge theory, and the corresponding Bethe ansatz equations have also been derived. Furthermore, all loop Bethe equations have been conjectured for certain asymptotic limits.

Since planar gauge theory possesses integrable structures, the corresponding string sigma model in the infinite volume limit on the string theory side should also possess integrable structures. However, calculations on string side do not go as smoothly as the gauge side, but nonetheless, integrable structures were discovered for classical strings and it is believed that integrability also survives quantum corrections. However, showing that such integrable structures indeed exist in the quantum corrections has proven to be a very difficult problem.

In fact, if both theories are completely integrable, then constructing the dilatation operator loop by loop or quantizing the string sigma model would be an impossible task. Therefore, it is not a bad idea to just simply assume integrability for the gauge and string theory and subsequently use the symmetry algebra to derive the S-matrix and the corresponding Bethe equations. In this thesis, we will do exactly this for the string sigma model in the light cone gauge.

This thesis is organized as follows. In chapter two we will introduce the various Bethe ansatz techniques like the nested (coordinate) Bethe ansatz and the nested (algebraic) Bethe ansatz using the spin-1/2 XXX Heisenberg model and the Hubbard model. Furthermore, we will use the underlying algebraic structure of the algebraic Bethe ansatz to prove the integrability of the XXX model.

In chapter three we will make an introduction to Lie superalgebras. The basics of Lie superalgebra theory and its representation theory will be treated and we will also describe

the $\mathfrak{psu}(2,2|4)$ Lie superalgebra which is of physical importance in the AdS/CFT correspondence. Note that for this chapter it is assumed that the reader has some basic knowledge of Lie algebras. A brief summary of Lie algebra theory can be found in appendix B.

In Chapter four we will describe the $\mathfrak{su}(2|2)$ Lie superalgebra and we will apply the Bethe ansatz machinery to obtain the Bethe ansatz equations and Bethe vectors of the (extended) $\mathfrak{su}(2|2)$-invariant S-matrix. The results are subsequently used to derive the full Bethe equations of the string sigma model.

In the final chapter we will summarize our results and state our conclusions.

Chapter 2

The Bethe Ansatz Method

In this chapter, we will introduce the coordinate Bethe ansatz by solving the one-dimensional spin-1/2 XXX Heisenberg model following the lines of [7, 18, 20]. Subsequently, we will prove the integrability of this model by solving it using the algebraic Bethe ansatz [8, 13, 14, 21]. We then move on to the nested coordinate Bethe ansatz [12, 16, 25, 26, 27], the nested algebraic Bethe ansatz [8, 11] and the graded (nested) algebraic Bethe ansatz [11, 12, 22]. This last technique will be illustrated using the Hubbard model and the results will be used in chapter 4. Finally, we will end this chapter with some remarks on the validity of the Bethe ansatz and how it is connected with S-matrices and integrable system. But before we continue, let us introduce some notations and definitions first:

- A spin chain of length N is a ordered set of points, also called lattice sites, labeled by integers n. We will take the possible values for n to be $n = 1, \ldots, N$. A spin chain of length N with periodic boundary conditions is defined as a spin chain of length N with the additional identification $n \equiv n + N$.

- Let $\mathcal{A}'$ be an algebra generated by elements X^α, where α assumes some finite number of values. Then the induced algebra $\mathcal{A}$ on a spin chain of length N will be defined as the algebra generated by elements X_n^α, where X_n^α is X^α associated with lattice site n, together with a set of commutation relations between X_n^α. These relations will be called *ultralocal* when X_n^α and X_m^β commute for $m \neq n$.

  If $\mathcal{A}$ is an ultralocal algebra, then the Hilbert space h for a representation π of $\mathcal{A}$ has the natural tensor product form

  $$h = \bigotimes_{n=1}^{N} h_n = h_1 \otimes h_2 \otimes \ldots \otimes h_N \tag{2.0.1}$$

  where each h_n is the Hilbert space for a representation π_n of $\mathcal{A}'$, and we can identify h_n with lattice site n. The elements X_n^α will be represented as

  $$\pi(X_n^\alpha) = \mathbb{I} \otimes \mathbb{I} \otimes \ldots \otimes \underset{\text{n-th place}}{\pi_n(X^\alpha)} \otimes \ldots \otimes \mathbb{I} \tag{2.0.2}$$

4

with $\mathbb{I}$ being the identity operator. So X_n^α will only act nontrivially on h_n:

$$\pi(X_n^\alpha)h = h_1 \otimes h_2 \otimes \ldots \otimes \pi_n(X^\alpha)h_n \otimes \ldots \otimes h_N \qquad (2.0.3)$$

To simplify the notation a bit, we will omit the symbol π in the rest of this thesis when there is no chance for confusion. Furthermore, suppose that $\mathcal{A}$ is an ultralocal algebra with representation space V and let W be a finite dimensional Hilbert space. We will now adopt the convention that unless stated otherwise, $\mathcal{A}$ will act on $V \otimes W$ with the canonically induced action $\mathcal{A} \otimes I_W$, where I_W is the identity operator on W. This will have the following impact on our notation. First of all, note that we can view $X_n^\alpha : h \to h$ as the induced action of $X_n^\alpha : h_n \to h_n$. As a second example, let $A : V \to V$ and $B : W \to W$ be operators acting on a finite dimensional vector space. Then the induced action for A on $V \otimes W$ is $A \otimes \mathbb{I}_W$ and it will be denoted again with A (we have something similar for B). Consequently, $(A \otimes B) = (A \otimes \mathbb{I}_W)(\mathbb{I}_V \otimes B)$ acts on $V \otimes W$ and it will be denoted by AB.

Finally, we have to warn the reader that these definitions are only valid when $\mathcal{A}'$ is a non-graded algebra. For a graded algebra we refer the reader to appendix A.4.

- The spin algebra is generated by the spin variables S_n^α, $\alpha = x, y, z$ with commutation relations

$$[S_m^\alpha, S_n^\beta] = i\hbar \sum_\gamma \varepsilon_{\alpha\beta\gamma} S_n^\gamma \delta_{mn} \qquad (2.0.4)$$

where $\varepsilon_{\alpha\beta\gamma}$ is a completely antisymmetric tensor with $\varepsilon_{xyz} = 1$. Here, the commutator $[\cdot, \cdot]$ is defined by $[a, b] = ab - ba$. Notice that the spin variables form the Lie algebra $\mathfrak{su}(2)$. Therefore, the finite dimensional representations of the spin algebra are labeled by half-integers $s = 0, 1/2, 1, \ldots$, and they are realized in $\mathbb{C}^{2s+1}$. The representation which we will be the most interested in, is the $s = 1/2$ representation. In this representation the spin variables S_n^α are represented as $S_n^\alpha = \hbar/2\sigma^\alpha$, with σ^α being the Pauli matrices which are defined as

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (2.0.5)$$

For simplicity, we will take $\hbar = 1$ in the remainder of this thesis.

- We will denote the permutation group over labels $\{1, 2, \ldots, r\}$ by S_r. A permutation $P \in S_r$ is a cycle and its action on an element $i \in \{1, 2, \ldots, r\}$ will be denoted by $P(i)$. For example, if $r = 4$ and $P = (1234)$, then $P(3) = 4$.

  For transpositions, we will use a slightly different notation. The transposition (ij) will be denoted by $P_{i,j}$. So, for example, $P_{1,3}(1) = 3$.

With these definitions in mind, we will start with the description of the XXX Heisenberg model.

2.1 The XXX Heisenberg Model

The one-dimensional spin-1/2 Heisenberg model is a periodic spin chain of length N with Hamiltonian

$$H = -\sum_{n=1}^{N} \left(J_x S_n^x S_{n+1}^x + J_y S_n^y S_{n+1}^y + J_z S_n^z S_{n+1}^z \right) \tag{2.1.1}$$

where S_n^α are spin variables in the $s = 1/2$ representation with the periodic boundary conditions $S_{N+1}^\alpha = S_1^\alpha$. When the real constants J_x, J_y and J_z are all different we will call the model the XYZ Heisenberg model. The other two cases are $J_x = J_y \neq J_z$ and $J_x = J_y = J_z$, and they are called the XXZ and XXX model respectively. In this section, we will consider the XXX model with $J = J_x = J_y = J_z$. In this case, the Hamiltonian simplifies to

$$H = -J \sum_{n=1}^{N} \left(S_n^x S_{n+1}^x + S_n^y S_{n+1}^y + S_n^z S_{n+1}^z \right) \tag{2.1.2}$$

Let us now define the raising and lowering operators $S_n^\pm \equiv S_n^x \pm i S_n^y$. They satisfy the commutation relations

$$[S_n^z, S_m^z] = [S_n^\pm, S_m^\pm] = 0 \,, \quad [S_n^z, S_m^\pm] = \pm S_n^\pm \delta_{nm} \,, \quad [S_n^+, S_m^-] = 2 S_n^z \delta_{nm} \tag{2.1.3}$$

With these operators, (2.1.2) can be written as

$$H = -J \sum_{n=1}^{N} \left(\frac{1}{2} \left(S_n^+ S_{n+1}^- + S_n^- S_{n+1}^+ \right) + S_n^z S_{n+1}^z \right) \tag{2.1.4}$$

Here, H acts on the 2^N-dimensional Hilbert space

$$h = \bigotimes_{n=1}^{N} h_n = h_1 \otimes h_2 \otimes \ldots \otimes h_N \quad \text{with} \quad h_1 = h_2 = \ldots = h_N = \mathbb{C}^2 \tag{2.1.5}$$

The Hilbert space $\mathbb{C}^2$ is two-dimensional vectorspace with orthogonal basis vectors

$$\mid \uparrow \rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{(spin up)} \quad \text{and} \quad \mid \downarrow \rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{(spin down)} \tag{2.1.6}$$

and it is equipped with the canonical hermitian inner product $\langle \cdot \,, \cdot \rangle$ defined by:

$$\langle \uparrow \mid \uparrow \rangle = \langle \downarrow \mid \downarrow \rangle = 1 \,, \qquad \langle \uparrow \mid \downarrow \rangle = \langle \uparrow \mid \downarrow \rangle = 0 \tag{2.1.7}$$

The action of the spin operators on these basis vectors is given by

$$\begin{aligned} S^+ \mid \uparrow \rangle &= 0 \,, & S^+ \mid \downarrow \rangle &= \mid \uparrow \rangle \,, & S^z \mid \uparrow \rangle &= \tfrac{1}{2} \mid \uparrow \rangle \\ S^- \mid \downarrow \rangle &= 0 \,, & S^- \mid \uparrow \rangle &= \mid \downarrow \rangle \,, & S^z \mid \downarrow \rangle &= -\tfrac{1}{2} \mid \downarrow \rangle \end{aligned} \tag{2.1.8}$$

Consequently, h will be spanned by the orthogonal basis vectors

$$|\boldsymbol{\sigma}\rangle \equiv |\sigma_1 \ldots \sigma_N\rangle \equiv \bigotimes_{n=1}^{N} |\sigma_n\rangle = |\sigma_1\rangle \otimes |\sigma_2\rangle \otimes \ldots \otimes |\sigma_N\rangle \quad \text{with} \ \sigma_n \in \{\uparrow, \downarrow\} \qquad (2.1.9)$$

and it is equipped with the (induced) hermitian inner product $\langle \cdot, \cdot \rangle$ defined by:

$$\langle \boldsymbol{\sigma}'|\boldsymbol{\sigma}\rangle = \langle \sigma_1|\sigma_1'\rangle \ldots \langle \sigma_N|\sigma_N'\rangle \qquad (2.1.10)$$

The canonically induced actions of $S_n^{\pm}$ and S_n^z on h are given by

$$\begin{aligned}
S_k^+|.. \uparrow_k ..\rangle &= 0, & S_k^+|.. \downarrow_k ..\rangle &= |.. \uparrow_k ..\rangle, & S_k^z|.. \uparrow_k ..\rangle &= \tfrac{1}{2}|.. \uparrow_k ..\rangle \\
S_k^-|.. \downarrow_k ..\rangle &= 0, & S_k^-|.. \uparrow_k ..\rangle &= |.. \downarrow_k ..\rangle, & S_k^z|.. \downarrow_k ..\rangle &= -\tfrac{1}{2}|.. \downarrow_k ..\rangle
\end{aligned} \qquad (2.1.11)$$

where $|.. \uparrow_k ..\rangle$ is the shorthand notation for $|\sigma_1 \ldots \sigma_N\rangle$ with $\sigma_k = \uparrow$. Notice that with these definitions, the Hamiltonian can be written as a $2^N \times 2^N$ real symmetric matrix which means that it has a complete orthogonal system of eigenvectors.

2.2 The Coordinate Bethe Ansatz (CBA)

Before we start constructing eigenvectors for H, let us remark that for the total spin operator $S_T^z \equiv \sum_{m=1}^{N} S_m^z$ we have

$$[H, S_T^z] = 0, \qquad S_T^z|\sigma_1 \ldots \sigma_N\rangle = \left(\frac{1}{2}N - K\right)|\sigma_1 \ldots \sigma_N\rangle \qquad (2.2.1)$$

with K the number of down spins of $|\sigma_1 \ldots \sigma_N\rangle$. This means that H acting on $|\sigma_1 \ldots \sigma_N\rangle$ yields a linear combination of basis vectors where each basis vector has the same number of down spins as $|\sigma_1 \ldots \sigma_N\rangle$. Therefore, we can block diagonalize H just by sorting the basis vectors according to their eigenvalue of S_T^z. The diagonalization of H is hereby reduced to the diagonalization of each of the blocks.

Let $K \leq N$ be an integer and let the K-subspace be the subspace of basis vectors with K down spins. If $K = 0$ we will have a subspace consisting of the single vector $|F\rangle \equiv |\uparrow \ldots \uparrow\rangle$. Notice that this vector is an eigenvector of H with eigenvalue $E_0 \equiv -JN/4$:

$$H|F\rangle = -J \sum_{n=1}^{N} S_n^z S_{n+1}^z |\uparrow \ldots \uparrow\rangle = -\frac{JN}{4}|\uparrow \ldots \uparrow\rangle \qquad (2.2.2)$$

Now let K be arbitrary. Since the K-subspace is $N!/[(N-K)!K!]$ dimensional, one should find the same number of eigenvectors of H in this subspace. So let us expand the eigenvectors in the form

$$|\psi\rangle = \sum_{\{1 \leq n_i \leq N\}} \tilde{a}(\mathbf{n})|n_1, \ldots, n_K\rangle \quad \text{where} \ \tilde{a}(\mathbf{n}) \equiv \tilde{a}(n_1, \ldots, n_K) \qquad (2.2.3)$$

with

$$|n_1, \ldots, n_K\rangle \equiv S^-_{n_1} S^-_{n_2} \ldots S^-_{n_K} |F\rangle \tag{2.2.4}$$

Furthermore, we will require $\tilde{a}(\mathbf{n})$ to satisfy the Pauli exclusion principle;

$$\tilde{a}(\mathbf{n}_R) = \text{sign}(R)\tilde{a}(\mathbf{n}) \quad \text{for every } R \in S_K \quad \text{where} \quad \tilde{a}(\mathbf{n}_R) \equiv \tilde{a}(n_{R(1)}, \ldots, n_{R(K)}) \tag{2.2.5}$$

Therefore it is enough to treat the eigenvalue problem for the situation that

$$\mathbf{n} \in I \quad \text{where} \quad I \equiv \{1 \leq n_1 < n_2 \ldots < n_K \leq N\} \tag{2.2.6}$$

In this case, we will denote the wave function by $a(\mathbf{n})$, i.e. $a(\mathbf{n}) \equiv \tilde{a}(\mathbf{n})|_I$, and we will also require $a(\mathbf{n})$ to satisfy the periodic boundary condition

$$a(n_1, \ldots, n_K) = a(n_2, \ldots, n_K, n_1 + N) \tag{2.2.7}$$

To determine these wave functions we will use the "coordinate Bethe ansatz" which postulates that

$$a(n_1, \ldots, n_K) = \sum_{P \in S_K} A_P \exp\left(i \sum_{j=1}^{K} k_{P(j)} n_j\right) \tag{2.2.8}$$

Here, the k_j's are pseudo-momenta (or wavenumbers) introduced for each of the K down spins, and S_K is the permutation group over labels $\{1, 2, \ldots, K\}$. So physically, we can view the j-th down spin as a particle with momentum k_j and strictly speaking, we should write $|n_1(p_1), \ldots, n_K(p_K)\rangle$ instead of $|n_1, \ldots, n_K\rangle$ to emphasize this. However, when there is no chance for confusion we will use the latter notation where it is understood that $n_j \equiv n_j(p_j)$.

What remains to do is to find the coefficients A_P and all possible sets of allowed pseudo-momenta $\{k_j\}$ using the eigenvalue equation for H and the periodic boundary conditions for $a(n_1, \ldots, n_K)$. Let us examine this procedure for the $K = 1$ and $K = 2$ cases first.

For the $K = 1$ case, the eigenvectors and coefficients are given by

$$|\psi\rangle = \sum_{n=1}^{N} a(n)|n\rangle \quad \text{and} \quad a(n) = Ae^{ikn} \tag{2.2.9}$$

The periodic boundary condition $a(n+N) = a(n)$ will fix the value of the pseudo-momentum k to be $k = 2\pi m/N$ with $m = 0, 1, \ldots, N - 1$. Furthermore, using the eigenvalue equation $E|\psi\rangle = H|\psi\rangle$ and the identification $|N + 1\rangle = |1\rangle$ we get

$$
\begin{aligned}
E \sum_{n=1}^{N} a(n)|n\rangle &= -J \sum_{n=1}^{N} a(n) \sum_{m=1}^{N} \left(\frac{1}{2}\left(S^+_m S^-_{m+1} + S^-_m S^+_{m+1}\right) + S^z_m S^z_{m+1}\right)|n\rangle \\
&= -J \sum_{n=1}^{N} a(n)\left(\frac{1}{2}|n + 1\rangle + \frac{1}{2}|n - 1\rangle + \left\{\frac{N}{4} - 1\right\}|n\rangle\right) \tag{2.2.10} \\
&= -J \sum_{n=1}^{N} \left(\frac{1}{2}a(n - 1) + \frac{1}{2}a(n + 1) + \left\{\frac{N}{4} - 1\right\}a(n)\right)|n\rangle
\end{aligned}
$$

From the above equation we obtain the relation

$$Ea(n) = -\frac{1}{2}Ja(n-1) - \frac{1}{2}Ja(n+1) + (E_0 + J)\,a(n) \tag{2.2.11}$$

which, after plugging in the expression for $a(n)$, gives the following expression for the eigenvalue:

$$E - E_0 = J(1 - \cos k) \tag{2.2.12}$$

This solves the $K = 1$ case.

When $K = 2$ however, it becomes a bit more complicated. In this case we have

$$|\psi\rangle = \sum_{1 \leq n_1 < n_2 \leq N} a(n_1, n_2)|n_1, n_2\rangle \quad \text{with} \quad a(n_1, n_2) = Ae^{i(k_1 n_1 + k_2 n_2)} + Be^{i(k_2 n_1 + k_1 n_2)} \tag{2.2.13}$$

The eigenvalue equation for H imposes conditions on $a(n_1, n_2)$ analogous to the $K = 1$ case with the only difference that we have to be a little careful in the situation when two down spins are sitting next to each other; if $n_2 > n_1 + 1$ we have

$$Ea(n_1, n_2) = -J\left(\frac{1}{2}\sum_{\sigma = \pm 1}\left\{a(n_1 + \sigma, n_2) + a(n_1, n_2 + \sigma)\right\} + \left\{\frac{N}{4} - 2\right\}a(n_1, n_2)\right)$$
$$\Leftrightarrow 2(E - E_0)a(n_1, n_2) = J\left(4a(n_1, n_2) - \sum_{\sigma = \pm 1}\left\{a(n_1 + \sigma, n_2) + a(n_1, n_2 + \sigma)\right\}\right) \tag{2.2.14}$$

and if $n_2 = n_1 + 1$ we have

$$2(E - E_0)a(n_1, n_2) = J\left(2a(n_1, n_2) - a(n_1 - 1, n_2) - a(n_1, n_2 + 1)\right) \tag{2.2.15}$$

Plugging in the expression for $a(n_1, n_2)$ into (2.2.14) gives the following expression for the eigenvalue:

$$E - E_0 = J(1 - \cos k_1 - \cos k_2) \tag{2.2.16}$$

Let us now extend the validity of $a(n_1, n_2)$ to the region $1 \leq n_1 \leq n_2 \leq N$. Then, because $a(n_1, n_2)$ is a smooth function, the equations (2.2.14) and (2.2.15) should agree with each other when $n_2 = n_1 + 1$. So subtracting (2.2.15) from (2.2.14) with $n_2 = n_1 + 1$ will give us the auxiliary condition

$$2a(n_1, n_1 + 1) = a(n_1, n_1) + a(n_1 + 1, n_1 + 1) \tag{2.2.17}$$

If we now plug in the expression for $a(n_1, n_2)$ we will obtain the following condition for the amplitude ratio:

$$\frac{B}{A} \equiv e^{-i\theta} = \frac{e^{i(k_1 + k_2)} + 1 - 2e^{ik_2}}{e^{i(k_1 + k_2)} + 1 - 2e^{ik_1}} \tag{2.2.18}$$

The ratio B/A will also be referred to as the S-matrix of the XXX model and it will be denoted by $S(k_1, k_2)$[1]. Notice that this definition implies that

$$\frac{A}{B} = \left(S(k_1, k_2)\right)^{-1} = S(k_2, k_1) \tag{2.2.19}$$

Finally, we have to impose the periodic boundary conditions on $a(n_1, n_2)$. This gives

$$Ae^{i(k_1 n_1 + k_2 n_2)} + Be^{i(k_1 n_2 + k_2 n_1)} = Ae^{i(k_1 n_2 + k_2 n_1)}e^{ik_2 N} + Be^{i(k_1 n_1 + k_2 n_2)}e^{ik_1 N} \tag{2.2.20}$$

which implies

$$e^{ik_1 N} = \frac{A}{B} = e^{i\theta} \quad \text{and} \quad e^{ik_2 N} = \frac{B}{A} = e^{-i\theta} \tag{2.2.21}$$

These last two equations are called the "Bethe (ansatz) equations". What remains to be done is to solve these equations which is an industry itself. Therefore we shall refrain from making an extensive analysis of solving the Bethe equations and only make some short comments on it.

First of all, taking the complex logarithm from (2.2.21) will give the equation

$$Nk_1 = 2\pi\lambda_1 + \theta, \qquad Nk_2 = 2\pi\lambda_2 - \theta \tag{2.2.22}$$

where the integers $\lambda_i \in \{0, 1, \ldots, N-1\}$ are called "Bethe quantum numbers". So the remaining task is to find all pairs (λ_1, λ_2) which yield solutions of equations (2.2.18) and (2.2.22). Note that we can restrict ourselves to pairs with $0 \leq \lambda_1 \leq \lambda_2 \leq N-1$ since switching λ_1 with λ_2 and letting $\theta \to -\theta$ produces the same solution.

The simplest solutions are the pairs for which one of the Bethe quantum numbers is zero; $\lambda_1 = 0$, $\lambda_2 = 0, 1, \ldots, N-1$. For these pairs we obtain $k_1 = 0$, $k_2 = 2\pi\lambda_2/N$, $\theta = 0$. More complicated solutions can be obtained by considering nonzero λ_1 and λ_2. The actual calculations will involve numerical techniques, but we will not go into this matter any further.

With the $K = 2$ case in mind, we can generalize the whole procedure to arbitrary K. First of all, the eigenvalue equation for H will impose conditions on $a(\mathbf{n}) \equiv a(n_1, \ldots, n_K)$. In the case that $n_{i+1} > n_i + 1$ for all i, this condition is

$$2[E - E_0]a(\mathbf{n}) = 2Jra(\mathbf{n}) - J\sum_{i=1}^{K}\sum_{\sigma=\pm 1} a(n_1, \ldots, n_i + \sigma, \ldots, n_K) \tag{2.2.23}$$

But if there is a j such that $n_{j+1} = n_j + 1$ and $n_{i+1} > n_i + 1$ for all $i \neq j$, this condition will be

$$2[E - E_0]a(\mathbf{n}) = 2J(K-1)a(\mathbf{n}) - J\sum_{i\neq j, j+1}^{K}\sum_{\sigma=\pm 1} a(n_1, \ldots, n_i + \sigma, \ldots, n_K) \tag{2.2.24}$$

$$- J\left(a(n_1, \ldots, n_j - 1, n_{j+1}, \ldots, n_K) + a(n_1, \ldots, n_j, n_{j+1} + 1, \ldots, n_K)\right)$$

---

[1]Note that in some literature the S-matrix is defined as $S(k_2, k_1)$ instead of $S(k_1, k_2)$

Plugging in the expression for $a(\mathbf{n})$ into (2.2.23) gives an expression for the eigenvalue:

$$E - E_0 = J \sum_{i=1}^{K} (1 - \cos k_i) \tag{2.2.25}$$

As in the $K = 2$ case, we can extend $a(\mathbf{n})$ to the region $1 \le n_1 \le n_2 \le \ldots \le n_K \le N$, since it is a smooth function. If we then take $n_{j+1} = n_j + 1$, $n_{i+1} > n_i + 1$ for all $i \ne j$ in (2.2.23) and subtract (2.2.24) from it we will get the auxiliary condition

$$2a(n_1, \ldots, n_j, n_j + 1, \ldots, n_K) = a(n_1, \ldots, n_j + 1, n_j + 1, \ldots, n_K) + a(n_1, \ldots, n_j, n_j, \ldots, n_K) \tag{2.2.26}$$

We will try to satisfy this equation by considering the following pair of terms in the expression for $a(\mathbf{n})$:

$$A_P \exp\left(i \sum_{i=1}^{K} k_{P(i)} n_i \right) \quad \text{and} \quad A_{P'} \exp\left(i \sum_{i=1}^{K} k_{P'(i)} n_i \right) \tag{2.2.27}$$

Here, P' is equal to the permutation P with $P(j)$ and $P(j+1)$ interchanged, i.e. $P' = P P_{j,j+1}$. For such a pair, equation (2.2.26) becomes

$$2A_P \exp\left(ik_{P(j)} n_j + ik_{P(j+1)}(n_j + 1) \right) + 2A_{P'} \exp\left(ik_{P'(j)} n_j + ik_{P'(j+1)}(n_j + 1) \right) =$$
$$A_P \exp\left(ik_{P(j)} n_j + ik_{P(j+1)} n_j \right) + A_P \exp\left(ik_{P(j)}(n_j + 1) + ik_{P(j+1)}(n_j + 1) \right)$$
$$+ A_{P'} \exp\left(ik_{P'(j)} n_j + ik_{P'(j+1)} n_j \right) + A_{P'} \exp\left(ik_{P'(j)}(n_j + 1) + ik_{P'(j+1)}(n_j + 1) \right) \tag{2.2.28}$$

If we now use that $P'(j) = P(j + 1)$ and $P'(j + 1) = P(j)$ we will obtain the following relation:

$$\frac{A_{P'}}{A_P} \equiv e^{-i\theta_{P(j),P(j+1)}} = -\frac{e^{i(k_{P(j)} + k_{P(j+1)})} + 1 - 2e^{ik_{P(j+1)}}}{e^{i(k_{P(j)} + k_{P(j+1)})} + 1 - 2e^{ik_{P(j)}}} \tag{2.2.29}$$

Reminding ourselves the $K = 2$ case, we recognize $A_{P'}/A_P$ as the S-matrix $S(k_{P(j)}, k_{P(j+1)})$.

Let us note (2.2.26) is the only constraint on $a(\mathbf{n})$; we will not get any new constraint equations by considering other particle configurations. For example, suppose that there is a j such that $n_{j+2} = n_{j+1} + 1 = n_j + 2$ and $n_{i+1} > n_i + 1$ for all $i \ne j, j + 1$. Then the condition on $a(\mathbf{n})$ will be

$$2[E - E_0]a(\mathbf{n}) = 2J(r - 2)a(\mathbf{n}) - J \sum_{i \ne j, j+1, j+2} \sum_{\sigma = \pm 1} a(n_1, \ldots, n_i + \sigma, \ldots, n_K) \tag{2.2.30}$$
$$- J\left(a(n_1, \ldots, n_j - 1, n_{j+1}, n_{j+2} \ldots, n_K) + a(n_1, \ldots, n_j, n_{j+1}, n_{j+2} + 1, \ldots, n_K) \right)$$

If we now take $n_{j+2} = n_{j+1} + 1 = n_j + 2$ and $n_{i+1} > n_i + 1$ for all $i \ne j, j + 1$ in (2.2.24) and subtract (2.2.30) from it we will get

$$2a(\mathbf{n}) = a(n_1, \ldots, n_j, n_{j+1} + 1, n_{j+1} + 1, \ldots, n_K) + a(n_1, \ldots, n_j, n_{j+1}, n_{j+1}, \ldots, n_K) \tag{2.2.31}$$

Comparing this with (2.2.26) we see that we do not get something new here.

Let us now consider the boundary conditions $a(n_1, \ldots, n_K) = a(n_2, \ldots, n_K, n_1 + N)$ which yield the equation

$$
\begin{aligned}
\sum_{P' \in S_K} A_{P'} \exp\left(i \sum_{j=1}^{K} k_{P'(j)} n_j\right) &= \sum_{P \in S_K} A_P \exp\left(i \sum_{j=1}^{K-1} k_{P(j)} n_{j+1} + i k_{P(K)}(n_1 + N)\right) \\
&= \sum_{P \in S_K} A_P \exp\left(i \sum_{j=2}^{K} k_{P(j-1)} n_j + i k_{P(K)}(n_1 + N)\right)
\end{aligned}
\tag{2.2.32}
$$

We can relate a permutation P' on the left hand side with a permutation P on the right hand side, where P is given by $P(j-1) = P'(j)$, $j = 2, \ldots, K$ and $P(K) = P'(1)$. For such a pair of permutations, equation (2.2.32) will reduce to

$$
A_{P'} = A_P e^{i k_{P(K)} N}
\tag{2.2.33}
$$

Let us now define $U_l = P_{K-1,K} P_{K-2,K-1} \ldots P_{l,l+1}$ and notice that $P' = P U_1$. So

$$
\begin{aligned}
A_{P'} &= A_P \frac{A_{P U_{K-1}}}{A_P} \frac{A_{P U_{K-2}}}{A_{P U_{K-1}}} \ldots \frac{A_{P U_1}}{A_{P U_2}} \\
&= A_P\, S\left(k_{P(K-1)}, k_{P(K)}\right) S\left(k_{P U_{K-1}(K-2)}, k_{P U_{K-1}(K-1)}\right) \ldots S\left(k_{P U_2(1)}, k_{P U_2(2)}\right) \\
&= A_P \prod_{j=1}^{K-1} S\left(k_{P(j)}, k_{P(K)}\right) = A_P e^{i k_{P(K)} N}
\end{aligned}
\tag{2.2.34}
$$

Taking $l \equiv P(K)$ gives the Bethe equations

$$
e^{i k_l N} = \prod_{\substack{j=1 \\ j \neq l}}^{K} S(k_j, k_l) = \prod_{\substack{j=1 \\ j \neq l}}^{K} e^{i \theta_{lj}} \quad \text{for all} \quad l = 1, \ldots, K
\tag{2.2.35}
$$

If we now take the logarithm of this equation we obtain

$$
N k_l = 2\pi \lambda_l + \sum_{\substack{j=1 \\ j \neq l}}^{K} \theta_{lj}
\tag{2.2.36}
$$

with Bethe quantum numbers $\lambda_l \in \{1, \ldots, N - K\}$. What remains to be done is to find the sets of Bethe quantum numbers $(\lambda_1, \ldots, \lambda_K)$ such that they yield solutions of the equations (2.2.29) and (2.2.36). Furthermore, we have to check whether we indeed get $N!/[(N - K)!K!]$ orthogonal eigenvectors with this method. These problems become rather tedious to solve as N becomes large and K approaches $N/2$. In many applications the Bethe ansatz method is therefore used to find selected solutions in limit cases.

Before we continue with the next section, let us first summarize our results in the following proposition

Proposition 2.2.1. *Let $K \leq N$. Then*

$$|\psi\rangle = \sum_{1 \leq n_1 < ... < n_K \leq N} a(n_1, \ldots, n_K)|n_1, \ldots, n_K\rangle \qquad (2.2.37)$$

is an eigenvector of H in the K-subspace with eigenvalue

$$E - E_0 = J \sum_{i=1}^{K} (1 - \cos k_i) \qquad (2.2.38)$$

The wave functions $a(n_1, \ldots, n_K)$ are given by the Bethe ansatz

$$a(n_1, \ldots, n_K) = \sum_{P \in S_K} A_p \exp\left(i \sum_{j=1}^{K} k_{P(j)} n_j \right)$$

with pseudo-momenta $\{k_i\}$ satisfying the Bethe equations

$$e^{ik_i N} = \prod_{\substack{j=1 \\ j \neq i}}^{K} \left(-\frac{e^{i(k_i + k_j)} + 1 - 2e^{ik_i}}{e^{i(k_i + k_j)} + 1 - 2e^{ik_j}} \right)$$

or, equivalently, a set of Bethe quantum numbers $\{\lambda_i\}$ satisfying the equations

$$Nk_i = 2\pi\lambda_i + \sum_{j \neq i} \theta_{ij} \quad and \quad e^{i\theta_{ij}} = -\frac{e^{i(k_i + k_j)} + 1 - 2e^{ik_i}}{e^{i(k_i + k_j)} + 1 - 2e^{ik_j}}$$

2.2.1 The Highest Weight Property of the Bethe Ansatz Eigenvectors

Let us consider the operators $S_T^{\pm} \equiv \sum_{i=1}^{N} S_i^{\pm}$ and S_T^z. Using (2.1.3) we obtain

$$[S_T^z, S_T^z] = [S_T^{\pm}, S_T^{\pm}] = 0, \qquad [S_T^+, S_T^-] = 2S_T^z, \qquad [S_T^z, S_T^{\pm}] = \pm S_T^{\pm} \qquad (2.2.39)$$

Furthermore, in addition to the earlier observation that $[H, S_T^z] = 0$, we also have that the operators $S_T^{\pm}$ commute with H; $[H, S_T^{\pm}] = 0$. Since the spin algebra form the Lie algebra $\mathfrak{su}(2)$ we see that the XXX model has $\mathfrak{su}(2)$ symmetry. Therefore, we expect that the eigenvectors of H will split into irreducible highest weight representations of $\mathfrak{su}(2)$. Each of these representations will have a highest weight vector $|\psi\rangle$, i.e. a nonzero vector $|\psi\rangle$ which satisfies $S_T^+|\psi\rangle = 0$, and the corresponding $\mathfrak{su}(2)$-module generated by applying S_T^- on $|\psi\rangle$ is $2s_z+1$ dimensional where s_z is the eigenvalue from the equation $S_T^z|\psi\rangle = s_z|\psi\rangle$. In the remaining section we will investigate this claim and find these highest weight vectors by studying the eigenvectors of the K-subspace with $K = 0, 1, 2$. To simplify things we will assume that N is large.

First of all, in the $K = 0$ subspace we only have the eigenvector $|F\rangle$ and it is easy to see that $|F\rangle$ is a highest weight vector. Furthermore, since $S_T^z|F\rangle = 1/2N|F\rangle$, the corresponding $\mathfrak{su}(2)$-module is $N + 1$ dimensional.

Now let us turn to the $K = 1$ subspace. Suppose that $|\psi\rangle_k = \sum_{n=1}^{N} e^{ikn}|n\rangle$ is an eigenvector in this subspace, then

$$S_T^+|\psi\rangle_k = \sum_{l=1}^{N} S_l^+ \sum_{n=1}^{N} e^{ikn}|n\rangle = \sum_{l,n=1}^{N} e^{ikn} S_l^+|n\rangle = \sum_{n=1}^{N} e^{ikn}|F\rangle \qquad (2.2.40)$$

Using $k = 2\pi m/N$ with $m = 0, 1, \ldots, N - 1$ we see that $|\psi\rangle_k$ is a highest weight vector if $m \neq 0$:

$$S_T^+|\psi\rangle_k = \sum_{n=1}^{N} e^{\frac{2\pi im}{N}n}|F\rangle = \frac{e^{\frac{2\pi im}{N}}(1 - e^{2\pi im})}{1 - e^{\frac{2\pi im}{N}}}|F\rangle = 0 \qquad (2.2.41)$$

Conversely, if $m = 0$, then $|\psi\rangle_0$ is not a highest weight vector since:

$$S_T^+|\psi\rangle_0 = \sum_{n=1}^{N} |F\rangle = N|F\rangle \neq 0 \qquad (2.2.42)$$

This is not a big surprise, because $|\psi\rangle_0$ is a "descendent" vector of the highest weight vector $|F\rangle$:

$$|\psi\rangle_0 = \sum_{n=1}^{N} |n\rangle = \sum_{n=1}^{N} S_n^-|F\rangle = S_T^-|F\rangle \qquad (2.2.43)$$

Let us now consider the $K = 2$ subspace. First notice that all the eigenvectors (2.2.13) with $k_1 = 0$ are descendent vectors of the eigenvectors from the $K = 1$ subspace. To see this, let $|\phi\rangle_k$ be an eigenvector from the $K = 2$ subspace with $k_1 = 0$ and $k_2 = k$. Then

$$S_T^-|\psi\rangle_k = \sum_{l=1}^{N} S_l^- \sum_{n=1}^{N} e^{ikn}|n\rangle = \sum_{1 \leq l < n \leq N} e^{ikn}|l, n\rangle + \sum_{1 \leq n < l \leq N} e^{ikn}|n, l\rangle$$

$$= \sum_{1 \leq n_1 < n_2 \leq N} (e^{ikn_1} + e^{ikn_2})|n_1, n_2\rangle = |\phi\rangle_k \qquad (2.2.44)$$

after renaming l, n with n_1, n_2 respectively in the first sum and n, l with n_1, n_2 respectively in the second sum. So $|\phi\rangle_k$ is indeed a descendant vector. Let us now check that $|\phi\rangle_k$ is not a highest weight vector, i.e. $S_T^+|\phi\rangle_k \neq 0$. Suppose that $k \neq 0$, then from the discussion above we know that $|\psi\rangle_k$ is a highest weight vector. So

$$S_T^+|\phi\rangle_k = S_T^+ S_T^-|\psi\rangle_k = \left(S_T^- S_T^+ + [S_T^+, S_T^-] \right)|\psi\rangle_k = \sum_{1 \leq n,m \leq N} [S_n^+, S_m^-]|\psi\rangle_k$$

$$= \sum_{1 \leq n,m \leq N} 2S_n^z \delta_{nm}|\psi\rangle_k = (N - 2)|\psi\rangle_k \neq 0 \qquad (2.2.45)$$

Now suppose that $k = 0$, then using (2.2.44), (2.2.43) and (2.2.39) gives

$$
\begin{aligned}
S_T^+|\phi\rangle_0 &= S_T^+ S_T^- S_T^-|F\rangle = \left(S_T^- S_T^- S_T^+ + [S_T^+, S_T^- S_T^-] \right)|F\rangle \\
&= \left([S_T^+, S_T^-] S_T^- + S_T^-[S_T^+, S_T^-] \right)|F\rangle = \left(2 S_T^z S_T^- + 2 S_T^- S_T^z \right)|F\rangle \neq 0
\end{aligned}
\tag{2.2.46}
$$

Therefore, eigenvectors of the form $|\phi\rangle_k$ are not highest weight vectors.

We will now prove that eigenvectors $|\phi\rangle$ with none-zero pseudo-momenta k_i *are* highest weight vectors. Suppose that $|\phi\rangle = \sum_{n_1 < n_2} a(n_1, n_2)|n_1, n_2\rangle$ is an eigenvector, then

$$
S_T^+|\phi\rangle = S_T^+ \sum_{n_1 < n_2} a(n_1, n_2)|n_1, n_2\rangle = \sum_{n_1 < n_2} a(n_1, n_2)|n_2\rangle + \sum_{n_1 < n_2} a(n_1, n_2)|n_1\rangle
\tag{2.2.47}
$$

$$
= \sum_{n_2 < n_1} a(n_2, n_1)|n_1\rangle + \sum_{n_1 < n_2} a(n_1, n_2)|n_1\rangle = \sum_{n_1=2}^{N} \sum_{n_2=1}^{n_1-1} a(n_2, n_1)|n_1\rangle + \sum_{n_1=1}^{N-1} \sum_{n_2=n_1+1}^{N} a(n_1, n_2)|n_1\rangle
$$

$$
= \sum_{n_1=2}^{N-1} \left(\sum_{n_2=1}^{n_1-1} a(n_2, n_1) + \sum_{n_2=n_1+1}^{N} a(n_1, n_2) \right)|n_1\rangle + \sum_{n_2=1}^{N-1} a(n_2, N)|N\rangle + \sum_{n_2=2}^{N} a(1, n_2)|1\rangle
$$

Since $\{|n\rangle\}$ is an orthogonal basis, the highest weight vector condition is equivalent with the vanishing of each of the terms in the equation above. To check that this is indeed the case, we will plug in the expressions for $a(n_1, n_2)$. We will first consider the last term of the above expression:

$$
\sum_{n_2=2}^{N} a(1, n_2) = \sum_{n_2=2}^{N} \left(A e^{i(k_1 + k_2 n_2)} + B e^{i(k_2 + k_1 n_2)} \right) = A \left(e^{ik_1} \sum_{n_2=2}^{N} e^{ik_2 n_2} + \frac{B}{A} e^{ik_2} \sum_{n_2=2}^{N} e^{ik_1 n_2} \right)
$$

$$
= A e^{i(k_1+k_2)} \left(\frac{e^{ik_2} - e^{ik_2 N}}{1 - e^{ik_2}} + \frac{B}{A} \frac{e^{ik_1} - e^{ik_1 N}}{1 - e^{ik_1}} \right)
\tag{2.2.48}
$$

Notice that the condition of nonzero k_i is used for the evaluation of the sums. If we now remind ourselves the relations

$$
\frac{A}{B} = e^{i\theta} = -\frac{e^{i(k_1+k_2)} + 1 - 2e^{ik_1}}{e^{i(k_1+k_2)} + 1 - 2e^{ik_2}}, \quad e^{ik_1 N} = e^{i\theta} \quad \text{and} \quad e^{ik_2 N} = e^{-i\theta}
\tag{2.2.49}
$$

we obtain

$$
\sum_{n_2=2}^{N} a(1, n_2) = A e^{i(k_1+k_2)} \left(\frac{e^{ik_2} - e^{-i\theta}}{1 - e^{ik_2}} + e^{-i\theta} \frac{e^{ik_1} - e^{i\theta}}{1 - e^{ik_1}} \right)
$$

$$
= A e^{i(k_1+k_2)} \left(\frac{(1 - e^{ik_1})(e^{ik_2} - e^{-i\theta}) + (1 - e^{ik_2}) e^{-i\theta}(e^{ik_1} - e^{i\theta})}{(1 - e^{ik_1})(1 - e^{ik_2})} \right)
\tag{2.2.50}
$$

$$
= A e^{i(k_1+k_2)} \left(\frac{\left(2e^{ik_2} - 1 - e^{i(k_1+k_2)} \right) + e^{-i\theta} \left(2e^{ik_1} - 1 - e^{i(k_1+k_2)} \right)}{(1 - e^{ik_1})(1 - e^{ik_2})} \right) = 0
$$

We have now proven that $\sum_{n_2=2}^{N} a(1, n_2)|1\rangle = 0$. Notice that for this calculation we have explicitly used the Bethe ansatz equations and the fact that the k_i are nonzero. Analogously, with the same kind of calculation we can show that the other terms of (2.2.47) are zero. So all eigenvectors $|\phi\rangle$ with none-zero pseudo-momenta k_i are highest weight vectors.

The generalization of these calculations to arbitrary r is straightforward but laborious. Without proof, let us formulate the following statements on highest weight vectors.

Proposition 2.2.2. *Suppose that $n < N/2$. Then there is a bijective correspondence between the eigenvectors in the subspace $K = n$ with the eigenvectors in the subspace $K = n + 1$ which are not highest weight vectors in the following sense.*

- *If $|\psi\rangle$ is an eigenvector in the subspace $K = n$ with pseudo-momenta $\{k_i\}_{i=1}^{n}$, then $S_T^-|\psi\rangle$ is a non highest weight eigenvector in the subspace $K = n + 1$ with pseudo-momenta $\{k_i'\}_{i=1}^{n+1}$ where $k_1' = 0$ and $k_{i+1}' = k_i$ for $i = 1, 2, \ldots, n$. The map S_T^- is injective.*

- *Conversely, every non highest weight eigenvector $|\psi'\rangle$ in the subspace $K = n + 1$ with pseudo-momenta $\{k_i\}_{i=1}^{n}$ has a pseudo-momentum k_j which is zero and is a descendent from an eigenvector $|\psi\rangle$ from the $K = n$ subspace; $|\psi'\rangle = S_T^-|\psi\rangle$.*

An eigenvector $|\psi'\rangle$ from the $K = n + 1$ subspace is a highest weight vector if and only if $|\psi'\rangle$ is not a descendent vector; there is no eigenvector $|\psi\rangle$ from the $K = n$ subspace such that $|\psi'\rangle = S_T^-|\psi\rangle$

2.3   The Algebraic Bethe Ansatz (ABA)

Let us solve the XXX Heisenberg model again, but this time using the algebraic Bethe ansatz. We will see that the ABA is a more powerful method than the CBA since it enables us to derive properties of the XXX model such as integrability. At this point we note that in the following sections we will make use of some tensor notations which the reader may be unfamiliar with. If this is the case, the reader is advised to have a look at appendix A.1.

2.3.1   The Integrability of the XXX Model

Let us define the notion of "integrability" first. In Hamiltonian mechanics, a Hamiltonian system with n degrees of freedom is called a completely integrable system if it has n conserved quantities such that the Poisson bracket between each pair of these conserved quantities vanishes. This notion of integrability can be carried over to quantum systems. In that case, a quantum system with n degrees of freedom is called a completely integrable system if it has n conserved quantities such that the commutator between each pair of these conserved quantities vanishes. This means that the XXX Heisenberg model of length N is integrable if it has N commuting operators. [2]

---

[2]This definition of quantum integrability is not entirely correct and it is still a subject of research on how to make a better definition. For example, the XXX model actually has 2^N degrees of freedom since each lattice

A completely integrable system is interesting because for such a system it is possible to calculate its eigenvalues and eigenvectors exactly. Furthermore, there is great number of topics and theorems that deal with integrability which enables us to investigate the system more properly.

Let us now show how we can determine the integrability of a quantum system. Suppose that we have a periodic spin chain of length N. In the theory of quantum integrable systems, a quantum integrable ultralocal model on this spin chain is characterized by an operator $L_{a,i}(\lambda) : V_a \otimes V_i \rightarrow V_a \otimes V_i$ where i refers to lattice point i, a being the label of the auxiliary space and λ is the spectral parameter of the operator. This operator is called a Lax operator and it satisfies the Fundamental Commutation Relations (FCR)[3]

$$R_{a,b}(\lambda - \mu)L_{a,i}(\lambda)L_{b,i}(\mu) = L_{b,i}(\mu)L_{a,i}(\lambda)R_{a,b}(\lambda - \mu) \tag{2.3.1}$$

where $R_{a,b}(\lambda - \mu) : V_a \otimes V_b \rightarrow V_a \otimes V_b$ is a R-matrix. Notice that we get the usual Yang-Baxter equation (A.2.5) back when $L(\lambda) = R(\lambda)$.

Next we define the monodromy matrix

$$T_a(\lambda) \equiv L_{a,N}(\lambda) \ldots L_{a,1}(\lambda) \tag{2.3.2}$$

The monodromy matrix is an operator on $V_a \otimes V_1 \otimes V_2 \otimes \ldots \otimes V_N$. If we take the partial trace over the auxiliary space of this operator, we will get an object called the transfer matrix and it is denoted by

$$\tau(\mu) = tr_a[T_a(\mu)] \tag{2.3.3}$$

The transfer matrix is an operator on $V_1 \otimes V_2 \otimes \ldots \otimes V_N$.

Now let us use the shorthand notations $R_{a,b} = R_{a,b}(\lambda - \mu)$, $L_a = L_{a,n}(\lambda)$, $L'_a = L_{a,n+1}(\lambda)$, $L_b = L_{b,n}(\mu)$ and $L'_b = L_{b,n+1}(\mu)$. Then using (2.3.1) we obtain

$$\begin{aligned} R_{a,b}L'_aL_aL'_bL_b = R_{a,b}L'_aL'_bL_aL_b = L'_bL'_aR_{a,b}L_aL_b \\ = L'_bL'_aL_bL_aR_{a,b} = L'_bL_bL'_aL_aR_{a,b} \end{aligned} \tag{2.3.4}$$

This procedure can easily be generalized to give the expression

$$R_{a,b}(\lambda - \mu)[T_a(\lambda)T_b(\mu)] = [T_b(\mu)T_a(\lambda)]R_{a,b}(\lambda - \mu) \tag{2.3.5}$$

Using this relation and using the cyclicity of the trace, we can derive the following expression for the transfer matrix:

$$\begin{aligned} \tau(\lambda)\tau(\mu) = tr_a[T_a(\lambda)]\, tr_b[T_b(\mu)] = tr_{a,b}[T_a(\lambda)T_b(\mu)] = tr_{a,b}[R_{a,b}T_a(\lambda)T_b(\mu)R_{a,b}^{-1}] \\ = tr_{a,b}[T_b(\mu)T_a(\lambda)] = tr_b[T_b(\mu)]\, tr_a[T_a(\lambda)] = \tau(\mu)\tau(\lambda) \end{aligned} \tag{2.3.6}$$

---

site can be occupied by a spin up or a spin down electron. But in this definition we only consider the lattice site of the XXX model of having one degree of freedom.

[3]An alternate form of the FCR can be found in Appendix A.3

which means that the transfer matrices commute with each other:

$$[\tau(\lambda), \tau(\mu)] = 0 \tag{2.3.7}$$

Generally, $\tau(\lambda)$ will be the generator of conserved quantities. The relation (2.3.7) will then imply that the conserved quantities are independent of each other which proves the integrability of the model.

Finally, the Hamiltonian of the model will be represented as a linear combination of logarithmic derivatives of the transfer matrix $\tau(\lambda)$ at some points λ_a:

$$H = \sum_k \sum_a c_{k,a} \frac{d^k}{d\lambda^k} \ln \tau(\lambda) \Big|_{\lambda=\lambda_a} \tag{2.3.8}$$

with coefficients $c_{k,a}$. This ensures that the Hamiltonian is part of the family of commuting conserved quantities.

After having sketched the general theory of quantum integrable systems, let us now apply this to the XXX model. The Lax operator $L_{a,n}(\lambda)$ for the XXX model is given by

$$L_{a,n}(\lambda) = \lambda(\mathbb{I}_a \otimes \mathbb{I}_n) + i \sum_{\alpha=x,y,z} (\sigma_a^\alpha \otimes S_n^\alpha) = \begin{pmatrix} \lambda + iS_n^z & iS_n^- \\ iS_n^+ & \lambda - iS_n^z \end{pmatrix} \tag{2.3.9}$$

where $L_{a,n}(\lambda)$ acts on $V_a \otimes h_n$ with V_a being the auxiliary space. We will choose the auxiliary space to be $\mathbb{C}^2$ so that $h_n = V_a = \mathbb{C}^2$ for all n. Let us now consider the permutation operator $P : \mathbb{C}^2 \otimes \mathbb{C}^2 : a \otimes b \to b \otimes a$. If we use

$$\{e_1 \equiv |\uparrow\rangle \otimes |\uparrow\rangle, \quad e_2 \equiv |\uparrow\rangle \otimes |\downarrow\rangle, \quad e_3 \equiv |\downarrow\rangle \otimes |\uparrow\rangle, \quad e_4 \equiv |\downarrow\rangle \otimes |\downarrow\rangle\} \tag{2.3.10}$$

as basis for $\mathbb{C}^2 \otimes \mathbb{C}^2$ with $|\uparrow\rangle$ and $|\downarrow\rangle$ as defined in (2.1.6), we see that P can be written as

$$P = \frac{1}{2}\left(\mathbb{I} \otimes \mathbb{I} + \sum_\alpha \sigma^\alpha \otimes \sigma^\alpha\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.3.11}$$

We will write $P_{a,b}$ to indicate that the permutation operator is acting on the space $V_a \otimes V_b$. Note that this only makes sense if $V_a = V_b = \mathbb{C}^2$, which will always be the case unless stated otherwise. It is easy to see that the permutation operator satisfies the following identities

$$tr_a P_{a,b} = tr_b P_{a,b} = \mathbb{I}, \quad P_{a,b} = P_{b,a}, \quad P_{a_1,b} P_{a_2,b} = P_{a_1,a_2} P_{a_1,b} = P_{a_2,b} P_{a_1,a_2} \tag{2.3.12}$$

where tr_b is the partial trace over the space V_b.

With the above relations, we see that $L_{a,n}(\lambda)$ can be written as

$$L_{a,n}(\lambda) = \left(\lambda - \frac{i}{2}\right)\mathbb{I}_{a,n} + iP_{a,n} \tag{2.3.13}$$

The corresponding R-matrix $R_{a,b}(\lambda)$ such that equation (2.3.1) is satisfied is given by

$$R_{a,b}(\lambda) = \lambda \mathbb{I}_{a,b} + i P_{a,b} \tag{2.3.14}$$

where $\mathbb{I}_{a,b}$ and $P_{a,b}$ are the unity and permutation operator in $V_a \otimes V_b$ respectively. Let us note that we can write $R_{a,b}(\lambda)$ and $L_{a,n}(\lambda)$ as

$$R_{a,b}(\lambda) = (\lambda + i) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\lambda}{\lambda+i} & \frac{i}{\lambda+i} & 0 \\ 0 & \frac{i}{\lambda+i} & \frac{\lambda}{\lambda+i} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad L_{a,n}(\lambda) = R_{a,n}(\lambda - i/2) \tag{2.3.15}$$

We will now consider the monodromy matrix $T_a(\lambda) = L_{a,N}(\lambda) \dots L_{a,1}(\lambda)$ and denote it by

$$T_a(\lambda) = \begin{pmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{pmatrix} \tag{2.3.16}$$

where the matrix elements are operators on the space h. Using (2.3.9) we see that the monodromy $T_a(\lambda)$ is a polynomial in λ of order N:

$$T_a(\lambda) = \lambda^N + i\lambda^{N-1} \sum_\alpha \sum_{n=1}^N (\sigma_a^\alpha \otimes S_n^\alpha) + \dots \tag{2.3.17}$$

So the expression for the transfer matrix $\tau(\lambda)$ becomes

$$\tau(\lambda) = tr_a[T_a(\lambda)] = A(\lambda) + D(\lambda) \equiv 2\lambda^N + \sum_{l=0}^{N-2} Q_l \lambda^l \tag{2.3.18}$$

Notice that there is no λ^{N-1} term since the Pauli matrices σ^α are traceless. Furthermore, since $[\tau(\lambda), \tau(\mu)] = 0$ for all λ and μ, we see that

$$[Q_k, Q_l] = \frac{d^k}{d\lambda^k} \frac{d^l}{d\mu^l} [\tau(\lambda), \tau(\mu)] \Big|_{\lambda=\mu=0} = 0 \tag{2.3.19}$$

So the operators Q_l commute with each other.

In order to prove that the XXX-model is completely integrable, we will show that its Hamiltonian H belongs to the family of the $N - 1$ commuting operators Q_l. First of all,

notice from (2.3.13) that $L_{n,a}(i/2) = iP_{n,a}$. Using (2.3.12) gives us the identities

$$T_a(i/2) = i^N P_{a,N} P_{a,N-1} \ldots P_{a,1} = i^N P_{1,2} P_{2,3} \ldots P_{N-1,N} P_{N,a}$$

$$\tau(i/2) = tr_a T_a(i/2) = i^N P_{1,2} P_{2,3} \ldots P_{N-1,N}$$

$$\frac{d}{d\lambda} T_a(\lambda)\Big|_{i/2} = i^{N-1} \sum_{n=1}^{N} P_{a,N} \ldots \hat{P}_{a,n} \ldots P_{a,1} = i^{N-1} \sum_{n=2}^{N-1} P_{1,2} P_{2,3} \ldots P_{n-1,n+1} \ldots P_{N-1,N} P_{N,a}$$

$$+ i^{N-1} \left(P_{2,3} P_{3,4} \ldots P_{N-1,N} P_{N,a} + P_{1,2} P_{2,3} \ldots P_{N-2,N-1} P_{N-1,a} \right)$$

$$\frac{d}{d\lambda} \tau(\lambda)\Big|_{i/2} = \frac{d}{d\lambda} tr_a T_a(\lambda)\Big|_{i/2} = i^{N-1} \left(P_{2,3} P_{3,4} \ldots P_{N-1,N} + P_{1,2} P_{2,3} \ldots P_{N-2,N-1} \right)$$

$$+ i^{N-1} \sum_{n=2}^{N-1} P_{1,2} P_{2,3} \ldots P_{n-1,n+1} \ldots P_{N-1,N} \tag{2.3.20}$$

where ˆ means that the corresponding permutation is absent. With these results we see that

$$\frac{d}{d\lambda} \ln \tau(\lambda)\Big|_{i/2} = \left(\frac{d}{d\lambda} \tau(\lambda) \right) \tau(\lambda)^{-1}\Big|_{i/2} = \frac{1}{i^N} \frac{d}{d\lambda} \tau(\lambda)\Big|_{i/2} (P_{N,N-1} P_{N-1,N-2} \ldots P_{2,1}) = \frac{1}{i} \sum_{n=1}^{N} P_{n,n+1} \tag{2.3.21}$$

Now notice that using (2.3.11), the Hamiltonian (2.1.2) can be written as

$$H = -J \left(\frac{1}{2} \sum_{n=1}^{N} P_{n,n+1} - \frac{N}{4} \right) \quad \text{with} \quad P_{N,N+1} \equiv P_{N,1} \tag{2.3.22}$$

which, in view of (2.3.21), can also be written as

$$H = -J \left(\frac{i}{2} \frac{d}{d\lambda} \ln \tau(\lambda)\Big|_{i/2} - \frac{N}{4} \right) \tag{2.3.23}$$

Therefore, H belongs to the family of $N-1$ commuting operators. To obtain N commuting operators, we add the operator S_T^z to this family. This completes the proof of integrability of the XXX-model.

2.3.2 The Spectrum of the XXX Model

We will now calculate the eigenvalues of H by diagonalizing $\tau(\lambda)$. This will be done by a procedure called the "Algebraic Bethe Ansatz" (ABA). First of all, let us derive commutation relations between the operators A, B, C and D using equation (2.3.5). Writing the

matrices $R(\lambda, \mu)$, $T_a(\lambda)$ and $T_b(\mu)$ in the matrix representation using the basis (2.3.10) gives

$$R(\lambda - \mu) = \begin{pmatrix} \lambda - \mu + i & 0 & 0 & 0 \\ 0 & \lambda - \mu & i & 0 \\ 0 & i & \lambda - \mu & 0 \\ 0 & 0 & 0 & \lambda - \mu + i \end{pmatrix} \tag{2.3.24}$$

$$T_a(\lambda) = \begin{pmatrix} A(\lambda) & 0 & B(\lambda) & 0 \\ 0 & A(\lambda) & 0 & B(\lambda) \\ C(\lambda) & 0 & D(\lambda) & 0 \\ 0 & C(\lambda) & 0 & D(\lambda) \end{pmatrix}, \quad T_b(\mu) = \begin{pmatrix} A(\mu) & B(\mu) & 0 & 0 \\ C(\mu) & D(\mu) & 0 & 0 \\ 0 & 0 & A(\mu) & B(\mu) \\ 0 & 0 & C(\mu) & D(\mu) \end{pmatrix}$$

Thus, equation (2.3.5) reads

$$\begin{pmatrix} (\alpha + i)A_\lambda A_\mu & (\alpha + i)A_\lambda B_\mu & (\alpha + i)B_\lambda A_\mu & (\alpha + i)B_\lambda B_\mu \\ \alpha A_\lambda C_\mu + iC_\lambda A_\mu & \alpha A_\lambda D_\mu + iC_\lambda B_\mu & \alpha B_\lambda C_\mu + iD_\lambda A_\mu & \alpha B_\lambda D_\mu + iD_\lambda B_\mu \\ iA_\lambda C_\mu + \alpha C_\lambda A_\mu & iA_\lambda D_\mu + \alpha C_\lambda B_\mu & iB_\lambda C_\mu + \alpha D_\lambda A_\mu & iB_\lambda D_\mu + \alpha D_\lambda B_\mu \\ (\alpha + i)C_\lambda C_\mu & (\alpha + i)C_\lambda D_\mu & (\alpha + i)D_\lambda C_\mu & (\alpha + i)D_\lambda D_\mu \end{pmatrix} =$$

$$= \begin{pmatrix} (\alpha + i)A_\mu A_\lambda & \alpha B_\mu A_\lambda + iA_\mu B_\lambda & iB_\mu A_\lambda + \alpha A_\mu B_\lambda & (\alpha + i)B_\mu B_\lambda \\ (\alpha + i)C_\mu A_\lambda & \alpha D_\mu A_\lambda + iC_\mu B_\lambda & iD_\mu A_\lambda + \alpha C_\mu B_\lambda & (\alpha + i)D_\mu B_\lambda \\ (\alpha + i)A_\mu C_\lambda & \alpha B_\mu C_\lambda + iA_\mu D_\lambda & iB_\mu C_\lambda + \alpha A_\mu D_\lambda & (\alpha + i)B_\mu D_\lambda \\ (\alpha + i)C_\mu C_\lambda & \alpha D_\mu C_\lambda + iC_\mu D_\lambda & iD_\mu C_\lambda + \alpha C_\mu D_\lambda & (\alpha + i)D_\mu D_\lambda \end{pmatrix} \tag{2.3.25}$$

where we have used the shorthand notations $A_\lambda \equiv A(\lambda)$ and $\alpha \equiv \lambda - \mu$. The relevant relations for the ABA follow from the (1,4), (1,3) and (3,4) matrix entries of the above equation:

$$[B(\lambda), B(\mu)] = 0 \tag{2.3.26}$$

$$A(\lambda)B(\mu) = \frac{\lambda - \mu - i}{\lambda - \mu}B(\mu)A(\lambda) + \frac{i}{\lambda - \mu}B(\lambda)A(\mu) \tag{2.3.27}$$

$$D(\lambda)B(\mu) = \frac{\lambda - \mu + i}{\lambda - \mu}B(\mu)D(\lambda) - \frac{i}{\lambda - \mu}B(\lambda)D(\mu) \tag{2.3.28}$$

Notice that we interchanged $\lambda \leftrightarrow \mu$ in the second equation.

Let us now consider the monodromy $T_a(\lambda)$ from (2.3.16). The main idea of the ABA is that there exists a pseudo-vacuum $|0\rangle \in h$ such that $C(\lambda)|0\rangle = 0$ and that the eigenvectors of $\tau(\lambda)$ with K spins down, have the form

$$|\lambda_1, \lambda_2, \ldots, \lambda_K\rangle = B(\lambda_1)B(\lambda_2)\ldots B(\lambda_K)|0\rangle \tag{2.3.29}$$

where $\{\lambda_i\}$ are the so called "Bethe roots" which are comparable with the pseudo-momenta $\{k_i\}$ of the CBA. We will see later how those are related with each other.

In order to find $|0\rangle$ we will look at the Lax operator $L_{a,n}(\lambda) : V_a \otimes h_n \to V_a \otimes h_n$ first. Let $|\uparrow\rangle_n \in h_n$ as defined in (2.1.6) and v an arbitrary vector in V_a, then we see that

$$L_n(\lambda)(v \otimes |\uparrow\rangle_n) = \begin{pmatrix} (\lambda + iS_n^3)|\uparrow\rangle_n & iS_n^-|\uparrow\rangle_n \\ iS_n^+|\uparrow\rangle_n & (\lambda - iS_n^3)|\uparrow\rangle_n \end{pmatrix} v = \begin{pmatrix} (\lambda + \frac{i}{2})|\uparrow\rangle_n & i|\downarrow\rangle_n \\ 0 & (\lambda - \frac{i}{2})|\uparrow\rangle_n \end{pmatrix} v$$

or in short: $\quad L_n(\lambda)|\uparrow\rangle_n = \begin{pmatrix} (\lambda + \frac{i}{2})|\uparrow\rangle_n & i|\downarrow\rangle_n \\ 0 & (\lambda - \frac{i}{2})|\uparrow\rangle_n \end{pmatrix}$ \hfill (2.3.30)

So if we now let

$$|0\rangle \equiv \bigotimes_{n=1}^{N} |\uparrow\rangle_n \tag{2.3.31}$$

then

$$T(\lambda)|0\rangle = \begin{pmatrix} (\lambda + \frac{i}{2})^N |0\rangle & * \\ 0 & (\lambda - \frac{i}{2})^N |0\rangle \end{pmatrix} \tag{2.3.32}$$

where $*$ denotes an expression not relevant for our discussion. So

$$A(\lambda)|0\rangle = \left(\lambda + \frac{i}{2}\right)^N |0\rangle \quad \text{and} \quad D(\lambda)|0\rangle = \left(\lambda - \frac{i}{2}\right)^N |0\rangle \tag{2.3.33}$$

Now notice that $|0\rangle$ is exactly the vector $|F\rangle$ from the CBA. Furthermore, we see that $|0\rangle$ is an eigenvector of $A(\lambda)$ and $D(\lambda)$ simultaneously, which means that it is also an eigenvector of $\tau(\lambda) = A(\lambda) + D(\lambda)$. The ABA then tells us that the other eigenvectors will be of the form

$$|\lambda_1, \lambda_2, \ldots, \lambda_K\rangle = B(\lambda_1)B(\lambda_2)\ldots B(\lambda_K)|0\rangle \tag{2.3.34}$$

The condition that $|\lambda_1, \lambda_2, \ldots, \lambda_K\rangle$ is an eigenvector of $\tau(\lambda)$ will generate restrictions on the parameters $\lambda_1, \ldots, \lambda_K$. Let us now derive these restrictions. First, using relation (2.3.27) we get

$$A(\lambda)B(\lambda_1)\ldots B(\lambda_K)|0\rangle = \left(\lambda + \frac{i}{2}\right)^N \left(\prod_{n=1}^{K} \frac{\lambda - \lambda_n - i}{\lambda - \lambda_n}\right) B(\lambda_1)\ldots B(\lambda_K)|0\rangle$$
$$+ \sum_{n=1}^{K} M_n(\lambda, \{\lambda_i\}_{i=1}^{K}) B(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{K} B(\lambda_j)|0\rangle \tag{2.3.35}$$

where the first term of the right hand side is obtained by using only the first term of the right hand side of (2.3.27) and where $M_n(\lambda, \{\lambda_i\}_{i=1}^{K})$ is a coefficient dependent on λ, λ_i with $1 \leq i \leq K$. To determine this coefficient let us first note that we can write

$$|\lambda_1, \lambda_2, \ldots, \lambda_K\rangle = B(\lambda_n) \prod_{\substack{j=1 \\ j \neq n}}^{K} B(\lambda_j)|0\rangle \quad \text{for} \quad 1 \leq n \leq K \tag{2.3.36}$$

since the operators $B(\lambda)$ commute with each other. So

$$A(\lambda)|\lambda_1, \lambda_2, \ldots, \lambda_K\rangle = \frac{\lambda - \lambda_n - i}{\lambda - \lambda_n} B(\lambda_n)A(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{K} B(\lambda_j)|0\rangle + \frac{i}{\lambda - \lambda_n} B(\lambda)A(\lambda_n) \prod_{\substack{j=1 \\ j \neq n}}^{K} B(\lambda_j)|0\rangle$$
$$\tag{2.3.37}$$

From this equation we see that only the second term of the right hand side will contribute to M_n since this term that does not contain $B(\lambda_n)$. If we now move $A(\lambda)$ past the $B(\lambda_j)$ we see that the only way to avoid the appearance of $B(\lambda_n)$ is by using the first term of the right hand side of (2.3.27). So the resulting term will be

$$\frac{i}{\lambda - \lambda_n}\left(\lambda_n + \frac{i}{2}\right)^N \left(\prod_{\substack{i=1 \\ i \neq n}}^{K} \frac{\lambda_n - \lambda_i - i}{\lambda_n - \lambda_i}\right) B(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{K} B(\lambda_j)|0\rangle \tag{2.3.38}$$

which means that

$$M_n(\lambda, \{\lambda_i\}_{i=1}^K) = \frac{i}{\lambda - \lambda_n}\left(\lambda_n + \frac{i}{2}\right)^N \prod_{\substack{j=1 \\ j \neq n}}^{K} \frac{\lambda_n - \lambda_j - i}{\lambda_n - \lambda_j} \tag{2.3.39}$$

In the same way we get

$$D(\lambda)B(\lambda_1)\dots B(\lambda_K)|0\rangle = \left(\lambda - \frac{i}{2}\right)^N \left(\prod_{n=1}^{K} \frac{\lambda - \lambda_n + i}{\lambda - \lambda_n}\right) B(\lambda_1)\dots B(\lambda_K)|0\rangle$$

$$+ \sum_{n=1}^{K} N_n(\lambda, \{\lambda_i\}_{i=1}^K) B(\lambda) \prod_{\substack{j=1 \\ j \neq n}}^{K} B(\lambda_j)|0\rangle \tag{2.3.40}$$

$$\text{with} \quad N_n(\lambda, \{\lambda_i\}_{i=1}^K) = -\frac{i}{\lambda - \lambda_n}\left(\lambda_n - \frac{i}{2}\right)^N \prod_{\substack{j=1 \\ j \neq n}}^{K} \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j}$$

From this we see that

$$(A(\lambda) + D(\lambda))|\lambda_1, \dots, \lambda_K\rangle = \Lambda(\lambda, \{\lambda_i\}_{i=1}^K)|\lambda_1, \dots, \lambda_K\rangle$$

$$\text{with} \quad \Lambda(\lambda, \{\lambda_i\}_{i=1}^K) = \left(\lambda + \frac{i}{2}\right)^N \prod_{j=1}^{K} \frac{\lambda - \lambda_j - i}{\lambda - \lambda_j} + \left(\lambda - \frac{i}{2}\right)^N \prod_{j=1}^{K} \frac{\lambda - \lambda_j + i}{\lambda - \lambda_j} \tag{2.3.41}$$

if $M_n + N_n = 0$ for all n, which written out is:

$$\left(\lambda_n + \frac{i}{2}\right)^N \prod_{\substack{j=1 \\ j \neq n}}^{K} \frac{\lambda_n - \lambda_j - i}{\lambda_n - \lambda_j} = \left(\lambda_n - \frac{i}{2}\right)^N \prod_{\substack{j=1 \\ j \neq n}}^{K} \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j} \tag{2.3.42}$$

or, written differently:

$$\left(\frac{\lambda_n + i/2}{\lambda_n - i/2}\right)^N = \prod_{\substack{j=1 \\ j \neq n}}^{K} \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j - i} \quad \text{for all} \quad n = 1, \dots, K \tag{2.3.43}$$

These are the Bethe (ansatz) equations and the eigenvectors are the Bethe vectors. Note that if we take $\lambda_i = \frac{1}{2}\cot\left(\frac{k_i}{2}\right)$, then

$$\lambda_n \pm \frac{i}{2} = \frac{1}{2}\left(\frac{\cos(k_n/2)}{\sin(k_n/2)} \pm i\frac{\sin(k_n/2)}{\sin(k_n/2)}\right) = \frac{e^{\pm\frac{1}{2}ik_n}}{2\sin(k_n/2)} \tag{2.3.44}$$

and

$$\lambda_n - \lambda_j + i = (\lambda_n + i/2) - (\lambda_j - i/2) = \frac{\left(e^{\frac{1}{2}ik_j} - e^{-\frac{1}{2}ik_j}\right)e^{\frac{1}{2}ik_n} - \left(e^{\frac{1}{2}ik_n} - e^{-\frac{1}{2}ik_n}\right)e^{-\frac{1}{2}ik_j}}{4i\sin(k_n/2)\sin(k_j/2)}$$

$$= \frac{e^{-\frac{1}{2}i(k_n+k_j)}}{4i\sin(k_n/2)\sin(k_j/2)}\left(e^{i(k_n+k_j)} - 2e^{ik_n} + 1\right) \tag{2.3.45}$$

With these relations we see that

$$\left(\frac{\lambda_n + i/2}{\lambda_n - i/2}\right)^N = e^{ik_n N} \quad \text{and} \quad \frac{\lambda_n - \lambda_j + i}{\lambda_n - \lambda_j - i} = -\frac{(\lambda_n + i/2) - (\lambda_j - i/2)}{(\lambda_j + i/2) - (\lambda_n - i/2)} = -\frac{e^{i(k_n+k_j)} - 2e^{ik_n} + 1}{e^{i(k_n+k_j)} - 2e^{ik_j} + 1} \tag{2.3.46}$$

Therefore, the Bethe equations of the ABA agree with the Bethe equations of the CBA. However, there is a slight difference in the solution sets between the CBA and the ABA. First of all, notice that the parametrization $\lambda_i = \frac{1}{2}\cot\left(\frac{k_i}{2}\right)$ has a singularity at $k_i = 0$. Since all λ_i must be finite, this means that only the solutions from the CBA for which all k_i are nonzero arise in the ABA as solutions. Therefore, since all the eigenvectors for which the k_i are nonzero are highest weight vectors of the spin algebra, we expect that all the eigenvectors obtained with the ABA are highest weight vectors. Let us investigate this claim.

Let us consider equation (2.3.5) in the limit $\lambda \to \infty$. Then using the expansion (2.3.17) for $T_a(\lambda)$ and the R-matrix (2.3.14) with the permutation P of the form (2.3.11) we see that this equation becomes

$$\left((\lambda - \mu) + \frac{1}{2}i\{\mathbb{I}_a \otimes \mathbb{I}_b + \sum_\alpha(\sigma_a^\alpha \otimes \sigma_b^\alpha)\}\right)\left(\lambda^N + i\lambda^{N-1}\sum_{\alpha,n}(\sigma_a^\alpha \otimes S_n^\alpha) + \dots\right)T_b(\mu) =$$

$$= T_b(\mu)\left(\lambda^N + i\lambda^{N-1}\sum_{\alpha,n}(\sigma_a^\alpha \otimes S_n^\alpha) + \dots\right)\left((\lambda - \mu) + \frac{1}{2}i\{\mathbb{I}_a \otimes \mathbb{I}_b + \sum_\alpha(\sigma_a^\alpha \otimes \sigma_b^\alpha)\}\right) \tag{2.3.47}$$

We see from this equation that the highest order term vanishes. Let us therefore look at the second highest order, λ^N, terms. If we factor out $i\lambda^N$, use the shorthand notation $S_T^\alpha \equiv \sum_n S_n^\alpha$ and carefully write out all the terms in the tensor space $V_a \otimes V_b \otimes h$, we obtain

$$\left(\sum_\alpha(\sigma_a^\alpha \otimes \mathbb{I}_b \otimes S_T^\alpha)\right)(\mathbb{I}_a \otimes T_b(\mu)) + \frac{1}{2}\sum_\alpha(\sigma_a^\alpha \otimes \sigma_b^\alpha \otimes \mathbb{I}_h)(\mathbb{I}_a \otimes T_b(\mu)) =$$

$$= (\mathbb{I}_a \otimes T_b(\mu))\left(\sum_\alpha(\sigma_a^\alpha \otimes \mathbb{I}_b \otimes S_T^\alpha)\right) + (\mathbb{I}_a \otimes T_b(\mu))\frac{1}{2}\sum_\alpha(\sigma_a^\alpha \otimes \sigma_b^\alpha \otimes \mathbb{I}_h) \tag{2.3.48}$$

which can be written as

$$\sum_{\alpha} \sigma_a^{\alpha} \otimes \left[T_b(\mu) , \mathbb{I}_b \otimes S_T^{\alpha} + \frac{1}{2} \left(\sigma_b^{\alpha} \otimes \mathbb{I}_h \right) \right] = 0 \tag{2.3.49}$$

which implies

$$\left[T_b(\mu) , \mathbb{I}_b \otimes S_T^{\alpha} + \frac{1}{2} \left(\sigma_b^{\alpha} \otimes \mathbb{I}_h \right) \right] = \left[\left(\begin{array}{cc} A(\mu) & B(\mu) \\ C(\mu) & D(\mu) \end{array} \right), \left(\begin{array}{cc} S_T^{\alpha} & 0 \\ 0 & S_T^{\alpha} \end{array} \right) + \frac{1}{2} \left(\sigma_b^{\alpha} \otimes \mathbb{I}_h \right) \right] = 0 \tag{2.3.50}$$

Taking $\alpha = z$ gives

$$\left[\left(\begin{array}{cc} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{array} \right), \left(\begin{array}{cc} S_T^z & 0 \\ 0 & S_T^z \end{array} \right) + \frac{1}{2} \left(\begin{array}{cc} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{array} \right) \right] = 0 \tag{2.3.51}$$

and the $(1, 2)$ matrix entry of this equation gives the relation

$$[S_T^z, B] = -B \tag{2.3.52}$$

In the same way we get the relation

$$[S_T^+, B] = A - D \tag{2.3.53}$$

by taking $\alpha = x$ and $\alpha = y$. Now notice that for our pseudo-vacuum $|0\rangle$ we have

$$S_T^+|0\rangle = 0 \quad \text{and} \quad S_T^z|0\rangle = \frac{N}{2}|0\rangle \tag{2.3.54}$$

So $|0\rangle$ is a highest weight vector for the spin algebra. Let us now look at the eigenvectors $|\lambda_1, \ldots, \lambda_K\rangle$. Using (2.3.52) and (2.3.53) we see that

$$S_T^z|\lambda_1, \ldots, \lambda_K\rangle = \left(\frac{N}{2} - K \right)|\lambda_1, \ldots, \lambda_K\rangle \tag{2.3.55}$$

and

$$
\begin{aligned}
S_T^+|\lambda_1, \ldots, \lambda_K\rangle &= \sum_{k=1}^{K} B(\lambda_1) \ldots B(\lambda_{k-1})(A(\lambda_k) - D(\lambda_k))B(\lambda_{k+1}) \ldots B(\lambda_K)|0\rangle \\
&= \sum_{j=1}^{K} O_j B(\lambda_1) \ldots B(\lambda_{j-1})\hat{B}(\lambda_j)B(\lambda_{j+1}) \ldots B(\lambda_K)|0\rangle
\end{aligned}
\tag{2.3.56}
$$

To calculate O_j we use the same sort of argument as for M_j and N_j; the only contributions to O_j will come from $B(\lambda_1) \ldots B(\lambda_{k-1})(A(\lambda_k) - D(\lambda_k))B(\lambda_{k+1}) \ldots B(\lambda_K)|0\rangle$ with $k \leq j$. If $k = j$ this contribution will be

$$\prod_{n=j+1}^{K} \frac{\lambda_j - \lambda_n - i}{\lambda_j - \lambda_n} \left(\lambda_j + \frac{i}{2} \right)^N - \prod_{n=j+1}^{K} \frac{\lambda_j - \lambda_n + i}{\lambda_j - \lambda_n} \left(\lambda_j - \frac{i}{2} \right)^N \tag{2.3.57}$$

and if $k < j$ the contribution will be

$$M_j(\lambda_k, \{\lambda_i\}_{i=k+1}^K) + N_j(\lambda_k, \{\lambda_i\}_{i=k+1}^K) \tag{2.3.58}$$

So in total we get

$$
\begin{aligned}
O_j &= \prod_{k=j+1}^{K} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k}\left(\lambda_j + \frac{i}{2}\right)^N + \sum_{k=1}^{j-1} M_j(\lambda_k, \{\lambda_i\}_{i=k+1}^K) \\
&\quad - \prod_{k=j+1}^{K} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k}\left(\lambda_j - \frac{i}{2}\right)^N + \sum_{k=1}^{j-1} N_j(\lambda_k, \{\lambda_i\}_{i=k+1}^K) \\
&= \prod_{q=j+1}^{K} \frac{\lambda_j - \lambda_q - i}{\lambda_j - \lambda_q}\left(\lambda_j + \frac{i}{2}\right)^N \left(1 + \sum_{k=1}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p}\right) \\
&\quad - \prod_{q=j+1}^{K} \frac{\lambda_j - \lambda_q + i}{\lambda_j - \lambda_q}\left(\lambda_j - \frac{i}{2}\right)^N \left(1 - \sum_{k=1}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p + i}{\lambda_j - \lambda_p}\right)
\end{aligned}
\tag{2.3.59}
$$

where we have used that

$$
\begin{aligned}
\sum_{k=1}^{j-1} M_j(\lambda_k, \{\lambda_i\}_{i=k+1}^K) &= \left(\lambda_j + \frac{i}{2}\right)^N \sum_{k=1}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{\substack{p=k+1 \\ p \neq j}}^{K} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} \\
&= \left(\lambda_j + \frac{i}{2}\right)^N \prod_{q=j+1}^{K} \frac{\lambda_j - \lambda_q - i}{\lambda_j - \lambda_q} \sum_{k=1}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p}
\end{aligned}
\tag{2.3.60}
$$

and an analogous expression for $N_j(\lambda_k, \{\lambda_i\}_{i=k+1}^K)$. Now note that

$$t_n \equiv 1 + \sum_{k=n}^{j-1} \frac{i}{\lambda_k - \lambda_j} \prod_{p=k+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \prod_{k=n}^{j-1} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k} \tag{2.3.61}$$

We will prove this by induction over n. For $n = j - 1$ and $n = j - 2$ we clearly have

$$
\begin{aligned}
t_{j-1} &= 1 + \frac{i}{\lambda_{j-1} - \lambda_j} = \frac{\lambda_j - \lambda_{j-1} - i}{\lambda_j - \lambda_{j-1}} \\
t_{j-2} &= 1 + \frac{i}{\lambda_{j-1} - \lambda_j} + \frac{i}{\lambda_{j-2} - \lambda_j} \frac{\lambda_j - \lambda_{j-1} - i}{\lambda_j - \lambda_{j-1}} = \frac{\lambda_j - \lambda_{j-1} - i}{\lambda_j - \lambda_{j-1}} \frac{\lambda_j - \lambda_{j-2} - i}{\lambda_j - \lambda_{j-2}}
\end{aligned}
\tag{2.3.62}
$$

Now suppose that our formula holds for $n = l$, then we have

$$t_{l-1} = t_l + \frac{i}{\lambda_{l-1} - \lambda_j} \prod_{p=l}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \left(1 + \frac{i}{\lambda_{l-1} - \lambda_j}\right) \prod_{p=l}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \prod_{p=l-1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p}$$

$$\tag{2.3.63}$$

which proves (2.3.61) and with the same formula we obtain

$$1 + \sum_{i=1}^{j-1} \frac{i}{\lambda_i - \lambda_j} \prod_{p=i+1}^{j-1} \frac{\lambda_j - \lambda_p - i}{\lambda_j - \lambda_p} = \prod_{k=1}^{j-1} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k} \tag{2.3.64}$$

In the same way we can prove that

$$1 - \sum_{i=1}^{j-1} \frac{i}{\lambda_i - \lambda_j} \prod_{p=i+1}^{j-1} \frac{\lambda_j - \lambda_p + i}{\lambda_j - \lambda_p} = \prod_{k=1}^{j-1} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k} \tag{2.3.65}$$

So using the Bethe equations we obtain

$$O_j = \left(\lambda_j + \frac{i}{2}\right)^N \prod_{\substack{k=1 \\ k \neq j}}^{K} \frac{\lambda_j - \lambda_k - i}{\lambda_j - \lambda_k} - \left(\lambda_j - \frac{i}{2}\right)^N \prod_{\substack{k=1 \\ k \neq j}}^{K} \frac{\lambda_j - \lambda_k + i}{\lambda_j - \lambda_k} = 0 \tag{2.3.66}$$

And this proves that the eigenvectors obtained from the ABA are highest weight vectors of the spin algebra.

Finally, let us calculate the eigenvalues for the corresponding eigenvectors. Using (2.3.23), we obtain after straight forward differentiation of (2.3.41):

$$E = -J\left(\frac{i}{2} \frac{d}{d\lambda} \ln \Lambda(\lambda)\Big|_{i/2} - \frac{N}{4}\right) = E_0 + J \sum_{j=1}^{N} \frac{1}{2} \frac{1}{\lambda_j^2 + 1/4} \tag{2.3.67}$$

If we now use the parameterization $\lambda_k = \frac{1}{2} \cot\left(\frac{k_j}{2}\right)$ we get

$$E - E_0 = J \sum_{j=1}^{N} \frac{2}{\cot(k_j/2)^2 + 1} = J \sum_{j=1}^{N} 2\sin^2\left(\frac{k_j}{2}\right) = J \sum_{j=1}^{N}(1 - \cos k_j) \tag{2.3.68}$$

and we see that this agrees with the eigenvalues from the CBA.

Let us summarize some of the important observations made during the application of the Bethe ansatz procedure. First of all, the XXX model has $\mathfrak{su}(2)$ symmetry which resulted into the fact that the eigenvectors calculated using the Bethe ansatz, splits into irreducible representations of $\mathfrak{su}(2)$. However, the way in which this result is proven differs between the ABA and the CBA. The ABA would only produce eigenvectors which were highest weight vectors of $\mathfrak{su}(2)$, whereas the CBA would also produce eigenvectors which were descendants of the highest weight vectors. This was caused by the way the pseudo-momenta $\{k_i\}$ and the Bethe roots $\{\lambda_i\}$ are related to each other. A second observation is that the ABA enabled us to prove the integrability of the XXX model, which is a feature the CBA could not give us. But in the end, the calculated eigenvalues and eigenvectors from the ABA and the CBA agree with each other which gives us more evidence of the correctness of both methods.

2.4 The Nested Coordinate Bethe Ansatz and the Hubbard Model

The nested coordinate Bethe ansatz (NCBA) is a generalized coordinate Bethe ansatz developed by C. N. Yang. We shall introduce this method using the one dimensional Hubbard model following closely to the article of Sutherland [26] and [12]. We will end this section by giving an alternative characterization of the NCBA following the methods displayed in [5] and [9].

The one dimensional Hubbard model is a spin chain model with Hamiltonian

$$H = -\sum_{i=1}^{N} \sum_{s \in \{\uparrow,\downarrow\}} \left(b_{i+1,s}^{\dagger} b_{i,s} + b_{i,s}^{\dagger} b_{i+1,s} \right) + 2c \sum_{i=1}^{N} b_{i,\uparrow}^{\dagger} b_{i,\uparrow} b_{i,\downarrow}^{\dagger} b_{i,\downarrow} \qquad (2.4.1)$$

where N is the length of the spin chain and $b_{i,s}^{\dagger}$, $b_{i,s}$ are the creation and annihilation operators respectively for an electron at site i with spin $s \in \{\uparrow, \downarrow\}$. These operators satisfy the anticommutation relations

$$\{b_{j,s}, b_{l,t}\} = \{b_{j,s}^{\dagger}, b_{l,t}^{\dagger}\} = 0, \qquad \{b_{j,s}^{\dagger}, b_{l,t}\} = \delta_{jl} \delta_{st} \quad \text{where} \quad \{A, B\} \equiv AB + BA \qquad (2.4.2)$$

and H acts on the module that is spanned by vectors of the form

$$\prod_{j=1}^{K} b_{n_j,s_j}^{\dagger} |0\rangle \quad \text{where} \quad |0\rangle = \bigotimes_{i=1}^{N} |0\rangle_i \quad \text{and} \quad 1 \le n_j \le N \qquad (2.4.3)$$

The pseudo-vacuum $|0\rangle_i$ is annihilated by the annihilation operators $b_{i,s}$. So physically, a lattice site i can be unoccupied ($|0\rangle_i$), occupied by a spin up electron ($b_{i,\uparrow}^{\dagger}|0\rangle$), occupied by a spin down electron ($b_{i,\downarrow}^{\dagger}|0\rangle$) or occupied by both a spin up and a spin down electron ($b_{i,\downarrow}^{\dagger} b_{i,\uparrow}^{\dagger}|0\rangle$).

To solve the Hubbard model we first remark that the Hamiltonian preserves the number of electrons:

$$[H, \hat{N}] = 0 \quad \text{where} \quad \hat{N} = \sum_{i=1}^{N} \left(b_{i,\uparrow}^{\dagger} b_{i,\uparrow} + b_{i,\downarrow}^{\dagger} b_{i,\downarrow} \right) \qquad (2.4.4)$$

Therefore, it is enough to consider the eigenvalue problem for a fixed number of particles and we will look for eigenvectors of the form

$$|\Psi\rangle = \sum_{\{1 \le n_i \le N\}} \phi(\mathbf{n}) \prod_{j=1}^{K_1} b_{n_j,s_j}^{\dagger} |0\rangle \quad \text{with} \quad \phi(\mathbf{n}) \equiv \phi(\mathbf{n}, \mathbf{s}) = \phi(n_1, \ldots, n_{K_1} ; s_1, \ldots, s_{K_1})$$

$$(2.4.5)$$

where n_i is the position of electron i with spin s_i and K_1 is the number of particles (we will omit the spin label $\mathbf{s}$ when there is no chance of confusion). Furthermore, we will require

the wavefunction $\phi(\mathbf{n}, \mathbf{s})$ to satisfy the Pauli exclusion principle;

$$\phi(\mathbf{n}_R, \mathbf{s}_R) \equiv \phi(n_{R(1)}, \ldots, n_{R(K_1)}, s_{R(1)}, \ldots, s_{R(K_1)}) = \text{sign}(R)\phi(\mathbf{n}, \mathbf{s}) \quad \text{for every} \quad R \in S_{K_1} \tag{2.4.6}$$

This means that it is enough to treat the eigenvalue problem for one particular particle ordering $\bar{D}_Q$, defined by

$$\bar{D}_Q \equiv \{1 \le n_{Q(1)} \le n_{Q(2)} \le \ldots \le n_{Q(K_1)} \le N\} \qquad \text{where} \quad Q \in S_{K_1} \tag{2.4.7}$$

We will define the restriction of $|\Psi\rangle$ to $\bar{D}_Q$ as

$$|\Psi\rangle_Q \equiv \sum_{\mathbf{n} \in \bar{D}_Q} \phi_Q(\mathbf{n}) \prod_{j=1}^{K_1} b_{n_j, s_j}^\dagger |0\rangle \tag{2.4.8}$$

where $\phi_Q = \phi|_{\bar{D}_Q}$ for which we will use the ansatz:

$$\phi_Q(\mathbf{n}) \equiv \phi_Q(\mathbf{n}, \mathbf{s}, \mathbf{k}) = \sum_{P \in S_{K_1}} \text{sign}(PQ) A(\mathbf{k}_P | \mathbf{s}_Q) \exp\left(i \sum_{j=1}^{K_1} k_{P(j)} n_{Q(j)} \right) \tag{2.4.9}$$

Here, $A(\mathbf{k}_P | \mathbf{s}_Q)$ is a coefficient which is to be determined by the boundary conditions on ϕ, and k_i is the pseudo-momentum introduced for i-th particle. Note that with this definition the Pauli principle is satisfied.

Now let us introduce the region D_Q, which is defined as

$$D_Q \equiv \{1 \le n_{Q(1)} < n_{Q(2)} < \ldots < n_{Q(K_1)} \le N\} \qquad \text{where} \quad Q \in S_{K_1} \tag{2.4.10}$$

If $\mathbf{n} \in D_Q$, then the local interaction term of the Hamiltonian (the second term of (2.4.1)) acts as zero, and from the eigenvalue equation $E|\Psi\rangle = H|\Psi\rangle$ we get an analogous expression of (2.2.23) for ϕ_Q:

$$E\phi_Q(\mathbf{n}) = -\sum_{i=1}^{K_1} \sum_{\sigma=\pm1} \phi_Q(n_1, \ldots, n_i + \sigma, \ldots, n_{K_1}) \tag{2.4.11}$$

If we now plug in the expression for ϕ_Q we obtain

$$E = -2 \sum_{j=1}^{K_1} \cos k_j \tag{2.4.12}$$

Let us now look at the situation when two particles are on top of each other. As in the section of the XXX model with the CBA, this will yield the relations between coefficients $A(\mathbf{k}_P | \mathbf{s}_Q)$.

First of all, consider the situation that there are two particles at the same place, $n_{Q(j)} = n_{Q(j+1)} = n$, while the other particles remain well separated; $n_{Q(i+1)} > n_{Q(i)}$ for all $i \ne j$. Secondly, let us define Q' as the permutation obtained from Q by interchanging $Q(j)$ and

$Q(j + 1)$, i.e. $Q' = QP_{j,j+1}$ with $P_{j,j+1}$ being the permutation operator. Then, since ϕ_Q is a smooth function, it should obey the continuity equation $\phi_Q(\mathbf{n}, \mathbf{s}) = \phi_{Q'}(\mathbf{n}, \mathbf{s})$. Furthermore, due to the eigenvalue equation $E|\Psi\rangle = H|\Psi\rangle$, ϕ should also satisfy the the following constraint:

$$E\phi(\mathbf{n}) = -\sum_{i=1}^{K_1} \sum_{\sigma=\pm 1} \phi(n_1, \ldots, n_i + \sigma, \ldots, n_{K_1}) + 2c\phi_Q(\mathbf{n}) \qquad (2.4.13)$$

or, in terms of ϕ_Q and $\phi_{Q'}$:

$$
\begin{aligned}
E\phi_Q(\mathbf{n}) = &- \sum_{\substack{i=1 \\ i \neq Q(j), Q(j+1)}}^{K_1} \sum_{\sigma=\pm 1} \phi_Q(n_1, \ldots, n_i + \sigma, \ldots, n_{K_1}) \\
&- \phi_Q(\ldots, n_{Q(j)} - 1, \ldots) - \phi_Q(\ldots, n_{Q(j+1)} + 1, \ldots) \\
&- \phi_{Q'}(\ldots, n_{Q(j+1)} - 1, \ldots) - \phi_{Q'}(\ldots, n_{Q(j)} + 1, \ldots) + 2c\phi_Q(\mathbf{n})
\end{aligned}
\qquad (2.4.14)
$$

If we now extend formula (2.4.11) to the situation that $n_{Q(j)} = n_{Q(j+1)} = n$, and subtract (2.4.14) from it, we obtain the constraint equation

$$
\begin{aligned}
\phi_{Q'}(\ldots, n_{Q(j+1)} - 1, \ldots) &+ \phi_{Q'}(\ldots, n_{Q(j)} + 1, \ldots) \\
&- \phi_Q(\ldots, n_{Q(j)} + 1, \ldots) - \phi_Q(.., n_{Q(j+1)} - 1, ..) = 2c\phi_Q(\mathbf{n})
\end{aligned}
\qquad (2.4.15)
$$

So combined with the continuity equation, we now have a pair of constraint equations for each pair of particles on the same site. We will try to satisfy these constraints by considering the following pair of terms in the expression for $\phi_Q(n_1, \ldots, n_{K_1})$:

$$\text{sign}(PQ)A(\mathbf{k}_P|\mathbf{s}_Q) \exp\left(i \sum_{i=1}^{K_1} k_{P(i)} n_{Q(i)}\right) \quad \text{and} \quad \text{sign}(P'Q)A(\mathbf{k}_{P'}|\mathbf{s}_Q) \exp\left(i \sum_{i=1}^{K_1} k_{P'(i)} n_{Q(i)}\right)$$
$$(2.4.16)$$

For $\phi_{Q'}(n_1, \ldots, n_{K_1})$ we will consider the same kind expression. Here, P' is equal to the permutation P with $P(j)$ and $P(j + 1)$ interchanged, i.e. $P' = PP_{j,j+1}$. So our constraint equations become

$$
\begin{aligned}
&A(\mathbf{k}_P|\mathbf{s}_Q) - A(\mathbf{k}_{P'}|\mathbf{s}_Q) = A(\mathbf{k}_{P'}|\mathbf{s}_{Q'}) - A(\mathbf{k}_P|\mathbf{s}_{Q'}) \qquad \text{and} \\
&A(\mathbf{k}_P|\mathbf{s}_Q)\left(2c + e^{ik_{P(j)}} + e^{-ik_{P(j+1)}}\right) - A(\mathbf{k}_{P'}|\mathbf{s}_Q)\left(2c + e^{-ik_{P(j)}} + e^{ik_{P(j+1)}}\right) = \\
&\qquad = -A(\mathbf{k}_P|\mathbf{s}_{Q'})\left(e^{-ik_{P(j)}} + e^{ik_{P(j+1)}}\right) + A(\mathbf{k}_{P'}|\mathbf{s}_{Q'})\left(e^{ik_{P(j)}} + e^{-ik_{P(j+1)}}\right)
\end{aligned}
\qquad (2.4.17)
$$

If we now solve $A(\mathbf{k}_{P'}|\mathbf{s}_Q)$ and $A(\mathbf{k}_{P'}|\mathbf{s}_{Q'})$ in terms of $A(\mathbf{k}_P|\mathbf{s}_Q)$ and $A(\mathbf{k}_P|\mathbf{s}_{Q'})$ we will get

$$
\begin{aligned}
A(\mathbf{k}_{P'}|\mathbf{s}_Q) &= \frac{cA(\mathbf{k}_P|\mathbf{s}_Q) - i\left(\sin(k_{P(j)}) - \sin(k_{P(j+1)})\right)A(\mathbf{k}_P|\mathbf{s}_{Q'})}{c - i\left(\sin(k_{P(j)}) - \sin(k_{P(j+1)})\right)} \\[2mm]
A(\mathbf{k}_{P'}|\mathbf{s}_{Q'}) &= \frac{cA(\mathbf{k}_P|\mathbf{s}_{Q'}) - i\left(\sin(k_{P(j)}) - \sin(k_{P_{j+1}})\right)A(\mathbf{k}_P|\mathbf{s}_Q)}{c - i\left(\sin(k_{P(j)}) - \sin(k_{P(j+1)})\right)}
\end{aligned}
\qquad (2.4.18)
$$

Notice that for fixed $\mathbf{k}_P$ the amplitude $A(\mathbf{k}_P|\mathbf{s}_Q)$ is a function of K_1 spin variables $s_j \in \{\uparrow, \downarrow\}$. Therefore, we can view $A(\mathbf{k}_P|\mathbf{s}_Q)$ as the components of a 2^{K_1} dimensional vector

$$A(\mathbf{k}_P) \equiv |\mathbf{k}_P\rangle \equiv \sum_{\sigma_1,\ldots,\sigma_{K_1}=\uparrow,\downarrow} A(\mathbf{k}_P|\sigma)|\sigma\rangle \tag{2.4.19}$$

where the spin basis σ is defined as (2.1.9). We see that with this definition we have

$$A(\mathbf{k}_P|\mathbf{s}_Q) = \langle \mathbf{s}_Q|\mathbf{k}_P\rangle \tag{2.4.20}$$

Furthermore, let us note that there is a naturel action of the symmetric group on our spin basis:

$$Q|\sigma\rangle = |\sigma_{Q^{-1}}\rangle \tag{2.4.21}$$

This defines a unitary representation of the symmetric group since

$$P\left(Q|\sigma\rangle\right) = P|\sigma_{Q^{-1}}\rangle = |\sigma_{Q^{-1}P^{-1}}\rangle = |\sigma_{(PQ)^{-1}}\rangle = (PQ)|\sigma\rangle \quad \text{and}$$
$$\langle\sigma|Q^\dagger|\sigma'\rangle = \langle\sigma_{Q^{-1}}|\sigma'\rangle = \delta_{\sigma_{Q^{-1}(1)},\sigma'_1}\cdots\delta_{\sigma_{Q^{-1}(K_1)},\sigma'_{K_1}} = \delta_{\sigma_1,\sigma'_{Q(1)}}\cdots\delta_{\sigma_{K_1},\sigma'_{Q(K_1)}} \tag{2.4.22}$$
$$= \langle\sigma|\sigma'_Q\rangle = \langle\sigma|Q^{-1}|\sigma'\rangle \quad \Rightarrow \quad Q^\dagger = Q^{-1}$$

Using these definitions and writing $\theta_i \equiv \sin(k_i)$, (2.4.18) becomes

$$A(\mathbf{k}_{P'}) = A(\mathbf{k}_{PP_{j,j+1}}) = S_{j,j+1}(\theta_{P(j)} - \theta_{P(j+1)})A(\mathbf{k}_P) \quad \text{where} \quad S_{a,b}(\theta) \equiv \frac{c - i\theta P_{a,b}}{c - i\theta} \tag{2.4.23}$$

We recognize $S_{a,b}(\theta)$ as the S-matrix of the model and it is easy to see that it satisfies the relations

$$S_{n,m}^{a,b}S_{m,n}^{a,b} = 1 \quad \text{and} \quad S_{m,l}^{a,b}S_{n,l}^{b,c}S_{n,m}^{a,b} = S_{n,m}^{b,c}S_{n,l}^{a,b}S_{m,l}^{b,c} \quad \text{where} \quad S_{n,m}^{a,b} \equiv S_{a,b}(\theta_n - \theta_m) \tag{2.4.24}$$

When there is no chance for confusion, we will also refer to θ as pseudo-momentum.

Before we move on to the boundary conditions, let us introduce the so called "transmission representation" for $A(\mathbf{k}_P)$ and $S_{a,b}(\theta)$. In this representation we use the objects $A'(\mathbf{k}_P)$ and $T_{a,b}$ which are defined as follows. We define the vector $A'(\mathbf{k}_P)$ as

$$A'(\mathbf{k}_P) \equiv PA(\mathbf{k}_P) \quad \text{with} \quad A'(\mathbf{k}_I) = A(\mathbf{k}_I) \tag{2.4.25}$$

Now suppose that $P' = PP_{l,l+1}$ and let $j = P(l)$, $k = P(l+1)$. Then

$$A'(\mathbf{k}_{P'}) = P'A(\mathbf{k}_{P'}) = P'S_{l,l+1}(\theta_j - \theta_k)A(\mathbf{k}_P) = P'S_{l,l+1}(\theta_j - \theta_k)P^{-1}A'(\mathbf{k}_P) \tag{2.4.26}$$

If we now write the S-matrix as

$$S_{l,l+1}(\theta_j - \theta_k) = r_{j,k} + t_{j,k}P_{l,l+1} \quad \text{where} \quad r_{j,k} = \frac{c}{c - i(\theta_j - \theta_k)} \text{ and } t_{j,k} = \frac{-i(\theta_j - \theta_k)}{c - i(\theta_j - \theta_k)} \tag{2.4.27}$$

then we see that

$$PP_{l,l+1}S_{l,l+1}(\theta_j - \theta_k)P^{-1} = t_{j,k} + r_{j,k}PP_{l,l+1}P^{-1} = t_{j,k} + r_{j,k}P_{j,k} \tag{2.4.28}$$

Therefore, we will define the operator $T_{j,k}$ as

$$T_{j,k} \equiv T(\theta_j - \theta_k) = PP_{l,l+1}S_{l,l+1}(\theta_j - \theta_k)P^{-1} = t_{j,k} + r_{j,k}P_{j,k} \tag{2.4.29}$$

and we see that

$$A'(\mathbf{k}_{P'}) = T_{j,k}A'(\mathbf{k}_P) \tag{2.4.30}$$

Furthermore, note that $T_{j,k}$ can also be written as

$$T_{j,k} = P_{j,k}S_{j,k}(\theta_j - \theta_k) \tag{2.4.31}$$

and therefore, we do not need to index both the θ's and the permutations anymore.

We can now impose the periodic boundary conditions on ϕ. Suppose that $\mathbf{n} \in \bar{D}_Q$ and let $j = Q(K_1)$. Then we require the wave function ϕ to satisfy the boundary condition

$$\phi(n_1, \ldots, n_j - N, \ldots, n_{K_1}) = \phi(\mathbf{n}) \tag{2.4.32}$$

Let us write $\mathbf{m} \equiv (n_1, \ldots, n_j - N, \ldots, n_{K_1})$ and notice that $m_{Q(K_1)} \leq m_{Q(1)} \leq \ldots \leq m_{Q(K_1-1)}$, which may also be equivalently stated as $m_{QU(1)} \leq \ldots \leq m_{QU(K_1)}$ where $U = P_{K_1-1,K_1} \ldots P_{1,2}$. Therefore, $\mathbf{m} \in \bar{D}_{QU}$ and (2.4.32) can be written as $\phi_{QU}(\mathbf{m}) = \phi_Q(\mathbf{n})$. If we now write $\phi_{QU}(\mathbf{m})$ as

$$\phi_{QU}(\mathbf{m}) = \sum_{P \in S_{K_1}} \text{sign}(PQU) A(\mathbf{k}_P | \mathbf{s}_{QU}) \exp\left(i \sum_{l=1}^{K_1} k_{P(l)} m_{QU(l)}\right)$$

$$= \sum_{P \in S_{K_1}} \text{sign}(PQ) A(\mathbf{k}_{PU} | \mathbf{s}_{QU}) \exp\left(i \sum_{l=1}^{K_1} k_{PU(l)} m_{QU(l)}\right) \tag{2.4.33}$$

$$= \sum_{P \in S_{K_1}} \text{sign}(PQ) A(\mathbf{k}_{PU} | \mathbf{s}_{QU}) e^{-ik_{P(K_1)}N} \exp\left(i \sum_{l=1}^{K_1} k_{P(l)} n_{Q(l)}\right)$$

and compare this with the expression for $\phi_Q(\mathbf{n})$, we see that the boundary conditions are satisfied if

$$e^{ik_{P(K_1)}N} A(\mathbf{k}_P | \mathbf{s}_Q) = A(\mathbf{k}_{PU} | \mathbf{s}_{QU}) \quad \text{or written differently:}$$
$$e^{ik_{P(K_1)}N} \langle \mathbf{s}_Q | \mathbf{k}_P \rangle = \langle \mathbf{s}_{QU} | \mathbf{k}_{PU} \rangle = \langle \mathbf{s}_Q | U | \mathbf{k}_{PU} \rangle \tag{2.4.34}$$

If we now define $U_r = P_{K_1-1,K_1} \ldots P_{r,r+1}$ (so that $U = U_1$) and write this condition in vector notation with (2.4.23) we will get

$$e^{ik_{P(K_1)}N} A(\mathbf{k}_P) = UA(\mathbf{k}_{PU}) = US_{1,2}(\theta_{PU_2(1)} - \theta_{PU_2(2)})A(\mathbf{k}_{PU_2})$$
$$= US_{1,2}(\theta_{PU_2(1)} - \theta_{PU_2(2)})S_{2,3}(\theta_{PU_3(2)} - \theta_{PU_3(3)}) \ldots S_{K_1-1,K_1}(\theta_{P(K_1-1)} - \theta_{P(K_1)})A(\mathbf{k}_P)$$
$$= P^{-1}PUS_{1,2}(\theta_{PU_2(1)} - \theta_{PU_2(2)})(PU_2)^{-1}(PU_2)S_{2,3}(\theta_{PU_3(2)} - \theta_{PU_3(3)})(PU_3)^{-1}(PU_3) \ldots$$
$$\ldots (PU_{K_1-1})^{-1}(PU_{K_1-1})S_{K_1-1,K_1}(\theta_{P(K_1-1)} - \theta_{P(K_1)})P^{-1}PA(\mathbf{k}_P)$$
$$= P^{-1}T(\theta_{P(1)} - \theta_{P(K_1)})T(\theta_{P(2)} - \theta_{P(K_1)}) \ldots T(\theta_{P(K_1-1)} - \theta_{P(K_1)})A'(\mathbf{k}_P) \tag{2.4.35}$$

Multiplying this expression by P gives an equation in the transmission representation:

$$e^{ik_{P(K_1)}N}A'(\mathbf{k}_P) = T(\theta_{P(1)} - \theta_{P(K_1)}) \ldots T(\theta_{P(K_1-1)} - \theta_{P(K_1)})A'(\mathbf{k}_P) \tag{2.4.36}$$

This is an eigenvalue problem which is referred to as the "auxiliary eigenvalue problem". To solve this equation we will use a generalized Bethe ansatz construction. Therefore, the name "nested coordinate Bethe ansatz" is perfectly suited, since a second Bethe ansatz construction is used (nested) inside the original Bethe ansatz construction.

2.4.1 The Nested Coordinate Bethe Ansatz I

Let us consider the eigenvalue equation

$$\lambda_{P(K_1)}|\Psi\rangle = T_{P(1),P(K_1)} \ldots T_{P(K_1-1),P(K_1)}|\Psi\rangle \tag{2.4.37}$$

It is easy to see that $A'(\mathbf{k}_P) = |\uparrow \ldots \uparrow\rangle$ is an eigenvector of this equation with eigenvalue $\lambda = 1$. Now suppose that we have one down spin while the rest remain an up spin. This is also called the one-magnon problem, and we will look for eigenvectors of the form

$$|\Psi\rangle = \sum_{i=1}^{K_1} \psi_{P(i)}|P(i)\rangle \quad \text{with} \quad |P(i)\rangle \equiv |\uparrow \ldots \uparrow\downarrow\uparrow \ldots \uparrow\rangle \tag{2.4.38}$$

where $\downarrow$ is sitting on the $P(i)$ position and $\psi_{P(i)}$ is a wavefunction to be determined. For this vector we see that

$$
\begin{aligned}
T_{P(K_1-1),P(K_1)}|\Psi\rangle &= \sum_{\substack{i=1 \\ i\neq K_1,K_1-1}}^{K_1} \psi_{P(i)}|P(i)\rangle + \psi_{P(K_1)}\Big(t_{P(K_1-1),P(K_1)}|P(K_1)\rangle + r_{P(K_1-1),P(K_1)}|P(K_1-1)\rangle\Big) \\
&\quad + \psi_{P(K_1-1)}\Big(t_{P(K_1-1),P(K_1)}|P(K_1-1)\rangle + r_{P(K_1-1),P(K_1)}|P(K_1)\rangle\Big) \\
&= \sum_{\substack{i=1 \\ i\neq K_1,K_1-1}}^{K_1} \psi_{P(i)}|P(i)\rangle + \psi_{P(K_1)}^{(1)}|P(K_1)\rangle \\
&\quad + \Big(\psi_{P(K_1-1)}t_{P(K_1-1),P(K_1)} + \psi_{P(K_1)}r_{P(K_1-1),P(K_1)}\Big)|P(K_1-1)\rangle
\end{aligned}
$$
$$\text{where} \quad \psi_{P(K_1)}^{(1)} \equiv \psi_{P(K_1)}t_{P(K_1-1),P(K_1)} + \psi_{P(K_1-1)}r_{P(K_1-1),P(K_1)} \tag{2.4.39}$$

The amplitude of $|P(K_1-1)\rangle$ is now $\psi_{P(K_1-1)}t_{P(K_1-1),P(K_1)} + \psi_{P(K_1)}r_{P(K_1-1),P(K_1)}$ and notice that the subsequent iterations $T(\theta_{P(1)} - \theta_{P(K_1)}) \ldots T(\theta_{P(K_1-2)} - \theta_{P(K_1)})$ wont change this anymore. Therefore, for the $|P(K_1-1)\rangle$ component of $|\Psi\rangle$ we have

$$\lambda_{P(K_1)}\psi_{P(K_1-1)} = \psi_{P(K_1-1)}t_{P(K_1-1),P(K_1)} + \psi_{P(K_1)}r_{P(K_1-1),P(K_1)} \tag{2.4.40}$$

We proceed with the next iteration:

$$
\begin{aligned}
T_{P(K_1-2),P(K_1)}T_{P(K_1-1),P(K_1)}|\Psi\rangle &= \sum_{i\neq K_1,K_1-1,K_1-2}\psi_{P(i)}|P(i)\rangle \\
&+ \left(\psi_{P(K_1-1)}t_{P(K_1-1),P(K_1)} + \psi_{P(K_1)}r_{P(K_1-1),P(K_1)}\right)|P(K_1-1)\rangle \\
&+ \psi_{P(K_1-2)}t_{P(K_1-2),P(K_1)}|P(K_1-2)\rangle + \psi_{P(K_1-2)}r_{P(K_1-2),P(K_1)}|P(K_1)\rangle \\
&+ \psi^{(1)}_{P(K_1)}r_{P(K_1-2),P(K_1)}|P(K_1-2)\rangle + \psi^{(1)}_{P(K_1)}t_{P(K_1-2),P(K_1)}|P(K_1)\rangle \qquad (2.4.41) \\
&= \sum_{i\neq K_1,K_1-1,K_1-2}\psi_{P(i)}|P(i)\rangle + \left(\psi_{P(K_1-1)}t_{P(K_1-1),P(K_1)} + \psi_{P(K_1)}r_{P(K_1-1),P(K_1)}\right)|P(K_1-1)\rangle \\
&+ \left(\psi_{P(K_1-2)}t_{P(K_1-2),P(K_1)} + \psi^{(1)}_{P(K_1)}r_{P(K_1-2),P(K_1)}\right)|P(K_1-2)\rangle + \psi^{(2)}_{P(K_1)}|P(K_1)\rangle
\end{aligned}
$$

with $\psi^{(2)}_{P(K_1)} = \psi_{P(K_1-2)}r_{P(K_1-2),P(K_1)} + \psi^{(1)}_{P(K_1)}t_{P(K_1-2),P(K_1)}$. The amplitude of $|P(K_1-2)\rangle$ is now $\psi_{P(K_1-2)}t_{P(K_1-2),P(K_1)} + \psi^{(1)}_{P(K_1)}r_{P(K_1-2),P(K_1)}$ and again notice that this amplitude wont be changed anymore by the subsequent iterations $T(\theta_{P(1)} - \theta_{P(K_1)})\dots T(\theta_{P(K_1-3)} - \theta_{P(K_1)})$. Therefore, the eigenvalue equation for the $|P(K_1-2)\rangle$ component reads

$$
\begin{aligned}
\lambda_{P(K_1)}\psi_{P(K_1-2)} &= \psi_{P(K_1-2)}t_{P(K_1-2),P(K_1)} + \psi^{(1)}_{P(K_1)}r_{P(K_1-2),P(K_1)} \qquad \Leftrightarrow \\
\lambda_{P(K_1)} &= t_{P(K_1-2),P(K_1)} + r_{P(K_1-2),P(K_1)}\frac{\psi^{(1)}_{P(K_1)}}{\psi_{P(K_1-2)}}
\end{aligned}
\qquad (2.4.42)
$$

The recursive nature of this scheme becomes apparent and we obtain:

$$
\begin{aligned}
\lambda_{P(K_1)} &= t_{P(K_1-j),P(K_1)} + r_{P(K_1-j),P(K_1)}\frac{\psi^{(j-1)}_{P(K_1)}}{\psi_{P(K_1-j)}} \\
\psi^{(j)}_{P(K_1)} &= \psi_{P(K_1-j)}r_{P(K_1-j),P(K_1)} + \psi^{(j-1)}_{P(K_1)}t_{P(K_1-j),P(K_1)}
\end{aligned}
\qquad (2.4.43)
$$

with $j = 1, 2, \dots, K_1 - 1$. With the first equation, we can find an expression for $\psi^{(j-1)}_k$:

$$
\psi^{(j-1)}_{P(K_1)} = \frac{\lambda_{P(K_1)} - t_{P(K_1-j),P(K_1)}}{r_{P(K_1-j),P(K_1)}}\psi_{P(K_1-j)}
\qquad (2.4.44)
$$

Plugging this expression into the second equation gives:

$$
\begin{aligned}
\frac{\lambda_{P(K_1)} - t_{P(K_1-j-1),P(K_1)}}{r_{P(K_1-j-1),P(K_1)}}\psi_{P(K_1-j-1)} &= \left(r_{P(K_1-j),P(K_1)} + t_{P(K_1-j),P(K_1)}\frac{\lambda_{P(K_1)} - t_{P(K_1-j),P(K_1)}}{r_{P(K_1-j),P(K_1)}}\right)\psi_{P(K_1-j)} \\
&= \left(\frac{t_{P(K_1-j),P(K_1)}\lambda_k + \Delta_{P(K_1-j),P(K_1)}}{r_{P(K_1-j),P(K_1)}}\right)\psi_{P(K_1-j)} \qquad (2.4.45)
\end{aligned}
$$

with

$$
\begin{aligned}
\Delta_{P(K_1-j),P(K_1)} &\equiv \left(r_{P(K_1-j),P(K_1)}\right)^2 - \left(t_{P(K_1-j),P(K_1)}\right)^2 \\
&= \left(r_{P(K_1-j),P(K_1)} + t_{P(K_1-j),P(K_1)}\right)\left(r_{P(K_1-j),P(K_1)} - t_{P(K_1-j),P(K_1)}\right) \qquad (2.4.46) \\
&= \left(r_{P(K_1-j),P(K_1)} - t_{P(K_1-j),P(K_1)}\right)
\end{aligned}
$$

So we obtain:

$$\frac{\psi_{P(K_1-j-1)}}{\psi_{P(K_1-j)}} = \frac{t_{P(K_1-j),P(K_1)}\lambda_{P(K_1)} + \Delta_{P(K_1-j),P(K_1)}}{\lambda_{P(K_1)} - t_{P(K_1-j-1),P(K_1)}} \frac{r_{P(K_1-j-1),P(K_1)}}{r_{P(K_1-j),P(K_1)}} \tag{2.4.47}$$

$$= \frac{t_{P(K_1-j),P(K_1)}\left(\lambda_{P(K_1)} - 1\right) + r_{P(K_1-j),P(K_1)}}{\lambda_{P(K_1)} - t_{P(K_1-j-1),P(K_1)}} \frac{r_{P(K_1-j-1),P(K_1)}}{r_{P(K_1-j),P(K_1)}} \tag{2.4.48}$$

If we now plug in the expressions for t and r we obtain

$$\frac{\psi_{P(K_1-j-1)}}{\psi_{P(K_1-j)}} = \frac{\frac{-\left(\lambda_{P(K_1)}-1\right)i\left(\theta_{P(K_1-j)}-\theta_{P(K_1)}\right)+c}{c-i\left(\theta_{P(K_1-j)}-\theta_{P(K_1)}\right)}}{\lambda_{P(K_1)} + \frac{i\left(\theta_{P(K_1-j-1)}-\theta_{P(K_1)}\right)}{c-i\left(\theta_{P(K_1-j-1)}-\theta_{P(K_1)}\right)}} \cdot \frac{\left(\frac{c}{c-i\left(\theta_{P(K_1-j-1)}-\theta_{P(K_1)}\right)}\right)}{\left(\frac{c}{c-i\left(\theta_{P(K_1-j)}-\theta_{P(K_1)}\right)}\right)}$$

$$= \frac{-\left(\lambda_{P(K_1)}-1\right)i\left(\theta_{P(K_1-j)}-\theta_{P(K_1)}\right)+c}{-\left(\lambda_{P(K_1)}-1\right)i\left(\theta_{P(K_1-j-1)}-\theta_{P(K_1)}\right)+c\lambda_{P(K_1)}} \tag{2.4.49}$$

$$= \frac{-c - i\theta_{P(K_1-j)} + \left(i\theta_{P(K_1)} + \frac{c\lambda_{P(K_1)}}{\lambda_{P(K_1)}-1}\right)}{-i\theta_{P(K_1-j-1)} + \left(i\theta_{P(K_1)} + \frac{c\lambda_{P(K_1)}}{\lambda_{P(K_1)}-1}\right)}$$

Now because the left hand side of this equation depends only on $K_1 - j$, the right hand side should also only depend on $K_1 - j$ and not on K_1 alone. Therefore we conclude that

$$i\mu + \frac{c}{2} \equiv i\theta_{P(K_1)} + \frac{c\lambda_{P(K_1)}}{\lambda_{P(K_1)} - 1} \tag{2.4.50}$$

must be constant. Subsequently, we can use this definition to obtain $\lambda_{P(K_1)}$. So

$$\frac{\psi_{P(K_1-j-1)}}{\psi_{P(K_1-j)}} = \frac{i\mu - i\theta_{P(K_1-j)} - \frac{c}{2}}{i\mu - i\theta_{P(K_1-j-1)} + \frac{c}{2}} \quad \text{and} \quad \lambda_{P(K_1)} = \frac{i\mu - i\theta_{P(K_1)} + \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}} \tag{2.4.51}$$

If we now normalize our eigenvector $|\Psi\rangle$ such that $\psi_{P(1)} = 1$, we can find $\psi_{P(i)}$ from the recursion relation:

$$\frac{\psi_{P(i-1)}}{\psi_{P(i)}} = \frac{i\mu - i\theta_{P(i)} - \frac{c}{2}}{i\mu - i\theta_{P(i-1)} + \frac{c}{2}} \quad \Leftrightarrow \quad \psi_{P(i)} \equiv \psi_{P(i)}(\mu) = \prod_{l=1}^{i-1} \frac{i\mu - i\theta_{P(l)} + \frac{c}{2}}{i\mu - i\theta_{P(l+1)} - \frac{c}{2}} \tag{2.4.52}$$

for all $2 \leq i \leq K_1$. Note that we also can write $\psi_{P(i)}$ as

$$\psi_{P(i)} = \frac{i\mu - i\theta_{P(1)} - \frac{c}{2}}{i\mu - i\theta_{P(i)} - \frac{c}{2}} \prod_{l=1}^{i-1} \lambda_{P(l)} \tag{2.4.53}$$

From this equation we obtain

$$\psi_{P(K_1)} = \psi_{P(K_1-j-1)} \frac{i\mu - i\theta_{P(K_1-j-1)} - \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}} \prod_{l=K_1-j-1}^{K_1-1} \lambda_{P(l)} \tag{2.4.54}$$

Notice also that (2.4.44) can now be written as

$$
\begin{aligned}
\frac{\psi_{P(K_1)}^{(j)}}{\psi_{P(K_1-j-1)}} &= \left(\lambda_{P(K_1)} + \frac{i\left(\theta_{P(K_1-j-1)} - \theta_{P(K_1)}\right)}{c - i\left(\theta_{P(K_1-j-1)} - \theta_{P(K_1)}\right)} \right) \frac{c - i\left(\theta_{P(K_1-j-1)} - \theta_{P(K_1)}\right)}{c} \\
&= \frac{-\left(\lambda_{P(K_1)} - 1\right) i\left(\theta_{P(K_1-j-1)} - \theta_{P(K_1)}\right) + c\lambda_{P(K_1)}}{c} \\
&= \frac{-i\left(\theta_{P(K_1-j-1)} - \theta_{P(K_1)}\right) + \frac{c\lambda_{P(K_1)}}{\left(\lambda_{P(K_1)}-1\right)}}{\frac{c}{\left(\lambda_{P(K_1)}-1\right)}} = \frac{i\mu - i\theta_{P(K_1-j-1)} + \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}}
\end{aligned}
\tag{2.4.55}
$$

since

$$
\lambda_k - 1 = \frac{i\mu - i\theta_k + \frac{c}{2}}{i\mu - i\theta_k - \frac{c}{2}} - \frac{i\mu - i\theta_k - \frac{c}{2}}{i\mu - i\theta_k - \frac{c}{2}} = \frac{c}{i\mu - i\theta_k - \frac{c}{2}}
\tag{2.4.56}
$$

If we now combine this result with (2.4.54) we will obtain

$$
\psi_{P(K_1)}^{(j)} = \frac{i\mu - i\theta_{P(K_1-j-1)} + \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}} \frac{i\mu - i\theta_{P(K_1)} - \frac{c}{2}}{i\mu - i\theta_{P(K_1-j-1)} - \frac{c}{2}} \frac{\psi_{P(K_1)}}{\prod_{l=K_1-j-1}^{K_1-1} \lambda_{P(l)}} = \frac{\psi_{P(K_1)}}{\prod_{l=K_1-j}^{K_1-1} \lambda_{P(l)}}
\tag{2.4.57}
$$

Let us now focus on the $P(K_1)$ component of the eigenvalue equation. From our construction above, the eigenvalue equation will be satisfied for each component of $|\Psi\rangle$ after performing all the iterations, except for the $P(K_1)$ component. To ensure that this component also satisfies our eigenvalue equation, we see that we have to require

$$
\lambda_{P(K_1)}\psi_{P(K_1)} = \psi_{P(K_1)}^{(K_1-1)} \quad \Leftrightarrow \quad \lambda_{P(K_1)}\psi_{P(K_1)} = \frac{\psi_{P(K_1)}}{\prod_{l=1}^{K_1-1} \lambda_{P(l)}} \quad \Leftrightarrow \quad \prod_{l=1}^{K_1} \lambda_{P(l)} = 1 \tag{2.4.58}
$$

We have now obtained the complete Bethe equations for the one magnon problem:

$$
e^{ik_{P(K_1)}N} = \lambda_{P(K_1)} = \frac{i\mu - i\theta_{P(K_1)} + \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}}, \qquad \prod_{l=1}^{K_1} \lambda_{P(l)} = 1 \tag{2.4.59}
$$

Before we continue, let us remark that we can view the one-magnon eigenvector as a particle with pseudo-momentum μ, on a spin chain of length K_1 with up spins as lattice sites. Or, in other words, the one-magnon eigenvector can be viewed as an excitation of the pseudo-vacuum $|\uparrow \ldots \uparrow\rangle$.

Using the above results, we can consider the two-magnon problem

$$
\Lambda_k|\Psi\rangle = T_{P(1),P(K_1)} \ldots T_{P(K_1-1),P(K_1)}|\Psi\rangle \tag{2.4.60}
$$

The eigenvector we will try is:

$$|\Psi\rangle = \sum_{1 \leq n_1 < n_2 \leq K_1} \psi_{P(n_1),P(n_2)}|P(n_1), P(n_2)\rangle \quad \text{with} \quad |k_1, k_2\rangle \equiv |\uparrow \dots \uparrow\downarrow\uparrow \dots \uparrow\downarrow\uparrow \dots \uparrow\rangle$$

(2.4.61)

where the $\downarrow$'s are sitting on the k_1 and k_2 position respectively. For ψ_{k_1,k_2} we will use the (generalized) Bethe ansatz

$$\psi_{k_1,k_2} \equiv \psi_{k_1,k_2}(\mu, \mu') = B\psi_{k_1}(\mu)\psi_{k_2}(\mu') + B'\psi_{k_1}(\mu')\psi_{k_1}(\mu)$$

(2.4.62)

For notational convenience, we will denote $\psi_k(\mu)$ and $\psi_k(\mu')$ with ψ_k and ψ'_k respectively. We see that after the first iteration we get

$$T_{P(K_1-1),P(K_1)}|\Psi\rangle = \sum_{1 \leq n_1 < n_2 \leq K_1-2} \psi_{P(n_1),P(n_2)}|P(n_1), P(n_2)\rangle + \psi_{P(K_1-1),P(K_1)}|P(K_1 - 1), P(K_1)\rangle$$

$$+ \sum_{n_1=1}^{K_1-2} \psi_{P(n_1),P(K_1-1)}\Big\{t_{P(K_1-1),P(K_1)}|P(n_1), P(K_1 - 1)\rangle + r_{P(K_1-1),P(K_1)}|P(n_1), P(K_1)\rangle\Big\}$$

$$+ \sum_{n_1=1}^{K_1-2} \psi_{P(n_1),P(K_1)}\Big\{r_{P(K_1-1),P(K_1)}|P(n_1), P(K_1 - 1)\rangle + t_{P(K_1-1),P(K_1)}|P(n_1), P(K_1)\rangle\Big\}$$

(2.4.63)

We can continue with the next iteration, but since these calculations become rather repetitive and cumbersome we will just state the results instead.

After performing the second iteration, the amplitude of the $|P(K_1 - 2), P(K_1 - 1)\rangle$ component will be

$$r_{P(K_1-2),P(K_1)}\psi_{P(K_1-1),P(K_1)}$$

(2.4.64)

$$+ t_{P(K_1-2),P(K_1)} \{t_{P(K_1-1),P(K_1)}\psi_{P(K_1-2),P(K_1-1)} + r_{P(K_1-1),P(K_1)}\psi_{P(K_1-2),P(K_1)}\}$$

Notice that the amplitude of this component will not be changed anymore by subsequent iterations. Therefore, the eigenvalue equation for this component reads

$$\Lambda_{P(K_1)}\psi_{P(K_1-2),P(Ny-1)} = r_{P(K_1-2),P(K_1)}\psi_{P(K_1-1),P(K_1)}+$$

$$+ t_{P(K_1-2),P(K_1)} \{t_{P(K_1-1),P(K_1)}\psi_{P(K_1-2),P(K_1-1)} + r_{P(K_1-1),P(K_1)}\psi_{P(K_1-2),P(K_1)}\}$$

(2.4.65)

To solve this equation we will choose B and B' such that

$$\psi_{P(K_1-1),P(K_1)} = B\psi_{P(K_1)}\psi'_{P(K_1-1)}\frac{\lambda'_{P(K_1)}}{\lambda_{P(K_1-1)}} + B'\psi'_{P(K_1)}\psi_{P(K_1-1)}\frac{\lambda_{P(K_1)}}{\lambda'_{P(K_1-1)}}$$

(2.4.66)

where $\lambda' \equiv \lambda(\mu')$. If we remind ourselves the relations

$$\lambda_{P(K_1)}\psi_{P(K_1-j)} = t_{P(K_1-j),P(K_1)}\psi_{P(K_1-j)} + r_{P(K_1-j),P(K_1)}\psi_{P(K_1)}^{(j-1)}$$

$$\psi_{P(K_1)}^{(j)} = \frac{\psi_{P(K_1)}}{\lambda_{P(K_1-1)} \dots \lambda_{P(K_1-j)}}$$

(2.4.67)

then we see that

$$
\begin{aligned}
t_{P(K_1-1),P(K_1)}\psi_{P(K_1-2),P(K_1-1)} &+ r_{P(K_1-1),P(K_1)}\psi_{P(K_1-2),P(K_1)} = \\
&= t_{P(K_1-1),P(K_1)}\left(B\psi_{P(K_1-2)}\psi'_{P(K_1-1)} + B'\psi'_{P(K_1-2)}\psi_{P(K_1-1)}\right) \\
&\quad + r_{P(K_1-1),P(K_1)}\left(B\psi_{P(K_1-2)}\psi'_{P(K_1)} + B'\psi'_{P(K_1-2)}\psi_{P(K_1)}\right) \\
&= B\psi_{P(K_1-2)}\left(t_{P(K_1-1),P(K_1)}\psi'_{P(K_1-1)} + r_{P(K_1-1),P(K_1)}\psi'_{P(K_1)}\right) \\
&\quad + B'\psi'_{P(K_1-2)}\left(t_{P(K_1-1),P(K_1)}\psi_{P(K_1-1)} + r_{P(K_1-1),P(K_1)}\psi_{P(K_1)}\right) \\
&= B\psi_{P(K_1-2)}\lambda'_{P(K_1)}\psi'_{P(K_1-1)} + B\psi'_{P(K_1-2)}\lambda_{P(K_1)}\psi_{P(K_1-1)}
\end{aligned}
\tag{2.4.68}
$$

and $\psi^{(1)}_{P(K_1)} = \psi_{P(K_1)}/\lambda_{P(K_1-1)}$. Therefore, (2.4.64) becomes

$$
\begin{aligned}
r_{P(K_1-2),P(K_1)} &\left\{\frac{B\psi_{P(K_1)}\psi'_{P(K_1-1)}\lambda'_{K_1}}{\lambda_{P(K_1-1)}} + \frac{B'\psi'_{P(K_1)}\psi_{P(K_1-1)}\lambda_{P(K_1)}}{\lambda'_{P(K_1-1)}}\right\} \\
&+ t_{P(K_1-2),P(K_1)}\left\{B\psi_{P(K_1-2)}\lambda'_{P(K_1)}\psi'_{P(K_1-1)} + B\psi'_{P(K_1-2)}\lambda_{P(K_1)}\psi_{P(K_1-1)}\right\} \\
&= B\psi'_{P(K_1-1)}\lambda'_{P(K_1)}\left(\frac{r_{P(K_1-2),P(K_1)}\psi_{P(K_1)}}{\lambda_{P(K_1-1)}} + t_{P(K_1-2),P(K_1)}\psi_{P(K_1-2)}\right) \\
&+ B'\psi_{P(K_1-1)}\lambda_{P(K_1)}\left(\frac{r_{P(K_1-2),P(K_1)}\psi'_{P(K_1)}}{\lambda'_{P(K_1-1)}} + t_{P(K_1-2),P(K_1)}\psi'_{P(K_1-2)}\right) \\
&= B\psi'_{P(K_1-1)}\lambda'_{P(K_1)}\lambda_{P(K_1)}\psi_{P(K_1-2)} + B'\psi_{P(K_1-1)}\lambda_{P(K_1)}\lambda'_{P(K_1)}\psi'_{P(K_1-2)} \\
&= \lambda'_{P(K_1)}\lambda_{P(K_1)}\psi_{P(K_1-2),P(K_1-1)}
\end{aligned}
\tag{2.4.69}
$$

So the eigenvalue equation (2.4.65) becomes

$$
\begin{aligned}
\Lambda_{P(K_1)}\psi_{P(K_1-2),P(K_1-1)} &= \lambda'_{P(K_1)}\lambda_{P(K_1)}\psi_{P(K_1-2),P(K_1-1)} \\
\implies \qquad \Lambda_{P(K_1)} &= \lambda'_{P(K_1)}\lambda_{P(K_1)}
\end{aligned}
\tag{2.4.70}
$$

However, we need to ensure that this relation also holds for all the other components in order for it to be an eigenvalue. This can be done as follows. It turns out that the amplitude of the $|P(K_1 - j), P(K_1)\rangle$ component after j iterations will be of the form

$$
B\psi_{P(K_1-j)}\frac{\psi'_{P(K_1)}}{\lambda'_{P(K_1-1)}\cdots\lambda'_{P(K_1-j+1)}} + B'\psi'_{P(K_1-j)}\frac{\psi_{P(K_1)}}{\lambda_{P(K_1-1)}\lambda_{P(K_1-j+1)}}
\tag{2.4.71}
$$

and that $\lambda'_{P(K_1)}\lambda_{P(K_1)}$ is indeed an eigenvalue if it is possibly to choose B and B' such that

$$
\begin{aligned}
B\frac{\psi_{P(K_1-j)}\psi'_{P(K_1)}}{\lambda'_{P(K_1-1)}\cdots\lambda'_{P(K_1-j+1)}} &+ B'\frac{\psi'_{P(K_1-j)}\psi_{P(K_1)}}{\lambda_{P(K_1-1)}\cdots\lambda_{P(K_1-j+1)}} = \\
&= B\frac{\psi_{P(K_1)}\psi'_{P(K_1-j)}\lambda'_{P(K_1)}}{\lambda_{P(K_1-1)}\cdots\lambda_{P(K_1-j)}} + B'\frac{\psi'_{P(K_1)}\psi_{P(K_1-j)}\lambda_{P(K_1)}}{\lambda'_{P(K_1-1)}\cdots\lambda'_{P(K_1-j)}}
\end{aligned}
\tag{2.4.72}
$$

for all $1 \leq j \leq K_1 - 1$. Let us find out if this is possible. Multiplying left and right with $\lambda_{P(K_1-1)} \ldots \lambda_{P(K_1-j)} \lambda'_{P(K_1-1)} \ldots \lambda'_{P(K_1-j)}$ gives

$$\frac{B}{B'} = -\frac{\psi'_{P(K_1)} \psi_{P(K_1-j)} \lambda_{P(K_1)} \ldots \lambda_{P(K_1-j)} - \lambda_{P(K_1-j)} \psi_{P(K_1)} \psi'_{P(K_1-j)} \lambda'_{P(K_1-1)} \ldots \lambda'_{P(K_1-j)}}{\psi_{P(K_1)} \psi'_{P(K_1-j)} \lambda'_{P(K_1)} \ldots \lambda'_{P(K_1-j)} - \lambda'_{P(K_1-j)} \psi'_{P(K_1)} \psi_{P(K_1-j)} \lambda_{P(K_1-1)} \ldots \lambda_{P(K_1-j)}}$$

(2.4.73)

Let us consider the numerator and substitute (2.4.53) and (2.4.51) for ψ and λ respectively. Then we obtain

$$\lambda_{P(K_1)} \frac{i\mu' - i\theta_{P(1)} - \frac{c}{2}}{i\mu' - i\theta_{P(K_1)} - \frac{c}{2}} \frac{i\mu - i\theta_{P(1)} - \frac{c}{2}}{i\mu - i\theta_{P(K_1-j)} - \frac{c}{2}} \left(\prod_{l=1}^{K_1-1} \lambda'_{P(l)} \right) \left(\prod_{l=1}^{K_1-1} \lambda_{P(l)} \right)$$

$$- \lambda_{P(K_1-j)} \frac{i\mu' - i\theta_{P(1)} - \frac{c}{2}}{i\mu' - i\theta_{P(K_1-j)} - \frac{c}{2}} \frac{i\mu - i\theta_{P(1)} - \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}} \left(\prod_{l=1}^{K_1-1} \lambda'_{P(l)} \right) \left(\prod_{l=1}^{K_1-1} \lambda_{P(l)} \right)$$

$$= A \left[\lambda_{P(K_1)} \left(i\mu - i\theta_{P(K_1)} - \frac{c}{2} \right) \left(i\mu' - i\theta_{P(K_1-j)} - \frac{c}{2} \right) \right.$$

$$\left. - \lambda_{P(K_1-j)} \left(i\mu' - i\theta_{P(K_1)} - \frac{c}{2} \right) \left(i\mu - i\theta_{P(K_1-j)} - \frac{c}{2} \right) \right]$$

(2.4.74)

$$= A \left[\left(i\mu - i\theta_{P(K_1)} + \frac{c}{2} \right) \left(i\mu' - i\theta_{P(K_1-j)} - \frac{c}{2} \right) - \left(i\mu' - i\theta_{P(K_1)} - \frac{c}{2} \right) \left(i\mu - i\theta_{P(K_1-j)} + \frac{c}{2} \right) \right]$$

$$= Ai \left(\theta_{P(K_1-j)} - \theta_{P(K_1)} \right) (i(\mu' - \mu) - c))$$

where $\quad A = \dfrac{\left(\prod_{l=1}^{K_1-1} \lambda'_{P(l)} \right) \left(\prod_{l=1}^{K_1-1} \lambda_{P(l)} \right) \left(i\mu' - i\theta_{P(1)} - \frac{c}{2} \right) \left(i\mu - i\theta_{P(1)} - \frac{c}{2} \right)}{\left(i\mu' - i\theta_{P(K_1)} - \frac{c}{2} \right) \left(i\mu - i\theta_{P(K_1)} - \frac{c}{2} \right) \left(i\mu' - i\theta_{P(K_1-j)} - \frac{c}{2} \right) \left(i\mu - i\theta_{P(K_1-j)} - \frac{c}{2} \right)}$

It is now easy to see, by simply switching the primes and non-primes, that the denominator of (2.4.73) is given by

$$Ai \left(\theta_{P(K_1-j)} - \theta_{P(K_1)} \right) (i(\mu - \mu') - c))$$

(2.4.75)

So

$$\frac{B}{B'} = -\frac{i(\mu' - \mu) - c}{i(\mu - \mu') - c} = \frac{i(\mu - \mu') + c}{i(\mu - \mu') - c}$$

(2.4.76)

This means that it is indeed possible to find B and B' such that (2.4.72) holds.

Let us now look at the components $|P(k), P(K_1)\rangle$ with $1 \leq k \leq K_1 - 1$. For example, consider the $|P(K_1 - 1), P(K_1)\rangle$ component. Without performing all the iterations explicitly, we will just state that after j iterations, the amplitude of the $|P(K_1 - 1), P(K_1)\rangle$ component becomes

$$\frac{B \psi_{P(K_1)} \psi'_{P(K_1-1)} \lambda'_{P(K_1)}}{\lambda_{P(K_1-1)} \lambda_{P(K_1-2)} \ldots \lambda_{P(K_1-j)}} + \frac{B' \psi'_{P(K_1)} \psi_{P(K_1-1)} \lambda_{P(K_1)}}{\lambda'_{P(K_1-1)} \lambda'_{P(K_1-2)} \ldots \lambda'_{P(K_1-j)}}$$

(2.4.77)

This procedure stops after $K_1 - 1$ iterations and the resulting expression should satisfy the eigenvalue equation with eigenvalue $\lambda_{P(K_1)}\lambda'_{P(K_1)}$:

$$\left(\frac{B\psi_{P(K_1)}\psi'_{P(K_1-1)}\lambda'_{P(K_1)}}{\lambda_{P(K_1-1)}\lambda_{P(K_1-2)}\ldots\lambda_{P(1)}} + \frac{B'\psi'_{P(K_1)}\psi_{P(K_1-1)}\lambda_{P(K_1)}}{\lambda'_{P(K_1-1)}\lambda'_{P(K_1-2)}\ldots\lambda'_{P(1)}}\right)|P(K_1-1), P(K_1)\rangle \qquad (2.4.78)$$

$$= \lambda_{P(K_1)}\lambda'_{P(K_1)}\left(B\psi_{P(K_1-1)}\psi'_{P(K_1)} + B'\psi'_{P(K_1-1)}\psi_{P(K_1)}\right)|P(K_1-1), P(K_1)\rangle$$

This will yield the equations:

$$\frac{B}{\lambda_{P(K_1-1)}\lambda_{P(K_1-2)}\ldots\lambda_{P(1)}} = B'\lambda_{P(K_1)} \quad \text{and} \quad \frac{B'}{\lambda'_{P(K_1-1)}\lambda'_{P(K_1-2)}\ldots\lambda'_{P(1)}} = B\lambda'_{P(K_1)} \quad (2.4.79)$$

which, using (2.4.76), becomes

$$\prod_{j=1}^{K_1}\lambda_{P(j)} = \frac{i(\mu - \mu') + c}{i(\mu - \mu') - c} \quad \text{and} \quad \prod_{j=1}^{K_1}\lambda'_{P(j)} = \frac{i(\mu' - \mu) + c}{i(\mu' - \mu) - c} \qquad (2.4.80)$$

The equation together with

$$e^{ik_{P(K_1)}N} = \Lambda_{P(K_1)} = \lambda_{P(K_1)}\lambda'_{P(K_1)} = \frac{i\mu - i\theta_{P(K_1)} + \frac{c}{2}}{i\mu - i\theta_{P(K_1)} - \frac{c}{2}}\frac{i\mu' - i\theta_{P(K_1)} + \frac{c}{2}}{i\mu' - i\theta_{P(K_1)} - \frac{c}{2}} \qquad (2.4.81)$$

form the complete Bethe equations which solves the two-magnon case completely.

This whole scheme can be generalized for the K_2-magnon problem where we have K_2 down spins while the rest remain up spins. The calculations are straightforward but laborious. Therefore, without proof, we will state that the following:

Proposition 2.4.1. *An eigenvector of the K_2-magnon problem will be of the form*

$$|\Psi\rangle = \sum_{1\leq n_1 < \ldots < n_{K_2} \leq L} \psi_{P(n_1),\ldots,P(n_{K_2})}(\mu_1,\ldots,\mu_{K_2})|P(n_1),\ldots,P(n_{K_2})\rangle \qquad (2.4.82)$$

where $\psi_{P(n_1),\ldots,P(n_{K_2})}(\mu_1,\ldots,\mu_{K_2})$ is given by the generalized Bethe ansatz:

$$\psi_{P(n_1),\ldots,P(n_{K_2})}(\mu_1,\ldots,\mu_{K_2}) = \sum_{P \in S_{K_2}} B(P)\prod_{j=1}^{K_2}\psi_{P(n_j)}(\mu_j) \qquad (2.4.83)$$

and the corresponding eigenvalue is given by

$$\Lambda_{P(N)} = \lambda_{P(N)}(\mu_1)\ldots\lambda_{P(N)}(\mu_{K_2}) \qquad (2.4.84)$$

Writing $n = P(N)$, the Bethe ansatz equations are:

$$e^{ik_n N} = \prod_{j=1}^{K_2}\frac{i\mu_j - i\theta_n + \frac{c}{2}}{i\mu_j - i\theta_n - \frac{c}{2}} \quad \text{for all} \quad n = 1,\ldots,K_1$$

$$\prod_{l=1}^{K_1}\frac{i\mu_j - i\theta_l + \frac{c}{2}}{i\mu_j - i\theta_l - \frac{c}{2}} = \prod_{\substack{l=1 \\ l \neq j}}^{K_2}\frac{i(\mu_j - \mu_l) + c}{i(\mu_j - \mu_l) - c} \quad \text{for all} \quad j = 1,\ldots,K_2 \qquad (2.4.85)$$

We have now solved the Hubbard model completely.

2.4.2 The Nested Coordinate Bethe Ansatz II

The NCBA can also be implemented in another way. To demonstrate this, let us return to the objects $A(\mathbf{k}_P)$ and $S_{a,b}(\theta)$. The use of $A(\mathbf{k}_P)$ and $S_{a,b}(\theta)$ to describe the model is also called the "reflection representation" as opposed to the use $A'(\mathbf{k}_P)$ and $T_{j,k}$, which we called the "transmission representation". Furthermore, let us remind ourselves that

$$A(\mathbf{k}_{P_{1,2}}) = S_{1,2}(\theta_1 - \theta_2)A(\mathbf{k}_I) \quad \text{where} \quad S_{1,2}(\theta_1 - \theta_2) = r_{1,2} + t_{1,2}P_{1,2}$$

$$\text{with} \quad r_{1,2} = \frac{c}{c - i(\theta_j - \theta_k)} \quad \text{and} \quad t_{1,2} = \frac{-i(\theta_1 - \theta_2)}{c - i(\theta_1 - \theta_2)} \tag{2.4.86}$$

The NCBA can now be implemented in the following way. First we note that in principle, we only need the S-matrix to derive the Bethe equations; we do not need to consider and solve equation (2.4.37) explicitly, although the alternative description of the NCBA does provide us with a solution as we will see later. The idea is that we will construct eigenvectors of the S-matrix by introducing auxiliary periodic systems at various "levels". Each level has its own pseudo-vacuum and excitation vectors, but the higher the level, the less types of excitation it has. This idea will be worked out in more detail in the coming paragraphs.

Let us start off by describing the first level. The first level system is just the original spin chain of length N and the pseudo-vacuum is (naturally) the empty spin chain. We will denote this by

$$|0\rangle^{\mathrm{I}} \equiv |\bullet \ldots \bullet\rangle \tag{2.4.87}$$

where "$\bullet$" stands for an empty lattice point. A general level I excitation vector of this system is of the form

$$|\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_{K_1}}\rangle^{\mathrm{I}} \equiv |\bullet \ldots \bullet \sigma_{i_1} \bullet \ldots \bullet \sigma_{i_{K_1}} \bullet \ldots \bullet\rangle \quad \text{with} \quad \sigma \in \{\uparrow, \downarrow\} \tag{2.4.88}$$

where $1 \leq i_j \leq N$ denotes the position of the j-th σ and K_1 is the total number of excitations out of the level I vacuum (the number of electrons). Furthermore, it is understood that each σ_{i_j} is parameterized by the pseudo-momentum θ_j, i.e. $\sigma_{i_j} = \sigma_{i_j}(\theta_j)$.

We now move on to the second level and define our first auxiliary system. The auxiliary system is defined by a spin chain of length K_1, and the pseudo-vacuum $|0\rangle^{\mathrm{II}}$ is just the empty spin chain which is given by:

$$|0\rangle^{\mathrm{II}} \equiv \bigotimes_{i=1}^{K_1} |\uparrow\rangle_i^{\mathrm{I}} = |\uparrow \ldots \uparrow\rangle^{\mathrm{I}} \tag{2.4.89}$$

A general level II excitation vector of this system is of the form

$$|\downarrow_{l_1} \ldots \downarrow_{l_{K_2}}\rangle^{\mathrm{II}} \equiv |\uparrow \ldots \uparrow\downarrow_{l_1}\uparrow \ldots \uparrow\downarrow_{l_{K_2}}\uparrow \ldots \uparrow\rangle^{\mathrm{I}} \tag{2.4.90}$$

where $1 \leq l_j \leq K_1$ denotes the position of the j-th $\downarrow$ and K_2 is the number of excitations out of the level II vacuum; the number of "down" spins. Finally, since there are no other type

of excitations anymore, the nesting stops here.

The next step of the alternative NCBA is to consider the equation

$$S_{a,b}(\theta_a - \theta_b)|\Psi\rangle = \lambda\,|\Psi\rangle_{(a,b)}\,, \qquad \lambda \in \mathbb{C} \tag{2.4.91}$$

where $|\Psi\rangle$ is a first level excitation vector and where $|\Psi_{(a,b)}\rangle$ is the vector $|\Psi\rangle$ with the pseudo-momenta θ_a and θ_b interchanged. Since the S-matrix $S_{a,b}$ acts on a first level excitation vector, $S_{a,b}$ is sometimes also referred to as the "first level S-matrix". We will see later that there are also higher level S-matrices.

At this point, we remark that a solution $|\Psi\rangle$ to (2.4.91) is also a solution to the equation

$$e^{ik_{P(K_1)}N}|\Psi\rangle = U S_{1,2}(\theta_{PU_2(1)} - \theta_{PU_2(2)}) S_{2,3}(\theta_{PU_3(2)} - \theta_{PU_3(3)}) \dots$$
$$\dots S_{K_1-1,K_1}(\theta_{P(K_1-1)} - \theta_{P(K_1)})|\Psi\rangle \tag{2.4.92}$$

So the alternative version of the NCBA does indeed provide us with a solution to the eigenvalue equation (2.4.35). We will now concentrate on solving (2.4.91).

First, note that the level II pseudo-vacuum is a solution since

$$S_{i,j}|0\rangle^{\mathrm{II}} = s^{\mathrm{I,I}}(\theta_i, \theta_j)|0\rangle^{\mathrm{II}}_{(i,j)} \quad \text{where} \quad s^{\mathrm{I,I}}(\theta_i, \theta_j) = 1 \tag{2.4.93}$$

Here, $|0\rangle^{\mathrm{II}}_{(i,j)}$ is $|0\rangle^{\mathrm{II}}$ with θ_i and θ_j interchanged. Let us now consider the level II excitations. We will use the ansatz that a general level II one-excitation solution is given by

$$|\downarrow\rangle^{\mathrm{II}} = \sum_{i=1}^{K_1} \psi_i |\uparrow \dots \uparrow\downarrow_i\uparrow \dots \uparrow\rangle^{\mathrm{I}} \quad \text{with} \quad \psi_i = f(\theta_i) \prod_{j=1}^{i-1} s^{\mathrm{II,I}}(\theta_j) \tag{2.4.94}$$

The unknown functions $f(\theta_i)$ and $s^{\mathrm{II,I}}(\theta_j)$ will be solved by imposing the compatibility condition

$$S_{i,j}|\downarrow\rangle^{\mathrm{II}} = s^{\mathrm{I,I}}(\theta_i, \theta_j)|\downarrow\rangle^{\mathrm{II}}_{(i,j)} \tag{2.4.95}$$

The compatibility condition is a natural condition since we want to view a level II one-excitation solution as an excitation out of the level II pseudo-vacuum. This acknowledges the remark in the previous section where we noted that the one-magnon eigenvector could be viewed as an excitation out of the pseudo-vacuum $|\uparrow \dots \uparrow\rangle$. It will become even more evident in a moment, when we show that the level II one-excitation solution is parameterized by pseudo-momenta μ and $\{\theta_i\}$. So the level II one-excitation solution will have the pseudo-vacuum structure with an additional excitation structure. The compatibility condition is then to ensure that the underlying pseudo-vacuum structure is preserved.

To solve equation (2.4.95) it is sufficient to consider a spin chain with only two sites:

$$|\downarrow\rangle^{\mathrm{II}} = f(\theta_1)|\downarrow\uparrow\rangle^{\mathrm{I}} + f(\theta_2)s^{\mathrm{II,I}}(\theta_1)|\uparrow\downarrow\rangle^{\mathrm{I}}$$
$$|\downarrow\rangle^{\mathrm{II}}_{(1,2)} = f(\theta_2)|\downarrow\uparrow\rangle^{\mathrm{I}} + f(\theta_1)s^{\mathrm{II,I}}(\theta_2)|\uparrow\downarrow\rangle^{\mathrm{I}} \tag{2.4.96}$$

So the compatibility condition becomes

$$r_{1,2}(\theta_1, \theta_2)f(\theta_1) + t_{1,2}(\theta_1, \theta_2)f(\theta_2)s^{\text{II,I}}(\theta_1) = f(\theta_2)$$
$$t_{1,2}(\theta_1, \theta_2)f(\theta_1) + r_{1,2}(\theta_1, \theta_2)f(\theta_2)s^{\text{II,I}}(\theta_1) = f(\theta_1)s^{\text{II,I}}(\theta_2) \tag{2.4.97}$$

From the first equation of (2.4.97) we get

$$s^{\text{II,I}}(\theta_1) = \frac{f(\theta_2) - r_{1,2}(\theta_1, \theta_2)f(\theta_1)}{t_{1,2}(\theta_1, \theta_2)f(\theta_2)} \tag{2.4.98}$$

Since the left hand side is θ_2 independent, the right hand side should also be θ_2 independent. Therefore, we can just fix $\theta_2 = 0$ which gives:

$$s^{\text{II,I}}(\theta) = \frac{f(0) - r_{1,2}(\theta, 0)f(\theta)}{t_{1,2}(\theta, 0)f(0)} = \frac{ic(f(0) - f(\theta)) + \theta f(0)}{\theta f(\theta)} \tag{2.4.99}$$

Plugging this into the second equation and solving to $f(\theta_1)$ gives:

$$f(\theta_1) = \frac{f(0)f(\theta_2)\theta_2}{f(0)\theta_1 - (\theta_1 - \theta_2)f(\theta_2)} \tag{2.4.100}$$

Since the left hand side is independent of θ_2, the right hand side should also be independent of θ_2. Therefore differentiating both sides to θ_2 will yield the differential equation $\frac{d}{d\theta_2}[f(\theta_1)] = 0$. Solving this equation then gives

$$f(\theta) = \frac{f(0)e^K}{e^K - \theta} \quad \text{with} \quad K \in \mathbb{C} \tag{2.4.101}$$

If we now fix the values $f(0) = 1$, $e^K = \mu + ic/2$ with $\mu \in \mathbb{C}$, then

$$f(\theta) = \frac{i\mu - \frac{c}{2}}{i\mu - i\theta - \frac{c}{2}} \quad \text{and} \quad s^{\text{II,I}}(\theta) = \frac{i\mu - i\theta + \frac{c}{2}}{i\mu - i\theta - \frac{c}{2}} \tag{2.4.102}$$

We immediately identify μ as the pseudo-momentum of the level II one-excitation solution and we will write

$$|\downarrow(\mu)\rangle^{\text{II}} = \sum_{i=1}^{K_1} \psi_i(\mu)|\uparrow \ldots \uparrow \downarrow_i \uparrow \ldots \uparrow\rangle^{\text{I}} \quad \text{with} \quad \psi_i(\mu) = f(\mu, \theta_i)\prod_{j=1}^{i-1} s^{\text{II,I}}(\mu, \theta_j) \tag{2.4.103}$$

instead of (2.4.94) to emphasize the μ dependence.

Comparing these results with (2.4.53) we notice that $s^{\text{II,I}}(\theta_i)$ is exactly λ_i, but $f(\theta)$ is not the same as $\frac{i\mu - i\theta_1 - \frac{c}{2}}{i\mu - i\theta - \frac{c}{2}}$. This is caused by the fact that we imposed a normalization on the wave function (2.4.53). Discounting this normalization, we can say that both methods yield the same eigenvector.

Let us now consider the level II two-excitations. First define the vector

$$|\mu_1, \mu_2\rangle^{\text{II}} \equiv \sum_{1 \le i_1 < i_2 \le K_1} \psi_{i_1}(\mu_1)\psi_{i_2}(\mu_2)|\ldots \uparrow\downarrow_{i_1}\uparrow \ldots \uparrow\downarrow_{i_2}\uparrow \ldots\rangle^{\text{I}} \tag{2.4.104}$$

A general level II two-excitation solution is then of the form

$$| \downarrow(\mu_1), \downarrow(\mu_2) \rangle^{\mathrm{II}} = |\mu_1, \mu_2\rangle^{\mathrm{II}} + S^{\mathrm{II}}(\mu_1, \mu_2)|\mu_1, \mu_2\rangle^{\mathrm{II}} \tag{2.4.105}$$

where $S^{\mathrm{II}}(\mu_1, \mu_2)$ is the "second level S-matrix" defined by

$$S^{\mathrm{II}}(\mu_1, \mu_2)|\mu_1, \mu_2\rangle^{\mathrm{II}} = s^{\mathrm{II,II}}(\mu_1, \mu_2)|\mu_2, \mu_1\rangle^{\mathrm{II}} \tag{2.4.106}$$

Analogous to the one-excitation case, we want $| \downarrow(\mu_1), \downarrow(\mu_2) \rangle^{\mathrm{II}}$ to satisfy the compatibility condition

$$S_{i,j}| \downarrow(\mu_1), \downarrow(\mu_2) \rangle^{\mathrm{II}} = | \downarrow(\mu_1), \downarrow(\mu_2) \rangle^{\mathrm{II}}_{(i,j)} \tag{2.4.107}$$

For a spin chain of two sites this means:

$$\begin{aligned} f(\mu_1, \theta_1) f(\mu_2, \theta_2) s^{\mathrm{II,I}}(\mu_2, \theta_1) + s^{\mathrm{II,II}}(\mu_1, \mu_2) f(\mu_2, \theta_1) f(\mu_1, \theta_2) s^{\mathrm{II,I}}(\mu_1, \theta_1) = \\ = f(\mu_1, \theta_2) f(\mu_2, \theta_1) s^{\mathrm{II,I}}(\mu_2, \theta_2) + s^{\mathrm{II,II}}(\mu_1, \mu_2) f(\mu_2, \theta_2) f(\mu_1, \theta_1) s^{\mathrm{II,I}}(\mu_1, \theta_2) \end{aligned} \tag{2.4.108}$$

which gives

$$s^{\mathrm{II,II}}(\mu_1, \mu_2) = \frac{f(\mu_1, \theta_2) f(\mu_2, \theta_1) s^{\mathrm{II,I}}(\mu_2, \theta_2) - f(\mu_1, \theta_1) f(\mu_2, \theta_2) s^{\mathrm{II,I}}(\mu_2, \theta_1)}{f(\mu_2, \theta_1) f(\mu_1, \theta_2) s^{\mathrm{II,I}}(\mu_1, \theta_1) - f(\mu_2, \theta_2) f(\mu_1, \theta_1) s^{\mathrm{II,I}}(\mu_1, \theta_2)} = \frac{i(\mu_1 - \mu_2) - c}{i(\mu_1 - \mu_2) + c} \tag{2.4.109}$$

Comparing this expression with (2.4.76) we see that it is exactly its inverse.

These results can easily be generalized to level II K_2-excitations; a level II K_2-excitation solution $|\Psi\rangle^{\mathrm{II}}$ will be of the form:

$$|\Psi\rangle^{\mathrm{II}} = |\mu_1, \ldots, \mu_{K_2}\rangle^{\mathrm{II}} + \sum_{P \in S_{K_2}} \mathbf{S}^{\mathrm{II}}_P |\mu_1, \ldots, \mu_{K_2}\rangle^{\mathrm{II}}$$

$$\text{where} \quad \mathbf{S}^{\mathrm{II}}_P \equiv \prod_{l=1}^{n} S^{\mathrm{II}}_{I_l, J_l}(\mu_{P_l(I_l)}, \mu_{P_l(J_l)}) \quad \text{and} \quad P_l \equiv P_{I_{l+1}, J_{l+1}} \ldots P_{I_n, J_n}$$

$$\text{when} \quad P = P_{I_1, J_1} \ldots P_{I_n, J_n} \tag{2.4.110}$$

$$\text{and where} \quad |\mu_1, \ldots, \mu_{K_2}\rangle^{\mathrm{II}} \equiv \sum_{1 \le i_1 < \ldots < i_k \le K_1} \prod_{l=1}^{K_2} \psi_{i_l}(\mu_l)| \ldots \uparrow \downarrow_{i_1} \uparrow \ldots \uparrow \downarrow_{i_{K_2}} \uparrow \ldots \rangle^{\mathrm{I}}$$

Finally, we remark that if we write out $|\Psi\rangle^{\mathrm{II}}$ in terms of level I excitation vectors, the resulting expression will be a solution of (2.4.91).

At this point, we note that the functions $s^{\mathrm{II,II}}_{i,j}(\mu_1, \mu_2)$, $s^{\mathrm{II,I}}_{i,j}(\mu, \theta)$ and $s^{\mathrm{I,I}}_{i,j}(\theta_1, \theta_2)$ have a physical interpretation as phase factors. To see this, we remind ourselves that the excitation $\downarrow$ is parameterized by pseudo-momenta θ and μ while the excitation $\uparrow$ is only parameterized by pseudo-momentum θ. The phase factor $s^{\mathrm{II,II}}(\mu_1, \mu_2)$ then arises when $\downarrow(\mu_1)$ scatters with $\downarrow(\mu_2)$, the factor $s^{\mathrm{II,I}}(\mu, \theta)$ arises when $\downarrow(\mu)$ scatters with $\uparrow(\theta)$ and the factor $s^{\mathrm{I,I}}(\theta_1, \theta_2)$

arises when $\uparrow(\theta_1)$ scatters with $\uparrow(\theta_2)$. Before we derive the Bethe ansatz equations let us also introduce the phase factor

$$s^{\mathrm{I},0}(\theta) = e^{ip} \quad \text{where} \quad \theta = \sin(p) \tag{2.4.111}$$

This factor arises when an excitation (with pseudo-momentum p) of the original spin chain scatters with a site of the original spin chain. Finally, we note that

$$s^{B,A}(\beta_B, \alpha_A) = \left(s^{A,B}(\alpha_A, \beta_B)\right)^{-1} \quad \text{where} \quad A, B \in \{0, \mathrm{I}, \mathrm{II}\} \tag{2.4.112}$$

We can now derive the Bethe equations by imposing periodicity conditions on the auxiliary spin chains. For example, a general level II excitation is of the form

$$|\mu_1 \ldots \mu_{K_2}\rangle^{\mathrm{II}} = \sum_{1 \leq i_1 < \ldots < i_k \leq K_1} \prod_{l=1}^{K_2} \psi_{i_l}(\mu_l) |\ldots \uparrow\downarrow_{i_1}\uparrow \ldots \uparrow\downarrow_{i_{K_2}}\uparrow \ldots\rangle^{\mathrm{I}} \tag{2.4.113}$$

Periodicity then means that when we shift $\downarrow(\mu_1)$ around the spin chain, the resulting vector should still be the same. However, during the shifting we will pick up phase factors $s^{\mathrm{II},\mathrm{II}}(\mu_1, \mu_i)$ and $s^{\mathrm{II},\mathrm{I}}(\mu_1, \theta_j)$. Therefore, the periodicity condition implies that

$$\prod_{l=1}^{K_1} s^{\mathrm{II},\mathrm{I}}(\mu_1, \theta_l) \prod_{i=2}^{K_2} s^{\mathrm{II},\mathrm{II}}(\mu_1, \mu_i) = \prod_{l=1}^{K_1} \frac{i\mu_1 - i\theta_l + \frac{c}{2}}{i\mu_1 - i\theta_l - \frac{c}{2}} \prod_{i=2}^{K_2} \frac{i(\mu_1 - \mu_i) - c}{i(\mu_1 - \mu_i) + c} = 1 \tag{2.4.114}$$

By also considering the other shiftings we obtain

$$\prod_{l=1}^{K_1} \frac{i\mu_j - i\theta_l + \frac{c}{2}}{i\mu_j - i\theta_l - \frac{c}{2}} \prod_{\substack{i=1 \\ i \neq j}}^{K_2} \frac{i(\mu_j - \mu_i) - c}{i(\mu_j - \mu_i) + c} = 1 \quad \text{for all} \quad 1 \leq i \leq K_2 \tag{2.4.115}$$

The same reasoning can be applied for the level I excitations. Consider the general level I excitation (2.4.88). When we shift $\sigma(\theta_j)$ around the spin chain we will pick up phase factors $s^{\mathrm{I},\mathrm{I}}(\theta_l, \theta_n)$, $s^{\mathrm{I},\mathrm{II}}(\theta_l, \mu_j)$ and $s^{\mathrm{I},0}(\theta_l)$. Therefore, the periodicity condition implies that

$$\prod_{i=1}^{N} s^{\mathrm{I},0}(\theta_l) \prod_{\substack{n=1 \\ n \neq l}}^{K_1} s^{\mathrm{I},\mathrm{I}}(\theta_l, \theta_n) \prod_{j=1}^{K_2} s^{\mathrm{I},\mathrm{II}}(\theta_l, \mu_j) = e^{ip_l N} \prod_{j=1}^{K_2} \frac{i\mu_j - i\theta_l - \frac{c}{2}}{i\mu_j - i\theta_l + \frac{c}{2}} = 1 \tag{2.4.116}$$

When we compare (2.4.115) and (2.4.116) with (2.4.85), then we see that they coincide completely with each other. We have now solved the Hubbard model using the alternative form of the NCBA.

2.5 The Nested Algebraic Bethe Ansatz (NABA)

In the ABA for the XXX model we had a Lax operator which was a 2×2 matrix (in auxiliary space), and the corresponding monodromy matrix consisted of 4 operators. However, when

the Lax operator is a $M \times M$ matrix, the corresponding monodromy matrix will consist of M^2 operators. The ABA in such cases is then implemented in a multistage manner which is referred to as the nested algebraic Bethe ansatz (NABA).

In this section we are going to introduce the NABA using a simple example. We will focus on the main ideas of the NABA and we will skip a lot of the proofs and only give the results, since most of the calculations are rather long and tedious, and because most of the techniques involved are already highlighted in the section on the algebraic Bethe ansatz.

Let us consider a 3×3 problem, which has all the features of the general $M \times M$ case. First, let $L_{a,i}(\lambda) : V_a \otimes V_i \to V_a \otimes V_i$ be a Lax operator with $V_a \simeq V_i \simeq \mathbb{C}^3$, let $T_a(\lambda) = L_{a,N}(\lambda)L_{a,N-1}(\lambda)\ldots L_{a,1}(\lambda)$ be the corresponding monodromy matrix and suppose that this operator satisfy the fundamental commutation relation (see Appendix A.3 for the notation)

$$\bar{R}_{a,b}(\lambda - \mu)[T_a(\lambda) \otimes T_b(\mu)] = [T_a(\mu) \otimes T_b(\lambda)]\bar{R}_{a,b}(\lambda - \mu) \tag{2.5.1}$$

where $\bar{R}(\lambda)$ is given by

$$\bar{R}(\lambda) = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & a(\lambda) & 0 & b(\lambda) & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & a(\lambda) & 0 & 0 & 0 & b(\lambda) & 0 & 0 \\
0 & b(\lambda) & 0 & a(\lambda) & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & a(\lambda) & 0 & b(\lambda) & 0 \\
0 & 0 & b(\lambda) & 0 & 0 & 0 & a(\lambda) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & b(\lambda) & 0 & a(\lambda) & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \tag{2.5.2}$$

with

$$a(\lambda) = \frac{\gamma}{\gamma - 2\lambda}, \quad b(\lambda) = \frac{2\lambda}{2\lambda - \gamma} \quad \text{and} \quad \gamma = i\hbar \tag{2.5.3}$$

Next, let use write the monodromy matrix as

$$T_a(\lambda) = \begin{pmatrix}
A(\lambda) & B_2(\lambda) & B_3(\lambda) \\
C_2(\lambda) & D_2^2(\lambda) & D_2^3(\lambda) \\
C_3(\lambda) & D_3^2(\lambda) & D_3^3(\lambda)
\end{pmatrix} \tag{2.5.4}$$

and let $|0\rangle$, which is the first level pseudo-vacuum, be a vector in $V_1 \otimes \ldots \otimes V_N$ such that

$$\begin{aligned}
B_i(\lambda)|0\rangle \neq 0, \quad && C_i(\lambda)|0\rangle = 0, \quad && D_2^3(\lambda)|0\rangle = D_3^2(\lambda)|0\rangle = 0 \\
A(\lambda)|0\rangle = e^{-i\lambda\Delta}|0\rangle \quad && \text{and} \quad && D_2^2(\lambda)|0\rangle = D_3^3(\lambda)|0\rangle = e^{i\lambda\Delta}|0\rangle
\end{aligned} \tag{2.5.5}$$

with $\Delta \in \mathbb{R}$. The name "first level" pseudo-vacuum suggests that there exist higher level pseudo-vacua. And indeed, this will be the case as we will see later on.

If we now write out the FCR we will get the commutation relations:

$$
\begin{aligned}
A(\lambda)B_i(\mu) &= \frac{1}{b(\mu - \lambda)}B_i(\mu)A(\lambda) - \frac{a(\mu - \lambda)}{b(\mu - \lambda)}B_i(\lambda)A(\mu) \\
B_i(\lambda)B_j(\mu) &= \tilde{R}^{i,j}_{k,l}(\lambda - \mu)B_k(\mu)B_l(\lambda) \\
D^j_k(\lambda)B_l(\mu) &= \frac{1}{b(\lambda - \mu)}\tilde{R}^{l,j}_{i,m}(\lambda - \mu)B_m(\mu)D^i_k(\lambda) - \frac{a(\lambda - \mu)}{b(\lambda - \mu)}B_j(\lambda)D^l_k(\mu)
\end{aligned}
\tag{2.5.6}
$$

where

$$
\tilde{R}^{a_2,b_2}_{a_1,b_1}(\lambda) = a(\lambda)\delta_{a_1,a_2}\delta_{b_1,b_2} + b(\lambda)\delta_{a_1,b_2}\delta_{a_2,b_1}
\tag{2.5.7}
$$

or, written in matrix form:

$$
\tilde{R}(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a(\lambda) & b(\lambda) & 0 \\ 0 & b(\lambda) & a(\lambda) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = a(\lambda)\mathbb{I} + b(\lambda)P
\tag{2.5.8}
$$

where $\mathbb{I}$ is the identity matrix and P is the permutation matrix (2.3.11). Next, we define the transfer matrix $\tau(\lambda)$ by:

$$
\tau(\lambda) = tr_a[T(\lambda)] = A(\lambda) + D^2_2(\lambda) + D^3_3(\lambda)
\tag{2.5.9}
$$

An eigenvector of $\tau(\lambda)$ is then built by applying the creation operators $B_2(\lambda)$ and $B_3(\lambda)$ on the pseudo-vacuum $|0\rangle$ and it is of the form

$$
|\mu_1, \mu_2, \ldots, \mu_{K_1}\rangle = \mathbf{B} \cdot \mathbf{X}|0\rangle \equiv \sum_{\{a_j\}} X^{a_1,\ldots,a_{K_1}} B_{a_1}(\mu_1) \ldots B_{a_{K_1}}(\mu_{K_1})|0\rangle
$$

$$
\text{where} \quad \sum_{\{a_j\}} \equiv \sum_{2 \le a_1,\ldots,a_{K_1} \le 3}
\tag{2.5.10}
$$

where $K_1 \le N$, $\mathbf{B}$ is a vector of creation operators and $\mathbf{F}$ is a vector of functions. We will call such an eigenvector a K_1-excitation eigenvector. If we now use (2.5.6) we will find that

the operators $A(\lambda)$ and $D_i^i(\lambda)$ act on a K_1-excitation eigenvector as

$$A(\lambda)|\mu_1,\mu_2,\ldots,\mu_{K_1}\rangle = e^{-i\lambda\Delta} \prod_{l=1}^{K_1} \frac{1}{b(\mu_l - \lambda)} |\mu_1,\mu_2,\ldots,\mu_{K_1}\rangle$$

$$+ \sum_{\{a_j\},\{b_j\}} \sum_{k=1}^{K_1} \left(\tilde{\Omega}_k\right)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} X^{a_1,\ldots,a_{K_1}} B_{b_k}(\mu_k) \prod_{\substack{i=1\\i\neq k}}^{K_1} B_{b_i}(\mu_i)|0\rangle$$

$$\sum_{a=2}^{3} D_a^a|\mu_1,\mu_2,\ldots,\mu_{K_1}\rangle = e^{i\lambda\Delta} \prod_{k=1}^{K_1} \frac{1}{b(\lambda - \mu_k)} \sum_{\{a_j\},\{b_j\}} \prod_{l=1}^{K_1} B_{b_l}(\mu_l)|0\rangle \tilde{\tau}(\lambda)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} X^{a_1,\ldots,a_{K_1}}$$

(2.5.11)

$$+ \sum_{\{a_j\},\{b_j\}} \sum_{k=1}^{K_1} \left(\Omega_k\right)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} X^{a_1,\ldots,a_{K_1}} B_{b_k}(\mu_k) \prod_{\substack{i=1\\i\neq k}}^{K_1} B_{b_i}(\mu_i)|0\rangle$$

where $\quad \tilde{\tau}(\lambda)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} \equiv \tilde{\tau}(\lambda,\{\mu_i\})_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}}$

$$= \sum_{c,\{c_j\}} \tilde{R}_{c,b_{K_1}}^{a_{K_1},c_{K_1-1}}(\lambda - \mu_{K_1}) \tilde{R}_{c_{K_1-1},b_{K_1-1}}^{a_{K_1-1},c_{K_1-2}}(\lambda - \mu_{K_1-1})\ldots \tilde{R}_{c_1,b_1}^{a_1,c}(\lambda - \mu_1)$$

The "coefficients" $\tilde{\Omega}_k$ and Ω_k form the higher dimensional matrix analogue of the coefficients M_n and N_n from the ABA for the XXX model and they can be computed in the same manner as M_n and N_n. But because these expressions are rather long and cumbersome, we will omit them here. Next, notice that $\tilde{R}_{a,c}^{b,d} = \hat{R}_{a,c}^{d,b}$, where $\hat{R}(\lambda) = \tilde{R}(\lambda)P$ with P being the permutation matrix. Therefore, $\tilde{\tau}(\lambda)$ can be written as

$$\tilde{\tau}(\lambda) = tr_a[\hat{R}_{a,K_1}(\lambda - \mu_{K_1})\hat{R}_{a,K_1-1}(\lambda - \mu_{K_1-1})\ldots \hat{R}_{a,1}(\lambda - \mu_1)] \quad \text{or}$$

$$\tilde{\tau}(\lambda)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} = \sum_{c,\{c_j\}} \hat{R}_{c,b_{K_1}}^{c_{K_1-1},a_{K_1}}(\lambda - \mu_{K_1}) \hat{R}_{c_{K_1-1},b_{K_1-1}}^{c_{K_1-2},a_{K_1-1}}(\lambda - \mu_{K_1-1})\ldots \hat{R}_{c_1,b_1}^{c,a_1}(\lambda - \mu_1) \quad (2.5.12)$$

where $\hat{R}_{a,i} \in End(\bar{V}_a \otimes \bar{V}_i)$ with $\bar{V}_a$ being an auxiliary space and $\bar{V}_a \simeq \bar{V}_i \simeq \mathbb{C}^2$. The objects $\tilde{R}$ and $\tilde{\tau}(\lambda)$ are also called the "reduced" R-matrix and the reduced transfer matrix respectively.

Now, in order to make $|\mu_1,\mu_2,\ldots,\mu_{K_1}\rangle$ an eigenvector of $\tau(\lambda)$ we will require the following two constraints:

$$\sum_{\{a_j\}} \tilde{\tau}(\lambda)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} X^{a_1,\ldots,a_{K_1}} = \Lambda(\lambda,\{\mu_i\})X^{b_1,\ldots,b_{K_1}} \qquad (2.5.13)$$

$$\sum_{\{a_j\}} \left(\tilde{\Omega}_k\right)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} X^{a_1,\ldots,a_{K_1}} = -\sum_{\{a_j\}} \left(\Omega_k\right)_{b_1,\ldots,b_{K_1}}^{a_1,\ldots,a_{K_1}} X^{a_1,\ldots,a_{K_1}} \qquad (2.5.14)$$

Notice that the first constraint is just an eigenvalue problem, which we will refer to as the auxiliary eigenvalue problem analogous to (2.4.36). If these two constraints are satisfied,

then $|\mu_1, \mu_2, \ldots, \mu_{K_1}\rangle$ is an eigenvector of $\tau(\lambda)$ with eigenvalue $\Upsilon(\lambda, \{\mu_i\})$ defined by

$$\Upsilon(\lambda, \{\mu_i\}) = e^{-i\lambda\Delta} \prod_{l=1}^{K_1} \frac{1}{b(\mu_l - \lambda)} + e^{i\lambda\Delta}\Lambda(\lambda, \{\mu_i\}) \prod_{k=1}^{K_1} \frac{1}{b(\lambda - \mu_k)} \tag{2.5.15}$$

Let us now focus on solving the constraint equations. Without proof, we will just state that (2.5.14) gives

$$e^{-i\mu_k\Delta} \prod_{\substack{l=1 \\ l \neq k}}^{K_1} \frac{b(\mu_k - \mu_l)}{b(\mu_l - \mu_k)} X^{b_1, \ldots, b_{K_1}} = \sum_{\{a_j\}} \tilde{\tau}(\mu_k)_{b_1, \ldots, b_{K_1}}^{a_1, \ldots, a_{K_1}} X^{a_1, \ldots, a_{K_1}} \tag{2.5.16}$$

If we combine this equation with (2.5.13) we will obtain

$$e^{-i\mu_k\Delta} \prod_{\substack{l=1 \\ l \neq k}}^{K_1} \frac{b(\mu_k - \mu_l)}{b(\mu_l - \mu_k)} = \Lambda(\mu_k, \{\mu_i\}) \tag{2.5.17}$$

Let us now consider (2.5.13) and try to solve that equation. First of all, define the reduced monodromy matrix $\hat{T}_a(\lambda, \{\mu_i\})$

$$\hat{T}_a(\lambda, \{\mu_i\}) = \hat{R}_{a,K_1}(\lambda - \mu_{K_1})\hat{R}_{a,K_1-1}(\lambda - \mu_{K_1-1}) \ldots \hat{R}_{a,1}(\lambda - \mu_1) \tag{2.5.18}$$

(where the label a now refers to auxiliary space $\bar{V}_a$) and notice that it obeys the FCR

$$\tilde{R}_{a,b}(\lambda_1 - \lambda_2)[\hat{T}_a(\lambda_1, \{\mu_i\}) \otimes \hat{T}_b(\lambda_2, \{\mu_i\})] = [\hat{T}_a(\lambda_2, \{\mu_i\}) \otimes \hat{T}_b(\lambda_1, \{\mu_i\})]\tilde{R}_{a,b}(\lambda_1 - \lambda_2) \tag{2.5.19}$$

If we now write

$$\hat{T}(\lambda, \{\mu_i\}) = \begin{pmatrix} \hat{A}(\lambda) & \hat{B}(\lambda) \\ \hat{C}(\lambda) & \hat{D}(\lambda) \end{pmatrix} \tag{2.5.20}$$

then it is easy to see from the structure of the matrix $\hat{R}$ that the vector

$$|1\rangle \equiv \bigotimes_{i=1}^{K_1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_i \in \bar{V}_1 \otimes \ldots \otimes \bar{V}_{K_1} \tag{2.5.21}$$

which is the second level pseudo-vacuum, obeys the relations

$$\hat{A}(\lambda)|1\rangle = |1\rangle, \qquad \hat{D}(\lambda)|1\rangle = \prod_{m=1}^{K_1} b(\lambda - \mu_m)|1\rangle \tag{2.5.22}$$

and we immediately recognize the similarities with the ABA of the XXX model. In fact, if we write (2.5.19) as

$$\hat{R}_{a,b}(\lambda_1 - \lambda_2)[\hat{T}_a(\lambda_1, \{\mu_i\})\hat{T}_b(\lambda_2, \{\mu_i\})] = [\hat{T}_b(\lambda_2, \{\mu_i\})\hat{T}_a(\lambda_1, \{\mu_i\})]\hat{R}_{a,b}(\lambda_1 - \lambda_2) \tag{2.5.23}$$

and compare $\hat{R}(\lambda)$ with (2.3.15) we will see that it is of the same form. Therefore, we can follow the construction of the ABA of the XXX model, and we will obtain the following expression for the eigenvalue $\Lambda(\lambda, \{\mu_i\})$:

$$\Lambda(\lambda, \{\mu_i\}) = \prod_{j=1}^{K_2} \frac{1}{b(\tilde{\mu}_j - \lambda)} + \frac{\prod_{k=1}^{K_1} b(\lambda - \mu_k)}{\prod_{l=1}^{K_2} b(\lambda - \tilde{\mu}_l)} \quad \text{with} \quad 1 \leq K_2 \leq K_1 \tag{2.5.24}$$

Furthermore, we will also get the Bethe equations

$$\prod_{k=1}^{K_1} b(\mu_k - \tilde{\mu}_l) = \prod_{\substack{j=1 \\ j \neq l}}^{K_2} \frac{b(\tilde{\mu}_j - \tilde{\mu}_l)}{b(\tilde{\mu}_l - \tilde{\mu}_j)} \quad \text{for all} \quad l = 1, \dots, K_2 \tag{2.5.25}$$

for the reduced transfer matrix. And (2.5.17) becomes

$$e^{-i\mu_k \Delta} \prod_{\substack{l=1 \\ l \neq k}}^{K_1} \frac{b(\mu_k - \mu_l)}{b(\mu_l - \mu_k)} = \prod_{j=1}^{K_2} \frac{1}{b(\tilde{\mu}_j - \mu_k)} \quad \text{for all} \quad k = 1, \dots, K_1 \tag{2.5.26}$$

So (2.5.25) and (2.5.26) form the complete set of Bethe equations for our original 3×3 problem, with Bethe roots $\{\mu_i\}$ and $\{\tilde{\mu}_i\}$. Note that if we started with a $M^2 \times M^2$ R-matrix and applied the method outlined above, we would end up with a $(M-1)^2 \times (M-1)^2$ reduced R-matrix. This procedure can be continued until we arrive at the lowest dimensional reduced 4×4 R-matrix, hence it is called the "nested" algebraic Bethe ansatz.

2.6 The (Nested) Algebraic Bethe Ansatz for Graded Models

In this section we will find the Bethe equation of the Hubbard model using the nested algebraic Bethe ansatz. However, the Hubbard model is a so called "graded" model which makes the Bethe ansatz a bit more involved. To make this more precise, we will first explain what a graded integrable model is analogous to section 2.3.1. The notation we use is explained in appendix A.4.

2.6.1 The Graded Integrable Model

Suppose that we have a periodic spin chain of length N to which each lattice point i is attached a finite dimensional graded Hilbert space V_i. A graded quantum integrable ultralocal model on this spin chain is then characterized by a graded Lax operator $L_{a,i}(\lambda, p_i)$: $V_a \hat{\otimes} V_i \rightarrow V_a \hat{\otimes} V_i$ where i refers to the lattice point i and a is the label of the auxiliary space. The graded Lax operator satisfies the fundamental commutation relation

$$\bar{R}_{a,b}(\lambda, \mu)[L_{a,i}(\lambda, \nu) \hat{\otimes} L_{b,i}(\mu, \nu)] = [L_{a,i}(\mu, \nu) \hat{\otimes} L_{b,i}(\lambda, \nu)]\bar{R}_{a,b}(\lambda, \mu) \tag{2.6.1}$$

where the matrix elements of $\bar{R}(\lambda, \mu)$ are given by $\bar{R}^{\alpha,\beta}_{\gamma,\delta} = R^{\gamma,\beta}_{\alpha,\delta}$, with $R(\lambda, \mu)$ being a R-matrix. At this point, the attentive reader will have noticed the extra parameter p_i in the

graded Lax operator which is absent in the ordinary Lax operator defined in section 2.3. However, the extra parameter does not turn out to be of any problem as we will see in a moment. In fact, we have already seen an example of a two parameter Lax operator in equation (2.5.18).

Next we define the graded monodromy matrix

$$T_a(\lambda, \{p_i\}) \equiv L_{a,N}(\lambda, p_N) \dots L_{a,1}(\lambda, p_1) \tag{2.6.2}$$

where $L_{a,i}(\lambda, p_i)$ is seen as the induced Lax operator on $V_a \hat{\otimes} V_1 \hat{\otimes} V_2 \hat{\otimes} \dots \hat{\otimes} V_N$. Therefore, the graded monodromy matrix is an operator on $V_a \hat{\otimes} V_1 \hat{\otimes} V_2 \hat{\otimes} \dots \hat{\otimes} V_N$. If we now take the partial supertrace over the auxiliary space of this operator, we will get the graded transfer matrix [4].

$$\tau(\lambda, \{p_i\}) = \mathrm{str}_a[T_a(\lambda, \{p_i\})] \tag{2.6.3}$$

which is an operator on $V_1 \hat{\otimes} V_2 \hat{\otimes} \dots \hat{\otimes} V_N$. It is easy to see that the graded monodromy matrix satisfies the relation

$$\bar{R}_{a,b}(\lambda, \mu)[T_a(\lambda, \{p_i\}) \hat{\otimes} T_b(\mu, \{p_i\})] = [T_a(\mu, \{p_i\}) \hat{\otimes} T_b(\lambda, \{p_i\})]\bar{R}_{a,b}(\lambda, \mu) \tag{2.6.4}$$

from which we derive the following expression for the graded transfer matrix:

$$[\tau(\lambda, \{p_i\}), \tau(\mu, \{p_i\})] = 0 \tag{2.6.5}$$

Finally, we can derive the Bethe ansatz equations by solving the eigenvalue equation

$$\tau(\lambda, \{p_i\})|\Phi\rangle = \Lambda(\lambda, \{p_i\})|\Phi\rangle \tag{2.6.6}$$

2.6.2   The Nested Algebraic Bethe Ansatz for the Hubbard Model

The Hubbard spin chain is modeled by attaching to each lattice site the graded vector space $V^{(2|2)} = V_{\bar{0}} \oplus V_{\bar{1}}$, where $\{e_1 = (1,0,0,0), e_4 = (0,0,0,1)\}$ and $\{e_2 = (0,1,0,0), e_3 = (0,0,1,0)\}$ span $V_{\bar{0}}$ and $V_{\bar{1}}$ respectively. The graded Lax matrix for the Hubbard model is given by

$$L_{a,i}(\lambda) = \left(L^{\sigma}_{a_1,i_1}(\lambda)L^{\sigma}_{a_2,i_2}(\lambda)\right)e^{h(\lambda)\sigma^z_{a_1}\sigma^z_{a_2}} \quad \text{where} \quad V_a = V_{a_1} \otimes V_{a_2}, \quad V_i = V_{i_1} \otimes V_{i_2},$$

$$\sigma_{a_1} = \sigma \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I}, \quad \sigma_{a_2} = \mathbb{I} \otimes \sigma \otimes \mathbb{I} \otimes \mathbb{I}, \quad \sigma_{i_1} = \mathbb{I} \otimes \mathbb{I} \otimes \sigma \otimes \mathbb{I}, \quad \sigma_{i_2} = \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes \sigma$$

$$L^{\sigma}_{a,i}(\lambda) = \frac{a(\lambda) + b(\lambda)}{2} + \frac{a(\lambda) - b(\lambda)}{2}\sigma^z_a\sigma^z_i + \left(\sigma^+_a\sigma^-_i + \sigma^-_a\sigma^+_i\right) \tag{2.6.7}$$

---

[4]We note that at first sight, it seems that we can also take the ordinary trace instead of the supertrace. However, because of the grading, certain symmetry properties then will look unnatural and hence, the supertrace is used [12]

where V_a and V_i are graded such that $V_a \simeq V_i \simeq V^{(2|2)}$. The corresponding R-matrix, which is often called Shastry's (graded) R-matrix, is given by:

$$
\bar{R}_{1,2}^{(s)}(\lambda,\mu) \equiv
\begin{pmatrix}
\alpha_2 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & \alpha_5 & 0 & 0 & | & -i\alpha_9 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha_5 & 0 & | & 0 & 0 & 0 & 0 & | & -i\alpha_9 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha_4 & | & 0 & 0 & -i\alpha_{10} & 0 & | & 0 & i\alpha_{10} & 0 & 0 & | & \alpha_7 & 0 & 0 & 0 \\
- & - & - & - & & - & - & - & - & & - & - & - & - & & - & - & - & - \\
0 & -i\alpha_8 & 0 & 0 & | & \alpha_5 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & | & 0 & \alpha_1 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & i\alpha_{10} & | & 0 & 0 & \alpha_3 & 0 & | & 0 & -\alpha_6 & 0 & 0 & | & -i\alpha_{10} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & | & 0 & 0 & 0 & \alpha_5 & | & 0 & 0 & 0 & 0 & | & 0 & -i\alpha_8 & 0 & 0 \\
- & - & - & - & & - & - & - & - & & - & - & - & - & & - & - & - & - \\
0 & 0 & -i\alpha_8 & 0 & | & 0 & 0 & 0 & 0 & | & \alpha_5 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -i\alpha_{10} & | & 0 & 0 & -\alpha_6 & 0 & | & 0 & \alpha_3 & 0 & 0 & | & i\alpha_{10} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & \alpha_1 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & \alpha_5 & | & 0 & 0 & -i\alpha_8 & 0 \\
- & - & - & - & & - & - & - & - & & - & - & - & - & & - & - & - & - \\
0 & 0 & 0 & \alpha_7 & | & 0 & 0 & i\alpha_{10} & 0 & | & 0 & -i\alpha_{10} & 0 & 0 & | & \alpha_4 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & | & 0 & 0 & 0 & -i\alpha_9 & | & 0 & 0 & 0 & 0 & | & 0 & \alpha_5 & 0 & 0 \\
0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & -i\alpha_9 & | & 0 & 0 & \alpha_5 & 0 \\
0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & 0 & | & 0 & 0 & 0 & \alpha_2
\end{pmatrix}
$$

$$\tag{2.6.8}$$

where the $\alpha_i \equiv \alpha_i(\lambda,\mu)$ are defined as:

$$
\alpha_1 = \left(e^{[h(\mu)-h(\lambda)]} a(\lambda)a(\mu) + e^{-[h(\mu)-h(\lambda)]} b(\lambda)b(\mu) \right) \alpha_5
$$

$$
\alpha_2 = \left(e^{-[h(\mu)-h(\lambda)]} a(\lambda)a(\mu) + e^{[h(\mu)-h(\lambda)]} b(\lambda)b(\mu) \right) \alpha_5
$$

$$
\alpha_3 = \frac{e^{[h(\mu)+h(\lambda)]} a(\lambda)b(\mu) + e^{-[h(\mu)+h(\lambda)]} b(\lambda)a(\mu)}{a(\lambda)b(\lambda) + a(\mu)b(\mu)} \left(\frac{\cosh[h(\mu)-h(\lambda)]}{\cosh[h(\mu)+h(\lambda)]} \right) \alpha_5
$$

$$
\alpha_4 = \frac{e^{-[h(\mu)+h(\lambda)]} a(\lambda)b(\mu) + e^{[h(\mu)+h(\lambda)]} b(\lambda)a(\mu)}{a(\lambda)b(\lambda) + a(\mu)b(\mu)} \left(\frac{\cosh(h(\mu)-h(\lambda))}{\cosh(h(\mu)+h(\lambda))} \right) \alpha_5
$$

$$
\alpha_6 = \left(\frac{e^{[h(\mu)+h(\lambda)]} a(\lambda)b(\mu) - e^{-[h(\mu)+h(\lambda)]} b(\lambda)a(\mu)}{a(\lambda)b(\lambda) + a(\mu)b(\mu)} \right) [b^2(\mu) - b^2(\lambda)] \frac{\cosh[h(\mu)-h(\lambda)]}{\cosh[h(\mu)+h(\lambda)]} \alpha_5
$$

$$
\alpha_7 = \left(\frac{-e^{-[h(\mu)+h(\lambda)]} a(\lambda)b(\mu) + e^{[h(\mu)+h(\lambda)]} b(\lambda)a(\mu)}{a(\lambda)b(\lambda) + a(\mu)b(\mu)} \right) [b^2(\mu) - b^2(\lambda)] \frac{\cosh[h(\mu)-h(\lambda)]}{\cosh[h(\mu)+h(\lambda)]} \alpha_5
$$

$$
\alpha_8 = \left(e^{[h(\mu)-h(\lambda)]} a(\lambda)b(\mu) - e^{-[h(\mu)-h(\lambda)]} b(\lambda)a(\mu) \right) \alpha_5
$$

$$
\alpha_9 = \left(-e^{-[h(\mu)-h(\lambda)]} a(\lambda)b(\mu) + e^{[h(\mu)-h(\lambda)]} b(\lambda)a(\mu) \right) \alpha_5
$$

$$
\alpha_{10} = \frac{b^2(\mu) - b^2(\lambda)}{a(\lambda)b(\lambda) + a(\mu)b(\mu)} \left(\frac{\cosh[h(\mu)-h(\lambda)]}{\cosh[h(\mu)+h(\lambda)]} \right) \alpha_5 \tag{2.6.9}
$$

where $a(\lambda) = \cos(\lambda), b(\lambda) = \sin(\lambda)$ and where h is a λ dependent coupling satisfying the constraint

$$
\sinh[2h(\lambda)] = -\frac{a(\lambda)b(\lambda)}{4c} \tag{2.6.10}
$$

We note that the Lax operator is proportional to a graded permutation of $\bar{R}_{12}^{(s)}(\lambda,0)$. Furthermore, it satisfies equation (2.6.1) and we can form the monodromy matrix $T_a(\lambda)$ and

transfer matrix $\tau(\lambda)$. To obtain the Bethe equations we will solve the eigenvalue problem $\tau(\lambda)|\Phi\rangle = \Lambda(\lambda)|\Phi\rangle$ using the NABA. First, we write the monodromy matrix as a 4×4 matrix in auxiliary space:

$$T(\lambda) = \begin{pmatrix} B(\lambda) & \mathbf{B}(\lambda) & F(\lambda) \\ \mathbf{C}(\lambda) & \hat{A}(\lambda) & \mathbf{B}^*(\lambda) \\ C(\lambda) & \mathbf{C}^*(\lambda) & D(\lambda) \end{pmatrix} \quad \text{where} \quad \hat{A}(\lambda) = \begin{pmatrix} A_1^1(\lambda) & A_1^2(\lambda) \\ A_2^1(\lambda) & A_2^2(\lambda) \end{pmatrix}$$

$$\mathbf{B}(\lambda) = (B_1(\lambda), B_2(\lambda)), \qquad \mathbf{C}^*(\lambda) = (C_1^*(\lambda), C_2^*(\lambda))$$

$$\mathbf{C}(\lambda) = (C_1(\lambda), C_2(\lambda))^T, \qquad \mathbf{B}^*(\lambda) = (B_1^*(\lambda), B_2^*(\lambda))^T$$

(2.6.11)

So $\mathbf{B}(\lambda)$, $\mathbf{C}^*(\lambda)$ are row vectors, $\mathbf{B}^*(\lambda)$, $\mathbf{C}(\lambda)$ are column vectors and $\hat{A}(\lambda)$ is a 2×2 matrix. With these definitions, the eigenvalue equation becomes

$$\left(B(\lambda) - \sum_{i=1}^{2} A_i^i(\lambda) + D(\lambda) \right) |\Phi\rangle = \Lambda(\lambda)|\Phi\rangle$$

(2.6.12)

The (first level) pseudo-vacuum $|0\rangle$ we shall use is defined by

$$|0\rangle = \overbrace{\bigotimes_{i=1}^{N}}^{N} |0\rangle_i \quad \text{where} \quad |0\rangle_i \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}_i \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_i \in V^{(2|2)}$$

(2.6.13)

and the action of the Lax matrix on this vector is

$$L_{a,i}(\lambda)|0\rangle_i = \begin{pmatrix} \omega_1(\lambda)|0\rangle_i & * & * & * \\ 0 & \omega_2(\lambda)|0\rangle_i & 0 & * \\ 0 & 0 & \omega_2(\lambda)|0\rangle_i & * \\ 0 & 0 & 0 & \omega_3(\lambda)|0\rangle_i \end{pmatrix} \quad \text{where}$$

$$\omega_1(\lambda) = a(\lambda)^2 e^{h(\lambda)} \qquad \omega_2(\lambda) = a(\lambda)b(\lambda)e^{-h(\lambda)} \qquad \omega_3(\lambda) = b(\lambda)^2 e^{h(\lambda)}$$

(2.6.14)

which implies that

$$T_a(\lambda)|0\rangle = \begin{pmatrix} \omega_1^N(\lambda) & * & * & * \\ 0 & \omega_2^N(\lambda) & 0 & * \\ 0 & 0 & \omega_2^N(\lambda) & * \\ 0 & 0 & 0 & \omega_3^N(\lambda) \end{pmatrix} |0\rangle$$

(2.6.15)

Here, the $*$ denote an expression not relevant for our discussion. So we see that $|0\rangle$ is an eigenvector with eigenvalue

$$\Lambda(\lambda) = \omega_1^N(\lambda) - 2\omega_2^N(\lambda) + \omega_3^N(\lambda)$$

(2.6.16)

As one would expect, the other eigenvectors will be constructed by letting the (creation) operators $\mathbf{B}(\lambda)$, $\mathbf{B}^*(\lambda)$ and $F(\lambda)$ act on the eigenvector $|0\rangle$. However, note that the number of creation operators is five, while on each lattice site we only have three possible (non-vacuum) configurations; we can have a spin up electron, spin down electron or we can have a pair of electrons with opposite spins. Therefore, the eigenvectors will be build by

operators $\mathbf{B}(\lambda)$ and $F(\lambda)$, or $\mathbf{B}^*(\lambda)$ and $F(\lambda)$, but not by a general combination of all those operators. Without loss of generality, we will build our eigenvectors with $\mathbf{B}(\lambda)$ and $F(\lambda)$. Finally, we remark that the pseudo-vacuum is annihilated by the following operators:

$$\mathbf{C}(\lambda)|0\rangle = 0, \qquad \mathbf{C}^*(\lambda)|0\rangle = 0, \qquad C(\lambda)|0\rangle = 0, \qquad \hat{A}_1^2|0\rangle = \hat{A}_2^1|0\rangle = 0 \qquad (2.6.17)$$

Before we build the eigenvectors, we first need to have the commutation relations between the various operators which are derived from equation (2.6.4). For the creation operators we have:

$$\mathbf{B}(\lambda) \otimes \mathbf{B}(\mu) = \frac{\alpha_1(\lambda,\mu)}{\alpha_2(\lambda,\mu)}[\mathbf{B}(\mu) \otimes \mathbf{B}(\lambda)] \cdot \hat{r}(\lambda,\mu) - i\frac{\alpha_{10}(\lambda,\mu)}{\alpha_7(\lambda,\mu)}[F(\lambda)B(\mu) - F(\mu)B(\lambda)]\boldsymbol{\xi}$$

$$[F(\lambda), F(\mu)] = 0$$

$$F(\lambda)\mathbf{B}(\mu) = \frac{\alpha_5(\lambda,\mu)}{\alpha_2(\lambda,\mu)}F(\mu)\mathbf{B}(\lambda) - i\frac{\alpha_8(\lambda,\mu)}{\alpha_2(\lambda,\mu)}\mathbf{B}(\mu)F(\lambda)$$

$$\mathbf{B}(\lambda)F(\mu) = \frac{\alpha_5(\lambda,\mu)}{\alpha_2(\lambda,\mu)}\mathbf{B}(\mu)F(\lambda) - i\frac{\alpha_9(\lambda,\mu)}{\alpha_2(\lambda,\mu)}F(\mu)\mathbf{B}(\lambda) \qquad (2.6.18)$$

$$\text{where} \quad \boldsymbol{\xi} = (0, 1, -1, 0) \quad \text{and} \quad \hat{r}(\lambda,\mu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bar{a}(\lambda,\mu) & \bar{b}(\lambda,\mu) & 0 \\ 0 & \bar{b}(\lambda,\mu) & \bar{a}(\lambda,\mu) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{with} \quad \bar{a}(\lambda,\mu) = \frac{\alpha_3(\lambda,\mu)\alpha_7(\lambda,\mu) + \alpha_{10}^2(\lambda,\mu)}{\alpha_1(\lambda,\mu)\alpha_7(\lambda,\mu)} \quad \text{and} \quad \bar{b}(\lambda,\mu) = -\frac{\alpha_6(\lambda,\mu)\alpha_7(\lambda,\mu) + \alpha_{10}^2(\lambda,\mu)}{\alpha_1(\lambda,\mu)\alpha_7(\lambda,\mu)}$$

Here we note that using (2.6.10) and introducing the reparameterizations

$$\tilde{\lambda} = \frac{a(\lambda)}{b(\lambda)}e^{2h(\lambda)} - \frac{b(\lambda)}{a(\lambda)}e^{-2h(\lambda)} + c \qquad (2.6.19)$$

we can write $\bar{a}(\lambda,\mu)$ and $\bar{b}(\lambda,\mu)$ as

$$\bar{a}(\tilde{\lambda},\tilde{\mu}) = \frac{2c}{\tilde{\mu} - \tilde{\lambda} + 2c} \qquad \text{and} \qquad \bar{b}(\tilde{\lambda},\tilde{\mu}) = \frac{\tilde{\mu} - \tilde{\lambda}}{\tilde{\mu} - \tilde{\lambda} + 2c} \qquad (2.6.20)$$

Continuing with the list of commutation relations, for $\mathbf{B}(\lambda)$ we have:

$$\hat{A}(\lambda) \otimes \mathbf{B}(\mu) = -i\frac{\alpha_1(\lambda,\mu)}{\alpha_9(\lambda,\mu)}[\mathbf{B}(\mu) \otimes \hat{A}(\lambda)] \cdot \hat{r}(\lambda,\mu) + i\frac{\alpha_5(\lambda,\mu)}{\alpha_9(\lambda,\mu)}\mathbf{B}(\lambda) \otimes \hat{A}(\mu)$$

$$- i\frac{\alpha_{10}(\lambda,\mu)}{\alpha_7(\lambda,\mu)}\left(\mathbf{B}^*(\lambda)B(\mu) + i\frac{\alpha_5(\lambda,\mu)}{\alpha_9(\lambda,\mu)}F(\lambda)\mathbf{C}(\mu) - i\frac{\alpha_2(\lambda,\mu)}{\alpha_9(\lambda,\mu)}F(\mu)\mathbf{C}(\lambda)\right) \otimes \boldsymbol{\xi}$$

$$B(\lambda)\mathbf{B}(\mu) = i\frac{\alpha_2(\mu,\lambda)}{\alpha_9(\lambda,\mu)}\mathbf{B}(\mu)B(\lambda) - i\frac{\alpha_5(\mu,\lambda)}{\alpha_9(\lambda,\mu)}\mathbf{B}(\lambda)B(\mu)$$

$$D(\lambda)\mathbf{B}(\mu) = -i\frac{\alpha_8(\lambda,\mu)}{\alpha_7(\lambda,\mu)}\mathbf{B}(\mu)D(\lambda) + \frac{\alpha_5(\lambda,\mu)}{\alpha_7(\lambda,\mu)}F(\mu)\mathbf{C}^*(\lambda) - i\frac{\alpha_4(\lambda,\mu)}{\alpha_7(\lambda,\mu)}F(\lambda)\mathbf{C}^*(\mu)$$

$$- i\frac{\alpha_{10}(\lambda,\mu)}{\alpha_7(\lambda,\mu)}\boldsymbol{\xi} \cdot [\mathbf{B}^*(\lambda) \otimes \hat{A}(\mu)]$$

$$(2.6.21)$$

and for the operator $F(\lambda)$ we have:

$$
A_a^b(\lambda)F(\mu) = \left(1 + \frac{\alpha_5^2(\lambda,\mu)}{\alpha_9(\lambda,\mu)\alpha_8(\lambda,\mu)}\right)F(\mu)A_a^b(\lambda) - \frac{\alpha_5^2(\lambda,\mu)}{\alpha_9(\lambda,\mu)\alpha_8(\lambda,\mu)}F(\lambda)A_a^b(\mu)
$$
$$
+ i\frac{\alpha_5(\lambda,\mu)}{\alpha_9(\lambda,\mu)}[\mathbf{B}(\lambda)\otimes\mathbf{B}^*(\mu)]_a^b + i\frac{\alpha_5(\lambda,\mu)}{\alpha_8(\lambda,\mu)}[\mathbf{B}^*(\lambda)\otimes\mathbf{B}(\mu)]_a^b
$$
$$
B(\lambda)F(\mu) = \frac{\alpha_2(\mu,\lambda)}{\alpha_7(\mu,\lambda)}F(\mu)B(\lambda) - \frac{\alpha_4(\mu,\lambda)}{\alpha_7(\mu,\lambda)}F(\lambda)B(\mu) + i\frac{\alpha_{10}(\mu,\lambda)}{\alpha_7(\mu,\lambda)}[\mathbf{B}(\lambda)\otimes\mathbf{B}(\mu)]\cdot\boldsymbol{\xi}^t
$$
$$
D(\lambda)F(\mu) = \frac{\alpha_2(\lambda,\mu)}{\alpha_7(\lambda,\mu)}F(\mu)D(\lambda) - \frac{\alpha_4(\lambda,\mu)}{\alpha_7(\lambda,\mu)}F(\lambda)D(\mu) - i\frac{\alpha_{10}(\lambda,\mu)}{\alpha_7(\lambda,\mu)}\boldsymbol{\xi}\cdot[\mathbf{B}^*(\lambda)\otimes\mathbf{B}^*(\mu)]
$$

$$(2.6.22)$$

With these relations we can start with the construction of the eigenvectors. A K_1-excitation eigenvector will be of the form

$$
|\Phi_{K_1}(\lambda_1,\ldots,\lambda_{K_1})\rangle = \boldsymbol{\Phi}_{K_1}(\lambda_1,\ldots,\lambda_{K_1})\cdot\mathbf{F}|0\rangle \tag{2.6.23}
$$

where the wave vector $\boldsymbol{\Phi}_{K_1}(\lambda_1,\ldots,\lambda_{K_1})$ is a vector described in terms of the creation operators, and $\mathbf{F}$ is a vector of functions whose components are the coefficients of the linear combination of the creation operators (analogous to (2.5.10)). This statement will become more clear in a few moments, when we consider the one and two-excitation eigenvectors.

For the one-excitation eigenvector we will use the ansatz $\boldsymbol{\Phi}_1(\lambda_1) = \mathbf{B}(\lambda_1)$, so that the one-excitation eigenvector is of the form

$$
|\Phi_1(\lambda_1)\rangle = \mathbf{B}(\lambda_1)\cdot\mathbf{F}|0\rangle = \sum_{i=1}^{2}B_i(\lambda_1)F^i|0\rangle \tag{2.6.24}
$$

Then using (2.6.21) we obtain:

$$
B(\lambda)|\Phi_1(\lambda_1)\rangle = i\frac{\alpha_2(\lambda_1,\lambda)}{\alpha_9(\lambda_1,\lambda)}[\omega_1(\lambda)]^N|\Phi_1(\lambda_1)\rangle
$$
$$
- i\frac{\alpha_5(\lambda_1,\lambda)}{\alpha_9(\lambda_1,\lambda)}[\omega_1(\lambda_1)]^N\mathbf{B}(\lambda)\cdot\mathbf{F}|0\rangle
$$

$$(2.6.25)$$

$$
D(\lambda)|\Phi_1(\lambda_1)\rangle = -i\frac{\alpha_8(\lambda,\lambda_1)}{\alpha_7(\lambda,\lambda_1)}[\omega_3(\lambda)]^N|\Phi_1(\lambda_1)\rangle
$$
$$
- i\frac{\alpha_{10}(\lambda,\lambda_1)}{\alpha_7(\lambda,\lambda_1)}[\omega_2(\lambda_1)]^N\big(\boldsymbol{\xi}\cdot(\mathbf{B}^*(\lambda)\otimes\mathbb{I}_{2\times2})\big)\cdot\mathbf{F}|0\rangle
$$

$$(2.6.26)$$

$$
\sum_{i=1}^{2}A_i^i(\lambda)|\Phi_1(\lambda_1)\rangle = -i\frac{\alpha_1(\lambda,\lambda_1)}{\alpha_9(\lambda,\lambda_1)}[\omega_2(\lambda)]^N\sum_{i,j,k=1}^{2}\hat{r}_{j,i}^{i,k}(\lambda,\lambda_1)B_j(\lambda_1)F^k|0\rangle
$$
$$
+ i\frac{\alpha_5(\lambda,\lambda_1)}{\alpha_9(\lambda,\lambda_1)}[\omega_2(\lambda_1)]^N\mathbf{B}(\lambda)\cdot\mathbf{F}|0\rangle
$$
$$
- i\frac{\alpha_{10}(\lambda,\lambda_1)}{\alpha_7(\lambda,\lambda_1)}[\omega_1(\lambda_1)]^N\big(\boldsymbol{\xi}\cdot(\mathbf{B}^*(\lambda)\otimes\mathbb{I}_{2\times2})\big)\cdot\mathbf{F}|0\rangle
$$

$$(2.6.27)$$

Notice that the first term of (2.6.25) and (2.6.26) are of the "right" form and that the first term of (2.6.27) is almost of the right form. All the other terms are unwanted terms and we will get rid of them by imposing conditions on the parameters λ_1. These conditions are the well-known Bethe ansatz equations.

First, notice from (2.6.9) that

$$\frac{\alpha_5(\lambda_1, \lambda)}{\alpha_9(\lambda_1, \lambda)} = -\frac{\alpha_5(\lambda, \lambda_1)}{\alpha_9(\lambda, \lambda_1)} \tag{2.6.28}$$

Therefore, if we impose the condition

$$\left(\frac{\omega_1(\lambda_1)}{\omega_2(\lambda_1)}\right)^N = 1 \tag{2.6.29}$$

then the second term of (2.6.25) cancels against the second term of (2.6.27) and the second term of (2.6.26) cancels against the third term of (2.6.27) in the eigenvalue equation. To obtain the eigenvalue of $|\Phi_1(\lambda_1)\rangle$ we have to impose one more condition; we want $\mathbf{F}$ to be an eigenvector of $\hat{r}$:

$$\sum_{k=1}^{2} \tau^{(1)}(\lambda, \lambda_1)_j^k F^k = \Lambda^{(1)}(\lambda, \lambda_1) F^j \qquad \text{where} \quad \tau^{(1)}(\lambda, \lambda_1)_j^k = \sum_{i=1}^{2} \hat{r}_{j,i}^{i,k}(\lambda, \lambda_1) \tag{2.6.30}$$

This is the auxiliary eigenvalue problem, similar to (2.5.13), and in this case it is easy to solve since

$$\tau^{(1)}(\lambda, \lambda_1) = \begin{pmatrix} 1 + \bar{b}(\lambda, \lambda_1) & 0 \\ 0 & 1 + \bar{b}(\lambda, \lambda_1) \end{pmatrix} \tag{2.6.31}$$

which implies that

$$\Lambda^{(1)}(\lambda, \lambda_1) = 1 + \bar{b}(\lambda, \lambda_1) \tag{2.6.32}$$

So the eigenvalue $\Lambda(\lambda, \lambda_1)$ of $|\Phi_1(\lambda_1)\rangle$ is

$$\Lambda(\lambda, \lambda_1) = i\frac{\alpha_2(\lambda_1, \lambda)}{\alpha_9(\lambda_1, \lambda)}[\omega_1(\lambda)]^N - i\frac{\alpha_8(\lambda, \lambda_1)}{\alpha_7(\lambda, \lambda_1)}[\omega_3(\lambda)]^N + i\frac{\alpha_1(\lambda, \lambda_1)}{\alpha_9(\lambda, \lambda_1)}\Lambda^{(1)}(\lambda, \lambda_1)[\omega_2(\lambda)]^N \tag{2.6.33}$$

Now that we have solved the one-excitation case, let us continue with the two-excitation case. Because the two-excitation case involves much more calculations then the one-excitation case, we will only mention the key points in the construction and refer the reader to [22] or [12] for more details.

For the two-excitation case we use the ansatz

$$\mathbf{\Phi}_2(\lambda_1, \lambda_2) = \mathbf{B}(\lambda_1) \otimes \mathbf{B}(\lambda_2) + \boldsymbol{\xi}\, F(\lambda_1) B(\lambda_2) \hat{g}_0^{(2)}(\lambda_1, \lambda_2) \qquad \text{so that}$$

$$|\Phi_2(\lambda_1, \lambda_2)\rangle = \sum_{i,j=1}^{2} B_i(\lambda_1) B_j(\lambda_2) F^{i,j}|0\rangle + [\omega_1(\lambda_2)]^N F(\lambda_1) \hat{g}_0^{(2)}(\lambda_1, \lambda_2)(F^{2,1} - F^{1,2})|0\rangle \tag{2.6.34}$$

where $\hat{g}_0^{(2)}(\lambda_1, \lambda_2)$ is an arbitrary function to be determined. As in the one-excitation case, the key point in obtaining the eigenvalue and the Bethe ansatz equations is to impose conditions on λ_1 and λ_2 such that the unwanted terms cancel each other out in the eigenvalue equation.

Among the set of unwanted terms there will be the terms of the form $F(\lambda)D(\lambda_1)B(\lambda_2)$, $\mathbf{B}(\lambda) \cdot \mathbf{B}^*(\lambda_1)B(\lambda_2)$ and $\boldsymbol{\xi} \cdot [\mathbf{B}^*(\lambda_1) \otimes \mathbf{B}(\lambda_1)]B(\lambda_2)$. Without proof we will mention that these terms are canceled if we take $\hat{g}_0^{(2)}(\lambda_1, \lambda_2)$ as

$$\hat{g}_0^{(2)}(\lambda_1, \lambda_2) = i\frac{\alpha_{10}(\lambda_1, \lambda_2)}{\alpha_7(\lambda_1, \lambda_2)} \tag{2.6.35}$$

Secondly, the action of $A_1^1 + A_2^2$ on $|\Phi_2(\lambda_1, \lambda_2)\rangle$ will yield a term similar to the first term of (2.6.27):

$$\prod_{j=1}^{2} -i\frac{\alpha_1(\lambda, \lambda_j)}{\alpha_9(\lambda, \lambda_j)}[\omega_2(\lambda)]^N \sum_{\substack{a_1,a_2,b_1,b_2 \\ c_1,c_2=1}} \hat{r}_{b_1,c_1}^{c_2,a_1}(\lambda, \lambda_1)\hat{r}_{b_2,c_2}^{c_1,a_2}(\lambda, \lambda_2)B_{b_1}(\lambda_1)B_{b_2}(\lambda_2)F^{a_1,a_2}|0\rangle \tag{2.6.36}$$

Consequently, we will have to solve the following auxiliary eigenvalue problem in order to get the eigenvalue $\Lambda(\lambda, \lambda_1, \lambda_2)$ of $|\Phi_2(\lambda_1, \lambda_2)\rangle$:

$$\sum_{a_1,a_2=1}^{2} \tau^{(1)}(\lambda, \lambda_1, \lambda_2)_{b_1,b_2}^{a_1,a_2}F^{a_1,a_2} \equiv \sum_{\substack{a_1,a_2 \\ c_1,c_2=1}}^{2} \hat{r}_{b_1,c_1}^{c_2,a_1}(\lambda, \lambda_1)\hat{r}_{b_2,c_2}^{c_1,a_2}(\lambda, \lambda_2)F^{a_1,a_2} = \Lambda^{(1)}(\lambda, \lambda_1, \lambda_2)F^{b_1,b_2}$$

$$\tag{2.6.37}$$

To solve this problem, let us consider the more general problem

$$\sum_{a_1,\dots,a_{K_1}} \tau^{(1)}(\lambda, \{\lambda_i\}_{i=1}^{K_1})_{b_1,\dots,b_{K_1}}^{a_1,\dots,a_{K_1}}F^{a_1,\dots,a_{K_1}} = \Lambda^{(1)}(\lambda, \{\lambda_i\}_{i=1}^{K_1})F^{b_1,\dots,b_{K_1}} \qquad \text{where}$$

$$\tau^{(1)}(\lambda, \{\lambda_i\}_{i=1}^{K_1})_{a_1,\dots,a_{K_1}}^{b_1,\dots,b_{K_1}} \equiv \sum_{c_1,\dots,c_{K_1}} \hat{r}_{b_1,c_1}^{c_{K_1},a_1}(\lambda, \lambda_1)\hat{r}_{b_2,c_2}^{c_1,a_2}(\lambda, \lambda_2)\dots\hat{r}_{b_{K_1},c_{K_1}}^{c_{K_1-1},a_{K_1}}(\lambda, \lambda_{K_1})$$

$$\tag{2.6.38}$$

If we introduce the matrix $\bar{r} \equiv P\hat{r}$ with P being the permutation matrix in $\mathbb{C}^2 \otimes \mathbb{C}^2$, then we notice that $\bar{r}_{a,c}^{b,d} = \hat{r}_{c,a}^{b,d}$. Therefore, $\tau^{(1)}(\lambda, \{\lambda_i\})$ can be written as

$$\tau^{(1)}(\lambda, \{\lambda_i\}) = tr_a[\bar{r}_{a,K_1}(\lambda, \lambda_{K_1})\bar{r}_{a,K_1-1}(\lambda, \lambda_{K_1-1})\dots\bar{r}_{a,1}(\lambda, \lambda_1)] \quad \text{or}$$

$$\tau^{(1)}(\lambda, \{\lambda_i\})_{b_1,\dots,b_{K_1}}^{a_1,\dots,a_{K_1}} = \sum_{c_1,\dots,c_{K_1}} \bar{r}_{c_{K_1},b_{K_1}}^{c_{K_1-1},a_{K_1}}(\lambda, \lambda_{K_1})\bar{r}_{c_{K_1-1},b_{K_1-1}}^{c_{K_1-2},a_{K_1-1}}(\lambda, \lambda_{K_1-1})\dots\bar{r}_{c_1,b_1}^{c_{K_1},a_1}(\lambda, \lambda_1) \tag{2.6.39}$$

where $\hat{r}_{a,i} \in End(\bar{V}_a \otimes \bar{V}_i)$ with $\bar{V}_a$ being an auxiliary space and $\bar{V}_a \simeq \bar{V}_i \simeq \mathbb{C}^2$. In this form we immediately recognize the nesting procedure of section 2.5. To solve this auxiliary eigenvalue problem we have to apply a second algebraic Bethe ansatz. Since the matrix $\bar{r}$

is of the same form as the reduced R-matrix of section 2.5, we will not repeat all the steps again and merely state the results; the eigenvalue $\Lambda^{(1)}(\lambda, \{\lambda_i\})$ is given by

$$\Lambda^{(1)}(\lambda, \{\lambda_i\}) = \prod_{j=1}^{K_2} \frac{1}{\bar{b}(\mu_j, \lambda)} + \frac{\prod_{j=1}^{K_1} \bar{b}(\lambda, \lambda_j)}{\prod_{l=1}^{K_2} \bar{b}(\lambda, \mu_l)} \quad \text{with} \quad 1 \leq K_2 \leq K_1 \tag{2.6.40}$$

and the corresponding Bethe equations are

$$\prod_{l=1}^{K_1} \bar{b}(\mu_j, \lambda_l) = \prod_{l=1}^{K_2} \frac{\bar{b}(\mu_j, \mu_l)}{\bar{b}(\mu_l, \mu_j)} \quad \text{for all} \quad j = 1, \ldots, K_2 \tag{2.6.41}$$

In fact, this generalized auxiliary eigenvalue problem is actually the auxiliary eigenvalue problem for the K_1-excitation case.

Returning to our two-excitation problem, we see that we can simply obtain the solution of the auxiliary eigenvalue problem by taking $K_1 = 2$ in (2.6.40) and (2.6.41). After this we get:

$$B(\lambda)|\Phi_2(\lambda_1, \lambda_2)\rangle = [\omega_1(\lambda)]^N \prod_{j=1}^{2} i \frac{\alpha_2(\lambda_j, \lambda)}{\alpha_9(\lambda_j, \lambda)} |\Phi_2(\lambda_1, \lambda_2)\rangle - \sum_{j=1}^{2} [\omega_1(\lambda_j)]^N |\Psi_1^{(1)}\rangle$$

$$+ H_1 [\omega_1(\lambda_1)\omega_1(\lambda_2)]^N |\Psi_0^{(3)}\rangle$$

$$D(\lambda)|\Phi_2(\lambda_1, \lambda_2)\rangle = [\omega_3(\lambda)]^N \prod_{j=1}^{2} -i \frac{\alpha_8(\lambda, \lambda_j)}{\alpha_7(\lambda, \lambda_j)} |\Phi_2(\lambda_1, \lambda_2)\rangle + H_2 [\omega_2(\lambda_1)\omega_2(\lambda_2)]^N |\Psi_0^{(3)}\rangle$$

$$- \sum_{j=1}^{2} [\omega_2(\lambda_j)]^N \Lambda^{(1)}(\lambda_j, \lambda_1, \lambda_2) |\Psi_1^{(2)}\rangle \tag{2.6.42}$$

$$\sum_{i=1}^{2} A_i^i(\lambda)|\Phi_2(\lambda_1, \lambda_2)\rangle = [\omega_2(\lambda)]^N \prod_{j=1}^{2} -i \frac{\alpha_1(\lambda, \lambda_j)}{\alpha_9(\lambda, \lambda_j)} \Lambda^{(1)}(\lambda, \lambda_1, \lambda_2) |\Phi_2(\lambda_1, \lambda_2)\rangle$$

$$- \sum_{j=1}^{2} [\omega_2(\lambda_j)]^N \Lambda^{(1)}(\lambda, \lambda_1, \lambda_2) |\Psi_1^{(1)}\rangle - \sum_{j=1}^{2} [\omega_1(\lambda_j)]^N |\Psi_1^{(2)}\rangle$$

$$+ \left(H_3 [\omega_1(\lambda_1)\omega_2(\lambda_2)]^N + H_4 [\omega_1(\lambda_2)\omega_2(\lambda_1)]^N \right) |\Psi_0^{(3)}\rangle$$

where $\{|\Psi_\beta^{(\alpha)}\rangle\}$ are unwanted terms whose explicit form are complicated expressions which we will not write out. Without proof, we will just state that these unwanted terms can be get rid of by imposing the condition

$$\left(\frac{\omega_1(\lambda_i)}{\omega_2(\lambda_i)} \right)^N = \Lambda^{(1)}(\lambda_i, \lambda_1, \lambda_2) \quad \text{with} \quad i = 1, 2 \tag{2.6.43}$$

Finally, collecting all the wanted terms yields the eigenvalue

$$\Lambda(\lambda, \lambda_1, \lambda_2) = [\omega_1(\lambda)]^N \prod_{j=1}^{2} i\frac{\alpha_2(\lambda_j, \lambda)}{\alpha_9(\lambda_j, \lambda)} + [\omega_3(\lambda)]^N \prod_{j=1}^{2} -i\frac{\alpha_8(\lambda, \lambda_j)}{\alpha_7(\lambda, \lambda_j)}$$

$$- [\omega_2(\lambda)]^N \prod_{j=1}^{2} -i\frac{\alpha_1(\lambda, \lambda_j)}{\alpha_9(\lambda, \lambda_j)} \Lambda^{(1)}(\lambda, \lambda_1, \lambda_2) \tag{2.6.44}$$

This solves the two-excitation problem.

Based on the expressions for the eigenvalues and Bethe equations of the one and two excitation eigenvectors, we can easily guess the expression for the eigenvalue and Bethe equations for the K_1-excitation eigenvector to be:

$$\Lambda(\lambda, \{\lambda_j\}_{j=1}^{K_1}) = [\omega_1(\lambda)]^N \prod_{j=1}^{K_1} i\frac{\alpha_2(\lambda_j, \lambda)}{\alpha_9(\lambda_j, \lambda)} + [\omega_3(\lambda)]^N \prod_{j=1}^{K_1} -i\frac{\alpha_8(\lambda, \lambda_j)}{\alpha_7(\lambda, \lambda_j)}$$

$$- [\omega_2(\lambda)]^N \Lambda^{(1)}(\lambda, \{\lambda_j\}_{j=1}^{K_1}) \prod_{j=1}^{K_1} -i\frac{\alpha_1(\lambda, \lambda_j)}{\alpha_9(\lambda, \lambda_j)} \tag{2.6.45}$$

$$\prod_{l=1}^{K_1} \bar{b}(\mu_j, \lambda_l) = \prod_{l=1}^{K_2} \frac{\bar{b}(\mu_j, \mu_l)}{\bar{b}(\mu_l, \mu_j)} \quad \text{for all} \quad j = 1, \ldots, K_2$$

$$\left(\frac{\omega_1(\lambda_n)}{\omega_2(\lambda_n)}\right)^N = \Lambda^{(1)}(\lambda_n, \{\lambda_j\}_{j=1}^{K_1}) \quad \text{for all} \quad n = 1, \ldots, K_1 \tag{2.6.46}$$

$$\text{where} \quad \Lambda^{(1)}(\lambda, \{\lambda_i\}_{i=1}^{K_1}) = \prod_{j=1}^{K_2} \frac{1}{\bar{b}(\mu_j, \lambda)} + \frac{\prod_{j=1}^{K_1} \bar{b}(\lambda, \lambda_j)}{\prod_{l=1}^{K_2} \bar{b}(\lambda, \mu_l)} \quad \text{with} \quad 1 \le K_2 \le K_1 \tag{2.6.47}$$

The expression of the eigenvector however, is a bit harder to guess and we will only mention that the wave vector $\mathbf{\Phi}_{K_1}(\lambda_1, \ldots, \lambda_{K_1})$ is of the form

$$\mathbf{\Phi}_{K_1}(\lambda_1, \ldots, \lambda_{K_1}) = \mathbf{B}(\lambda_1) \otimes \mathbf{\Phi}_{K_1-1}(\lambda_2, \ldots, \lambda_{K_1}) + \sum_{j=2}^{K_1} i\frac{\alpha_{10}(\lambda_1, \lambda_j)}{\alpha_7(\lambda_1, \lambda_j)} \prod_{\substack{k=2 \\ k \ne j}}^{K_1} i\frac{\alpha_2(\lambda_k, \lambda_j)}{\alpha_9(\lambda_k, \lambda_j)}$$

$$\times \left[\boldsymbol{\xi} \otimes F(\lambda_1) \mathbf{\Phi}_{K_1-2}(\lambda_2, \ldots, \lambda_{j-1}, \lambda_{j+1}, \ldots \lambda_{K_1}) B(\lambda_j) \right] \prod_{l=2}^{j-1} \frac{\alpha_1(\lambda_k, \lambda_j)}{\alpha_2(\lambda_k, \lambda_j)} \hat{r}_{k,k+1}(\lambda_k, \lambda_j) \tag{2.6.48}$$

Note that this expression is in accordance with the one and two-excitation wave vector when we take $K_1 = 1$ and $K_1 = 2$ respectively.

Finally, if we introduce the variable $k(\lambda)$ defined by $e^{ik(\lambda)} = \frac{a(\lambda)}{b(\lambda)} e^{2h(\lambda)}$, then

$$\frac{a(\lambda)}{b(\lambda)} e^{2h(\lambda)} - \frac{b(\lambda)}{a(\lambda)} e^{-2h(\lambda)} = e^{ik(\lambda)} - e^{-ik(\lambda)} = 2i \sin[k(\lambda)] \tag{2.6.49}$$

Therefore, using (2.6.14), (2.6.19) and (2.6.20) we see that the Bethe equations (2.6.46) can be written as

$$\prod_{l=1}^{K_1} \frac{\frac{\tilde{\mu}_j}{2} - i\sin(k_l) - \frac{c}{2}}{\frac{\tilde{\mu}_j}{2} - i\sin(k_l) + \frac{c}{2}} = \prod_{l=1}^{K_2} \frac{\frac{\tilde{\mu}_j}{2} - \frac{\tilde{\mu}_l}{2} - c}{\frac{\tilde{\mu}_j}{2} - \frac{\tilde{\mu}_l}{2} + c} \quad \text{for all} \quad j = 1, \ldots, K_2 \tag{2.6.50}$$

$$e^{ik_n N} = \prod_{j=1}^{K_2} \frac{\frac{\tilde{\mu}_j}{2} - i\sin(k_n) + \frac{c}{2}}{\frac{\tilde{\mu}_j}{2} - i\sin(k_n) - \frac{c}{2}} \quad \text{for all} \quad n = 1, \ldots, K_1 \tag{2.6.51}$$

where $k_i \equiv k(\lambda_i)$ If we reparameterize $\tilde{\mu}$ to $2i\tilde{\mu}$, we will recover the Bethe equations (2.4.85).

2.6.3   Some Highest Weight Properties of the Hubbard Eigenvectors

In section 2.3.2 we have derived a symmetry property for the monodromy matrix and showed that the Bethe vectors were highest weight vectors for the given symmetry. Here, we will do the same.

First, let us remark that the Hubbard model has $\mathfrak{su}(2) \oplus \mathfrak{su}(2) \subseteq \mathfrak{su}(2|2)$ symmetry. This will become apparent in chapter 4 when we derive the (extended) $\mathfrak{su}(2|2)$-invariant S-matrix and show that it is equivalent with the Hubbard model R-matrix. The symmetry of the S-matrix is then also a symmetry of the Hubbard model R-matrix, and since we noted that the Hubbard model Lax matrix is just a permutation of its R-matrix, we have that the S-matrix symmetry is also the symmetry of the Hubbard model monodromy matrix.

To describe the symmetry properties, we first define the generators of each of the two $\mathfrak{su}(2)$ copies, which we denote with $\Sigma = \{\Sigma^z, \Sigma^+, \Sigma^-\}$ and $S = \{S^z, S^+, S^-\}$, by:

$$\Sigma^z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \Sigma^+ = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \Sigma^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$S^z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad S^+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad S^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.6.52}$$

Next, using the analogue of (2.3.49):

$$[T_a(\lambda), \Sigma_a^\alpha \otimes \mathbf{I}_h + \mathbf{I}_a \otimes \Sigma_h^\alpha] = 0 \qquad \text{and} \qquad [T_a(\lambda), S_a^\alpha \otimes \mathbf{I}_h + \mathbf{I}_a \otimes S_h^\alpha] = 0 \tag{2.6.53}$$

where a denotes the auxiliary space V_a and h denotes the pseudo-vacuum $|0\rangle$, we obtain:

$$
\begin{aligned}
[\Sigma^z, \mathbf{B}(\lambda)] &= -\mathbf{B}(\lambda), & [\Sigma^z, F(\lambda)] &= -2F(\lambda) \\
[\Sigma^+, \mathbf{B}(\lambda)] &= -\mathbf{C}^*(\lambda), & [\Sigma^+, F(\lambda)] &= B(\lambda) - D(\lambda) \\
[S^z, \mathbf{B}(\lambda)] &= \mathbf{B}(\lambda) \cdot \sigma^z, & [S^z, F(\lambda)] &= 0 \\
[S^\pm, \mathbf{B}(\lambda)] &= \mathbf{B}(\lambda) \cdot \frac{1}{2}\sigma^\pm, & [S^\pm, F(\lambda)] &= 0
\end{aligned} \tag{2.6.54}
$$

where $\sigma^{\pm} \equiv \sigma^x \pm i\sigma^y$ with σ^α, $\alpha = x, y, z$ being the Pauli matrices (2.0.5). Furthermore, from the matrix representation (2.6.52) we see that it acts on $|0\rangle_i$ as

$$\Sigma^+|0\rangle_i = S^+|0\rangle_i = S^z|0\rangle_i = 0 \qquad \text{and} \qquad \Sigma^z|0\rangle_i = |0\rangle_i \qquad (2.6.55)$$

With these identities we will now show that the Bethe vectors are highest weight vectors of the two $\mathfrak{su}(2)$ subalgebras and calculate its weight (i.e. the eigenvalues of Σ^z and S^z).

Let us denote the weight corresponding to the algebras Σ and S with s_1 and s_2 respectively. It is clear from (2.6.55) that the vacuum $|0\rangle$ is a highest weight vector with respect to Σ and S, and that it has weights $s_1 = N$ and $s_2 = 0$.

Next, we consider the one-excitation Bethe vector (2.6.24). Using (2.6.54) and (2.6.17) we then have

$$\Sigma^z |\Phi_1(\lambda_1)\rangle = (\mathbf{B}(\lambda_1)\Sigma^z - \mathbf{B}(\lambda_1)) \cdot \mathbf{F} |0\rangle = (N-1)|\Phi_1(\lambda_1)\rangle$$
$$\Sigma^+ |\Phi_1(\lambda_1)\rangle = (\mathbf{B}(\lambda_1)\Sigma^+ - \mathbf{C}^*(\lambda_1)) \cdot \mathbf{F} |0\rangle = 0 \qquad (2.6.56)$$

So the one-excitation Bethe vector $|\Phi_1(\lambda_1)\rangle$ is a highest weight vector of Σ with weight $s_1 = N - 1$.

Before we continue with the calculations for S, we first note that for the K_1-excitation case, $\mathbf{F}$ is a solution of the auxiliary eigenvalue problem (2.6.38). More precisely, $\mathbf{F}$ is a Bethe vector. Now because the auxiliary problem (2.6.38) is equivalent with the XXX model, and because every Bethe vector (which is found with the ABA) of the XXX model is a highest weight vector of $\mathfrak{su}(2)$ we have that

$$\sigma^+ \mathbf{F} \equiv \sum_{i=1}^{K_1} S_i^+ \mathbf{F} = 0 \qquad \text{and}$$

$$\sigma^z \mathbf{F} \equiv \sum_{i=1}^{K_1} S_i^z \mathbf{F} = (K_1 - 2K_2)\mathbf{F} \quad \text{where} \quad 0 \le K_2 \le K_1 \qquad (2.6.57)$$

analogous to (2.3.55) and (2.3.56). Therefore, returning to the one-excitation Bethe vector $|\Phi_1(\lambda_1)\rangle$ we get:

$$S^z|\Phi_1(\lambda_1)\rangle = (\mathbf{B}(\lambda_1)S^z + \mathbf{B}(\lambda_1) \cdot \sigma^z) \cdot \mathbf{F} |0\rangle = (1 - 2K_2)|\Phi_1(\lambda_1)\rangle \quad \text{where} \quad 0 \le K_2 \le 1$$

$$S^+|\Phi_1(\lambda_1)\rangle = \left(\mathbf{B}(\lambda_1)S^+ + \mathbf{B}(\lambda_1) \cdot \frac{1}{2}\sigma^+\right) \cdot \mathbf{F} |0\rangle = 0 \qquad (2.6.58)$$

Let us continue with two-excitation case. For the two-excitation Bethe vector (2.6.34) we have

$$\Sigma^z |\Phi_2(\lambda_1, \lambda_2)\rangle = [\mathbf{B}(\lambda_1) \otimes \mathbf{B}(\lambda_2)\Sigma^z - 2\mathbf{B}(\lambda_1) \otimes \mathbf{B}(\lambda_2)] \cdot \mathbf{F} |0\rangle$$
$$+ F(\lambda_1)[\omega_1(\lambda_2)]^L \hat{g}_0^{(2)}(\lambda_1, \lambda_2)[\Sigma^z - 2]\boldsymbol{\xi} \cdot \mathbf{F} |0\rangle \qquad (2.6.59)$$
$$= (N-2)|\Phi_2(\lambda_1, \lambda_2)\rangle$$

To show that the two-excitation Bethe vector is a highest weight vector of Σ we write:

$$\Sigma^+ |\Phi_2(\lambda_1, \lambda_2)\rangle = \mathbf{B}(\lambda_1) \otimes \Sigma^+ \mathbf{B}(\lambda_2) \cdot \mathbf{F} |0\rangle - \mathbf{C}^*(\lambda_1) \otimes \mathbf{B}(\lambda_2) \cdot \mathbf{F} |0\rangle$$
$$+ i\frac{\alpha_{10}(\lambda_1, \lambda_2)}{\alpha_7(\lambda_1, \lambda_2)}[B(\lambda_1) - D(\lambda_1)][\omega_1(\lambda_2)]^L \boldsymbol{\xi} \cdot \mathbf{F} |0\rangle \tag{2.6.60}$$

The first term will vanish with the same arguments as for the one-excitation Bethe vector. The third term can be simplified using (2.6.15), and without proof, we will mention that the second term can be can also be simplified [22]. The resulting expression is

$$\Sigma^+ |\Phi_2(\lambda_1, \lambda_2)\rangle = i\frac{\alpha_{10}(\lambda_1, \lambda_2)}{\alpha_7(\lambda_1, \lambda_2)} \left([\omega_1(\lambda_1)\omega_1(\lambda_2)]^L - [\omega_2(\lambda_1)\omega_2(\lambda_2)]^L \right) \boldsymbol{\xi} \cdot \mathbf{F} |0\rangle \tag{2.6.61}$$

Now from the Bethe equations (2.6.46) and (2.6.47) we see that

$$\prod_{j=1}^{K_1} \Lambda^{(1)}(\lambda_j, \{\lambda_i\}_{i=1}^{K_1}) = 1 \tag{2.6.62}$$

so that

$$[\omega_2(\lambda_1)\omega_2(\lambda_2)]^L = [\omega_2(\lambda_1)\omega_2(\lambda_2)]^L \Lambda^{(1)}(\lambda_1, \{\lambda_1, \lambda_2\}) \Lambda^{(1)}(\lambda_2, \{\lambda_1, \lambda_2\}) = [\omega_1(\lambda_1)\omega_1(\lambda_2)]^L \tag{2.6.63}$$

And this implies that

$$\Sigma^+ |\Phi_2(\lambda_1, \lambda_2)\rangle = 0 \tag{2.6.64}$$

For the second $\mathfrak{su}(2)$ we have

$$S^z |\Phi_2(\lambda_1, \lambda_2)\rangle = [\mathbf{B}(\lambda_1)\sigma^z \otimes \mathbf{B}(\lambda_2) + \mathbf{B}(\lambda_1) \otimes \sigma^z \mathbf{B}(\lambda_2)] \cdot \mathbf{F} |0\rangle$$
$$= [\mathbf{B}(\lambda_1) \otimes \mathbf{B}(\lambda_2)] \cdot [\sigma^z \otimes \mathbb{I} + \mathbb{I} \otimes \sigma^z] \cdot \mathbf{F} |0\rangle$$
$$= (2 - 2K_2) |\Phi_2(\lambda_1, \lambda_2)\rangle \qquad \text{where} \qquad 0 \le K_2 \le 2$$
$$S^+ |\Phi_2(\lambda_1, \lambda_2)\rangle = [\mathbf{B}(\lambda_1)\frac{1}{2}\sigma^+ \otimes S^+\mathbf{B}(\lambda_2)] \cdot \mathbf{F} |0\rangle + i\frac{\alpha_{10}(\lambda_1, \lambda_2)}{\alpha_7(\lambda_1, \lambda_2)}[\omega_1(\lambda_2)]^L S^+ \boldsymbol{\xi} \cdot \mathbf{F} |0\rangle$$
$$= [\mathbf{B}(\lambda_1)\frac{1}{2}\sigma^+ \otimes \mathbf{B}(\lambda_2) + \mathbf{B}(\lambda_1) \otimes \frac{1}{2}\sigma^+\mathbf{B}(\lambda_2)] \cdot \mathbf{F} |0\rangle$$
$$= \frac{1}{2}[\mathbf{B}(\lambda_1) \otimes \mathbf{B}(\lambda_2)] \cdot [\sigma^+ \otimes \mathbb{I} + \mathbb{I} \otimes \sigma^+] \cdot \mathbf{F} |0\rangle = 0$$

where we used that $\mathbf{F}$ is a highest weight vector of $\mathfrak{su}(2)$.

These calculation can easily be generalized to a K_1-excitation Bethe vector. In that case the weight will be $[N - K_1, K_1 - 2K_2]$.

2.7 The Bethe Ansatz, S-Matrices and Integrable Systems

In solving the XXX model and the Hubbard model, we have applied the coordinate Bethe ansatz right from the beginning without thinking about the validity of the ansatz. Fortunately the CBA does solve both models, but one may wonder why the whole procedure works. The answer to that question is "integrability". Both models are integrable (we have shown this for the XXX model using the algebraic Bethe ansatz) and that is the reason the CBA works.

Let us elaborate on this a bit further. Suppose that we have a system with various kinds of particles which admits scattering. Furthermore, let $|0\rangle$ be the vacuum quantum state, and let $A_a^\dagger(p)$ and $A_a(p)$ be the creation and annihilation operator respectively of a particle of type a with momentum p. Then, a scattering process will be described by using the "in" state and "out" state basis which are defined as:

$$
\begin{aligned}
|p_1, \ldots, p_n\rangle_{i_1,\ldots,i_n}^{(in)} &\equiv A_{i_1}^\dagger(p_1) \ldots A_{i_n}^\dagger(p_n)|0\rangle, \quad p_1 > p_2 > \ldots > p_n \\
|p_1, \ldots, p_n\rangle_{i_1,\ldots,i_n}^{(out)} &\equiv A_{i_n}^\dagger(p_n) \ldots A_{i_1}^\dagger(p_1)|0\rangle, \quad p_1 > p_2 > \ldots > p_n
\end{aligned}
\tag{2.7.1}
$$

In the scattering process, the in states will go to the out states. In particular, the in and out states are related by a so called S-matrix $\mathbf{S}$ as follows:

$$
|p_1, \ldots, p_m\rangle_{i_1,\ldots,i_m}^{(in)} = \mathbf{S} |p'_1, \ldots, p'_n\rangle_{j_1,\ldots,j_n}^{(out)}
\tag{2.7.2}
$$

By using an explicit basis we can write this as

$$
|p_1, \ldots, p_m\rangle_{i_1,\ldots,i_m}^{(in)} = \sum_{n=2}^{\infty} \sum_{p'_1 < \ldots < p'_n} \sum_{\{j_1,\ldots,j_n\}} S_{i_1,\ldots,i_m}^{j_1,\ldots,j_n}(p_1, \ldots, p_m; p'_1, \ldots, p'_n)|p'_1, \ldots, p'_n\rangle_{j_1,\ldots,j_n}^{(out)}
\tag{2.7.3}
$$

where $S(p_1, \ldots, p_m; p'_1, \ldots, p'_n)$ is the matrix form of the action of $\mathbf{S}$.

Now, if the system is also integrable, we will have the following properties

- There is no particle production; the number of particles before and after the scattering process is the same. This means that $n = m$ in (2.7.3)

- The sets of momenta of the particles before and after the scattering process is the same. This means that $p_i = p'_i$ for all i in (2.7.3)

- The S-matrix of a n-particle scattering process factorizes into a product of 2-particle S-matrices.

- The 2-particle S-matrix satisfies the Yang-Baxter equation (A.2.3)

- The wave functions of the eigenvectors are asymptotically, i.e. when $n_1 \ll n_2 \ll \ldots \ll n_N$, of the coordinate Bethe ansatz form.

When the system is integrable, we are generally referring to the two-particle S-matrix when we talk about S-matrices (note that this is in accordance with the way we defined our S-matrices in the previous sections; they are all two-particle S-matrices). Next, notice the appearance of the word "asymptotically" in the last property. This means that in general, the CBA will not hold everywhere and therefore, the CBA is also called the asymptotic Bethe ansatz. For the XXX model and the Hubbard model the CBA did hold everywhere, but that is because the interaction in those two models is a nearest neighbor interaction.

We see how important the notion of integrability is. In particular, we see that the S-matrix plays a key role in integrable systems. Therefore, we can ask ourselves the following interesting question: in what degree is the two-particle S-matrix uniquely defined and what is the corresponding integrable system. This is a question which has been investigated by Zamolodchikov [28] and which resulted into the introduction of the Zamolodchikov-Faddeev algebra. It turns out that the form of the S-matrix is heavily restricted by its symmetry properties and the Yang-Baxter equation, and, depending on the symmetry algebra, these restrictions can even define the S-matrix uniquely (up to a constant).

Let us outline this construction in a bit more detail. Suppose that we have a multiplet of particles and suppose that we have a symmetry algebra $\mathfrak{J}$ which acts linearly on the particle states. Then, for generators $\mathbf{J} \in \mathfrak{J}$ which preserve the number of particles we have the action:

$$\mathbf{J}\,|0\rangle = 0\,, \quad \mathbf{J}\,A^{\dagger}_{i_1}(p_1)\ldots A^{\dagger}_{i_k}(p_k)|0\rangle = J^{j_1,\ldots,j_k}_{i_1,\ldots,i_k}(p_1,\ldots,p_k)A^{\dagger}_{j_1}(p_1)\ldots A^{\dagger}_{j_k}(p_k)|0\rangle \quad (2.7.4)$$

Now suppose that the scattering process of the particles satisfies the first three properties of integrability. Then, for two particles, equations (2.7.3) and (2.7.4) become

$$A^{\dagger}_i(p_1)A^{\dagger}_j(p_2)|0\rangle = S^{k,l}_{i,j}A^{\dagger}_l(p_2)A^{\dagger}_k(p_1)|0\rangle$$
$$\mathbf{J}\,A^{\dagger}_i(p_1)A^{\dagger}_j(p_2)|0\rangle = J^{k,l}_{i,j}A^{\dagger}_l(p_2)A^{\dagger}_k(p_1)|0\rangle$$
$$(2.7.5)$$

If we now let $\mathbf{J}$ act on both sides of the first equation, we will get the condition

$$J^{k,l}_{i,j}(p_1,p_2)S^{m,n}_{k,l}(p_1,p_2) = S^{k,l}_{i,j}(p_1,p_2)J^{n,m}_{l,k}(p_2,p_1) \quad (2.7.6)$$

which will be called the "invariance condition" for the (two-particle) S-matrix. This together with the Yang-Baxter equation will put restrictions on the form of the two-particle S-matrix.

When we have obtained a S-matrix we can construct an integrable system by simply imposing the integrability properties outlined above, and its corresponding two-body potential can subsequently be found using inverse scattering methods. Finally, we can calculate the Bethe equations of our system by diagonalizing the transfer matrix constructed from the two-particle S-matrix.

Chapter 3

The $\mathfrak{psu}(2,2|4)$ Lie Superalgebra

3.1 General Theory of Lie superalgebras

A Lie superalgebra $\mathfrak{g}$ is a $\mathbb{Z}_2$-graded algebra over a field k of characteristic 0, that is also a vector space which can be written as the direct sum of two vector spaces $\mathfrak{g}_{\bar{0}}$ and $\mathfrak{g}_{\bar{1}}$. Furthermore, it is equipped with a product (or supercommutator) $[\![\cdot\,,\cdot]\!] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which satisfies

- $\mathbb{Z}_2$-gradation: $[\![\mathfrak{g}_i, \mathfrak{g}_j]\!] \subset \mathfrak{g}_{i+j}$

- graded skew-symmetry: $[\![X_i, X_j]\!] = -(-1)^{ij}[\![X_j, X_i]\!]$

- graded Jacobi identity:

$$(-1)^{ik}[\![X_i, [\![X_j, X_k]\!]]\!] + (-1)^{ij}[\![X_j, [\![X_k, X_i]\!]]\!] + (-1)^{jk}[\![X_k, [\![X_i, X_j]\!]]\!] = 0$$

for all $X_i \in \mathfrak{g}_i$, $X_j \in \mathfrak{g}_j$, $i, j \in \mathbb{Z}_2$. In this and the upcoming sections, we will use the following supercommutator

$$[\![X_i, X_j]\!] = X_i X_j - (-1)^{ij} X_j X_i \quad \text{for all} \quad X_i \in \mathfrak{g}_i, \ X_j \in \mathfrak{g}_j, \ i, j \in \mathbb{Z}_2 \tag{3.1.1}$$

In particular, we can write

$$[\![X_i, X_j]\!] = \begin{cases} [X_i, X_j] = X_i X_j - X_j X_i & \text{if } ij = 0 \\ \{X_i, X_j\} = X_i X_j + X_j X_i & \text{if } ij = 1 \end{cases} \tag{3.1.2}$$

and we call $[\cdot\,,\cdot]$ the commutator, and $\{\cdot\,,\cdot\}$ the anti-commutator. Notice that from these definitions it follows that $\mathfrak{g}_{\bar{0}}$ is a Lie algebra.

An important class of Lie superalgebras are the so called basic Lie superalgebras. A basic Lie superalgebra $\mathfrak{g} = \mathfrak{g}_{\bar{0}} \oplus \mathfrak{g}_{\bar{1}}$ is a Lie superalgebra for which:

- $\mathfrak{g}$ is simple: $\dim(\mathfrak{g}) \geq 3$, and the only subalgebras $\mathfrak{k}$ of $\mathfrak{g}$ with the property $[\![\mathfrak{g}, \mathfrak{k}]\!] \subset \mathfrak{k}$ are $\mathfrak{g}$ and $\{0\}$.

- $\mathfrak{g}_{\bar{0}}$ is a reductive Lie algebra: $\mathfrak{g}_{\bar{0}}$ is isomorphic to a direct sum of indecomposable Lie algebras.

- there exists a bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ which is:

  - Non-degenerate: If $B(X, Y) = 0$ for all $X \in \mathfrak{g}$, then $Y = 0$
  - Consistent: $B(X, Y) = 0$, for all $X \in \mathfrak{g}_{\bar{0}}$, $Y \in \mathfrak{g}_{\bar{1}}$.
  - Supersymmetric: $B(X_i, X_j) = (-1)^{ij} B(X_j, X_i)$ for all $X_i \in \mathfrak{g}_i$.
  - Invariant: $B(\llbracket X, Y \rrbracket, Z) = B(X, \llbracket Y, Z \rrbracket)$.

In the rest of the section, we shall assume that $k = \mathbb{C}$ and that $\mathfrak{g}$ is a basic Lie superalgebra. The results and definitions can also be applied to Lie superalgebras $\mathfrak{g} = \mathfrak{g}_{\bar{0}} \oplus \mathfrak{g}_{\bar{1}}$ with $k = \mathbb{R}$ if we consider $\mathfrak{g}$ as the real part of its complexification $\mathfrak{g}_{\mathbb{C}} \equiv \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$; $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}_{\mathbb{C}_{\bar{0}}} \oplus \mathfrak{g}_{\mathbb{C}_{\bar{1}}}$. Therefore, we will use the convention that when we consider a Lie superalgebra $\mathfrak{g}$ over $\mathbb{R}$, we will actually refer to the complexified Lie superalgebra $\mathfrak{g}_{\mathbb{C}}$.

A Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a maximum commuting subalgebra of $\mathfrak{g}$. For the basic Lie superalgebras, this will coincide with the Cartan subalgebra of the Lie algebra $\mathfrak{g}_{\bar{0}}$. Now let $\alpha \in \mathfrak{h}^*$ and define

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} | \llbracket H, X \rrbracket = \alpha(H)X, \ H \in \mathfrak{h}\} \tag{3.1.3}$$

If $\mathfrak{g}_\alpha \neq 0$ and $\alpha \neq 0$, then we call α a root of $\mathfrak{g}$ (with respect to $\mathfrak{h}$) and $\mathfrak{g}_\alpha$ is the corresponding root space. More precisely, a root α is called even (resp. odd) if $\mathfrak{g}_\alpha \cap \mathfrak{g}_{\bar{0}} \neq 0$ (resp. $\mathfrak{g}_\alpha \cap \mathfrak{g}_{\bar{1}} \neq 0$). Notice that a root can be both even and odd, but for the basic Lie superalgebras, this will not be the case. We will denote the set of all roots with Δ, the set of even roots with $\Delta_{\bar{0}}$ and the set of odd roots with $\Delta_{\bar{1}}$; so

$$\Delta = \{\alpha \in \mathfrak{h}^* | \mathfrak{g}_\alpha \neq 0\}, \quad \Delta_{\bar{0}} = \{\alpha \in \Delta | \mathfrak{g}_\alpha \cap \mathfrak{g}_{\bar{0}} \neq 0\}, \quad \Delta_{\bar{1}} = \{\alpha \in \Delta | \mathfrak{g}_\alpha \cap \mathfrak{g}_{\bar{1}} \neq 0\} \tag{3.1.4}$$

The set Δ spans $\mathfrak{h}^*$ and is called a root system. Roots have the property that if α is a root, then $-\alpha$ is also a root. Additionally, for basic Lie superalgebras we also have that $\mathfrak{g}_\alpha$ is 1-dimensional.

Following the definition of basic Lie superalgebras, we can define a non-degenerate consistent supersymmetric invariant bilinear form $\langle \cdot , \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ on $\mathfrak{g}$. The restriction of this form to $\mathfrak{h}$ remains non-degenerate consistent supersymmetric invariant and for a given $\alpha \in \mathfrak{h}^*$ (not necessarily a root), there exists a unique element $H^\alpha \in \mathfrak{h}$ such that $\alpha(H) = \langle H^\alpha, H \rangle$ for all $H \in \mathfrak{h}$. Using this bijective correspondence $\alpha \longleftrightarrow H^\alpha$ between $\mathfrak{h}^*$ and $\mathfrak{h}$ we can extend the bilinear form to $\mathfrak{h}^*$ by defining $\langle \alpha, \beta \rangle \equiv \langle H^\alpha, H^\beta \rangle$, for all $\alpha, \beta \in \mathfrak{h}^*$.

For a given root system we can choose a subset $\Pi = \{\alpha_1, \ldots, \alpha_r\}$, $r = \dim(\mathfrak{h})$ of Δ such that for each $\alpha \in \Delta$ we have $\alpha = n_1 \alpha_1 + \ldots + n_r \alpha_r$ where either $n_j \in \mathbb{N}_{\geq 0}$ for all j, or $n_j \in \mathbb{N}_{<0}$ for all j. The elements of Π are then called simple roots and Π is a simple root system. The α's for which $n_j \geq 0$ are called the positive roots and the α's for which $n_j < 0$

are called the negative roots. The set of positive and negative roots will be denoted by Δ_+ and Δ_- respectively.

For Lie algebras we know that all the simple root systems (for a given root system) are equivalent with each other; they can be transformed to each other with a Weyl group transformation. For Lie superalgebras however, there exist no fully working analogue of the Weyl group and so in general there are many inequivalent simple root systems. For basic Lie superalgebras there exist simple root systems for which the number of simple odd roots is the smallest one. These root systems are called distinguished simple root systems.

For a given simple root system, we can construct Cartan numbers a_{ij} for $\alpha_i, \alpha_j \in \Pi$. They are defined as $a_{ij} = \langle \alpha_i, \alpha_j \rangle$. The Cartan numbers form a square matrix called the Cartan matrix. Notice that this definition differs from the one for Lie algebras since it is possible to have $\langle \alpha_i, \alpha_i \rangle = 0$.

An elegant way to describe a basic Lie superalgebra is to use a Chevalley basis. Let $\mathfrak{g}$ be a basic Lie superalgebra with Cartan subalgebra $\mathfrak{h}$ and let us choose a non-degenerate consistent supersymmetric invariant bilinear form $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ (this bilinear form will be unique up to a constant factor). We then find the root system Δ and choose a simple root system Π. Subsequently, we calculate the Cartan matrix $a_{ij} = \langle \alpha_i, \alpha_j \rangle$ and for each $\alpha_i \in \Pi$ we can find elements $E_i^{\pm} \in \mathfrak{g}_{\pm\alpha_i}$ and $H_i \in \mathfrak{h}$ such that

$$[\![H_i, H_j]\!] = 0, \qquad [\![H_i, E_j^{\pm}]\!] = \pm a_{ij} E_i^{\pm}, \qquad [\![E_i^+, E_j^-]\!] = \delta_{ij} H_i \qquad (3.1.5)$$

The set of elements $\{H_i, E_i^{\pm}\}_{i=1}^r$ is called a Chevalley basis[1] and it generates the whole Lie superalgebra $\mathfrak{g}$. It is interesting to note that using these generators, we can describe the root spaces as follows. If $\alpha \in \Delta$, then $\mathfrak{g}_\alpha$ is the linear span of elements of the form $[\![E_{i_1}^+, [\![E_{i_2}^+, [\![\ldots, E_{i_s}^+]\!] \ldots]\!]]\!]$ such that $\sum_s \alpha_{i_s} = \alpha$, or of elements of the form $[\![E_{i_1}^-, [\![E_{i_2}^-, [\![\ldots, E_{i_s}^-]\!] \ldots]\!]]\!]$ such that $\sum_s \alpha_{i_s} = -\alpha$.

Let V be a $\mathbb{Z}_2$-graded vector space over k; that is, V is the direct sum of two vector spaces $V_{\bar{0}}$ and $V_{\bar{1}}$. A (finite-dimensional) representation $\pi = (\pi, V)$ of $\mathfrak{g}$ in V consists of a (finite-dimensional) $\mathbb{Z}_2$-graded vector space V and a homomorphism $\pi : \mathfrak{g} \to End(V)$ such that

$$\pi(X_i)V_j \in V_{i+j} \quad \text{for all} \quad i, j \in \mathbb{Z}_2$$
$$\pi([\![X, Y]\!]) = [\![\pi(X), \pi(Y)]\!] \quad \text{and} \quad \pi(aX) = a\pi(X) \quad \text{for all } X, Y \in \mathfrak{g}, a \in k \qquad (3.1.6)$$

If (π, V) is a representation of $\mathfrak{g}$, then we say that $\mathfrak{g}$ acts on V and that V is a $\mathfrak{g}$-module.

Now suppose that (π, V) and (π', V') are representations of $\mathfrak{g}$. Then the tensor product of these two representations, denoted by $(\pi \bar\otimes \pi', V \bar\otimes V')$, is also a representation of $\mathfrak{g}$ if we define the action of $\mathfrak{g}$ as:

$$((\pi \bar\otimes \pi')X)(v \bar\otimes w) = \pi(X)v \bar\otimes w + (-1)^{ij} v \bar\otimes \pi'(X)w \quad \text{for all } X \in \mathfrak{g}_i, v \in V_j \qquad (3.1.7)$$

---

[1]Additionally, the elements of the Chevalley basis will also be related via the so called Serre relations. But we will not consider them here.

A finite dimensional representation (π, V) is called irreducible, if the only subspaces W of V with the property that $\pi(g)w \in W$ for all $w \in W$, $g \in \mathfrak{g}$, are V and $\{0\}$.

Analogously to roots and root spaces, we can introduce the notion of weight and weight spaces. First, let $\mathfrak{g}$ be a Lie superalgebra with Cartan subalgebra $\mathfrak{h}$ and (π, V) a finite dimensional representation of $\mathfrak{g}$. Then $\Lambda \in \mathfrak{h}^*$ is a weight of $\mathfrak{h}$ in V if

$$V_\Lambda \equiv \{v \in V | \pi(H)v = \Lambda(H)v, \, H \in \mathfrak{h}\} \neq \{0\}$$

V_Λ will be called a weight space.

Now let $\mathfrak{g}$ be a basic Lie superalgebra with Cartan subalgebra $\mathfrak{h}$ and simple root system Π. Then an irreducible representation (π, V) is called a highest weight representation with highest weight $\Lambda \in \mathfrak{h}^*$ if:

- there exists a non-zero vector $v_\Lambda \in V$ such that

$$
\begin{aligned}
\pi(X)v_\Lambda &= 0 \quad \text{for all} \quad X \in \{\cup_\alpha \mathfrak{g}_\alpha | \alpha \text{ is positive root}\} \\
\pi(H)v_\Lambda &= \Lambda(H)v_\Lambda \quad \text{for all} \quad H \in \mathfrak{h}
\end{aligned}
\tag{3.1.8}
$$

- the smallest invariant subspace of V containing v_Λ is V itself.

Then the vector v_Λ will be called a highest weight vector and V will be called a highest weight module and is also denoted by V_Λ. Furthermore, for each simple root $\alpha_i \in \Pi$ we can define the so called Dynkin label a_i by $a_i = \Lambda(H_i)$, where $H_i \in \mathfrak{h}$ is an element of the Chevalley basis.

It is also possible to construct a representation for $\mathfrak{g}$ by first specifying a highest weight vector v_Λ and highest weight Λ. In this construction, the module V_Λ is called a Verma module and it is defined as the linear span of vectors of the form

$$\pi(Y_{i_1})\pi(Y_{i_2})\ldots\pi(Y_{i_s})v_\lambda \qquad \text{for all } Y_i \in \mathfrak{g}_\alpha, \, \alpha \in \Delta_- \tag{3.1.9}$$

This module is infinite dimensional and so we end up with an infinite dimensional representation which may or may not be irreducible. Fortunately, we can obtain an irreducible representation from V_Λ by the following construction. We define U_Λ to be the subspace of V_Λ consisting of all vectors v such that

- the v_Λ-component of v is zero.

- the v_Λ-component of $\pi(X_{i_1})\pi(X_{i_2})\ldots\pi(X_{i_s})v$ is zero for any collection of $X_{i_1}\ldots X_{i_s}$ with $X_i \in \mathfrak{g}_\alpha, \, \alpha \in \Delta_+$.

Then the quotients space V_Λ/U_Λ is an irreducible representation with highest weight Λ and highest weight vector v_Λ. Depending on Λ, this representation may or may not be infinite-dimensional.

In the remainder of this thesis we will drop the notation with π. That is, we will simply write Xv instead of $\pi(X)v$ when there is no chance for confusion.

3.2 The $\mathfrak{psu}(2, 2|4)$ Lie superalgebra

The Lie superalgebra $\mathfrak{su}(2, 2|4)$ is a graded algebra over $\mathbb{R}$ and it is spanned by matrices of 4×4 blocks of the form

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{3.2.1}$$

The even, $\mathfrak{su}(2, 2|4)_{\bar{0}}$ and odd, $\mathfrak{su}(2, 2|4)_{\bar{1}}$ part of $\mathfrak{su}(2, 2|4)$ is generated by matrices of the form

$$M_{\bar{0}} = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \quad \text{and} \quad M_{\bar{1}} = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \quad \text{respectively.} \tag{3.2.2}$$

The matrices satisfy the conditions:

$$\text{str}(M) \equiv tr(A) - tr(D) = 0 \quad \text{and} \quad HM + M^{\dagger}H = 0$$
$$\text{with} \quad H = \begin{pmatrix} \Sigma & 0 \\ 0 & \mathbb{I}_4 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} \tag{3.2.3}$$

This last condition means that

$$\begin{pmatrix} \Sigma A + A^{\dagger}\Sigma & \Sigma B + C^{\dagger} \\ C + B^{\dagger}\Sigma & D + D^{\dagger} \end{pmatrix} = 0 \tag{3.2.4}$$

so A and D span the subalgebras $\mathfrak{u}(2, 2)$ and $\mathfrak{u}(4)$ respectively. Since the $\mathfrak{u}(1)$ generator $i\mathbb{I}$ obeys the above conditions, it is also contained in $\mathfrak{su}(2, 2|4)$. Therefore, $\mathfrak{su}(2, 2|4)_{\bar{0}} = \mathfrak{su}(2, 2) \oplus \mathfrak{su}(4) \oplus \mathfrak{u}(1)$. The Lie superalgebra $\mathfrak{psu}(2, 2|4)$ is now defined as the quotient $\mathfrak{su}(2, 2|4)/\mathfrak{u}(1)$ [2]. Because $\mathfrak{psu}(2, 2|4)$, as opposed to $\mathfrak{su}(2, 2|4)$, cannot be realized in terms of 8×8 matrices, we will continue to work with $\mathfrak{su}(2, 2|4)$ (to be more precise, we will work with the complexified Lie superalgebra $\mathfrak{su}(2, 2|4)_{\mathbb{C}} = \mathfrak{sl}(4|4)$).

3.2.1 Chevalley Basis I

To describe $\mathfrak{su}(2, 2|4)$, we will give a Chevalley basis $\{H_i, E_i^+, E_i^- | i = 1, \dots, 7\}$ for it. Here, H_i are the Cartan generators and $E_i^{\pm}$ are the simple root space generators such that:

$$[\![H_i, H_j]\!] = 0, \qquad [\![E_i^+, E_j^-]\!] = \delta_{ij}H_j, \qquad [\![H_i, E_j^{\pm}]\!] = \pm a_{ij}E_j^{\pm} \tag{3.2.5}$$

with a_{ij} the elements of the Cartan matrix.

First, let $\langle \cdot, \cdot \rangle$ be a non-degenerate consistent supersymmetric invariant bilinear form on $\mathfrak{su}(2, 2|4)$ defined by:

$$\langle X, Y \rangle = -\text{str}(XY) \quad \text{for all} \quad X, Y \in \mathfrak{sl}(4|4) \tag{3.2.6}$$

---

[2]In fact, in general we have $\mathfrak{psu}(n, m|n + m) = \mathfrak{su}(n, m|n + m)/\mathfrak{u}(1)$.

Next, let $\mathfrak{h}$ be the Cartan subalgebra of $\mathfrak{su}(2,2|4)$ generated by

$$
H_2 = \begin{pmatrix} -1 & & & & & & & \\ & 1 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix}, \quad
H_3 = \begin{pmatrix} 0 & & & & & & & \\ & -1 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & -1 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix}, \quad
H_4 = \begin{pmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & 0 & & & \\ & & & & & 1 & & \\ & & & & & & -1 & \\ & & & & & & & 0 \end{pmatrix},
$$

$$
H_5 = \begin{pmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 1 & \\ & & & & & & & 0 \end{pmatrix}, \quad
H_6 = \begin{pmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & -1 & & & & & \\ & & & 1 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix}
$$

$$
H_1 = \begin{pmatrix} 1 & & & & & & & \\ & 0 & & & & & & \\ & & 0 & & & & & \\ & & & 0 & & & & \\ & & & & 1 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix}, \quad
H_7 = \begin{pmatrix} 0 & & & & & & & \\ & 0 & & & & & & \\ & & 0 & & & & & \\ & & & -1 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & -1 \end{pmatrix}
$$

$$(3.2.7)$$

Notice that $\{H_2, H_3, H_4\}$ and $\{H_4, H_5, H_6\}$ form two Cartan subalgebras of $\mathfrak{su}(2|2)$. We will come back to this in more detail in the next chapter. Now, let Π be the simple root system with simple root $\alpha_i \in \Pi$ defined by $\alpha_i(H_j) = \langle H_i, H_j \rangle$. Then we see that Π consists of even simple roots $\{\alpha_1, \alpha_3, \alpha_5, \alpha_7\}$ and odd simple roots $\{\alpha_2, \alpha_4, \alpha_6\}$ defined by:

$$
\begin{aligned}
\alpha_1 &= \delta_1 - \epsilon_1\,, & \alpha_2 &= \epsilon_1 - \epsilon_2\,, & \alpha_3 &= \epsilon_2 - \delta_2\,, & \alpha_4 &= \delta_2 - \delta_3 \\
\alpha_5 &= \delta_3 - \epsilon_3\,, & \alpha_6 &= \epsilon_3 - \epsilon_4\,, & \alpha_7 &= \epsilon_4 - \delta_4
\end{aligned}
$$

$$(3.2.8)$$

Here, $\epsilon_i, \delta_i \in \mathfrak{h}^*$ are defined as $\delta_i = \epsilon_{4+i}$, $\epsilon_i(H) = H_{ii}$ for all matrices $H \in \mathfrak{h}$, where H_{ii} is the (i,i)-th matrix element of H. We can now construct the symmetric Cartan matrix a by $a_{ij} = \langle \alpha_i, \alpha_j \rangle = -\mathrm{str}(H_i H_j)$:

$$
a = \begin{pmatrix}
0 & +1 & 0 & 0 & 0 & 0 & 0 \\
+1 & -2 & +1 & 0 & 0 & 0 & 0 \\
0 & +1 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & -1 & 2 & -1 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & +1 & 0 \\
0 & 0 & 0 & 0 & +1 & -2 & +1 \\
0 & 0 & 0 & 0 & 0 & +1 & 0
\end{pmatrix}
$$

$$(3.2.9)$$

This Cartan matrix can be represented graphically by a Dynkin diagram using the following rules:

- One draws for each simple even root a white dot, for each simple odd root of non-zero length a black dot and for each simple odd root of zero length a crossed dot.

- the i-th and j-th dots will be joined by n_{ij} lines where

$$n_{ij} = \frac{2|a_{ij}|}{\min(|a_{ii}|, |a_{jj}|)} \quad \text{if } a_{ii} \cdot a_{jj} \neq 0$$

$$n_{ij} = \frac{2|a_{ij}|}{\min_{a_{kk} \neq 0} |a_{kk}|} \quad \text{if } a_{ii} \neq 0 \text{ and } a_{jj} = 0$$

$$n_{ij} = |a_{ij}| \quad \text{if } a_{ii} = a_{jj} = 0$$

- We will add an arrow on the lines connecting the i-th and j-th dots when $n_{ij} > 1$. The arrow points from i to j

  - if $a_{ii} \cdot a_{jj} \neq 0$ and $|a_{ii}| > |a_{jj}|$, or

  - if $a_{ii} = 0$, $a_{jj} \neq 0$ and $|a_{jj}| < 2$.

The arrow points from j to i

  - if $a_{ii} = 0$, $a_{jj} \neq 0$ and $|a_{jj}| > 2$.

Using these rules we obtain the following Dynkin diagram for our Cartan matrix:

$$(3.2.10)$$

We complete our Chevalley basis by giving the corresponding simple root generators. The $\{E_i^-\}$ are given by:

$$E_2^- = \begin{pmatrix} \begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix} & & \\ & 0 & \\ & & 0 \\ & & & 0_{4\times 4} \end{pmatrix}, \quad E_3^- = \begin{pmatrix} 0_{4\times 4} & & 0_{4\times 4} \\ \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} & & 0_{4\times 4} \end{pmatrix}, \quad E_4^- = \begin{pmatrix} 0_{4\times 4} & & \\ & 0 & \\ & & \begin{smallmatrix} 0 & 0 \\ -1 & 0 \end{smallmatrix} \\ & & & 0 \end{pmatrix}$$

$$E_5^- = \begin{pmatrix} 0_{4\times 4} & \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} \\ 0_{4\times 4} & 0_{4\times 4} \end{pmatrix}, \quad E_6^- = \begin{pmatrix} 0 & & \\ & 0 & \\ & & \begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix} \\ & & & 0_{4\times 4} \end{pmatrix}$$

$$E_1^- = \begin{pmatrix} 0_{4\times 4} & \begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{smallmatrix} \\ 0_{4\times 4} & 0_{4\times 4} \end{pmatrix}, \quad E_7^- = \begin{pmatrix} 0_{4\times 4} & & 0_{4\times 4} \\ \begin{smallmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{smallmatrix} & & 0_{4\times 4} \end{pmatrix}$$

$$(3.2.11)$$

and the $\{E_i^+\}$ by

$$
E_2^+ = \begin{pmatrix} \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} & & \\ & 0 & \\ & & 0 \\ & & & 0_{4\times4} \end{pmatrix}, \quad
E_3^+ = \begin{pmatrix} 0_{4\times4} & \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \\ 0_{4\times4} & 0_{4\times4} \end{pmatrix}, \quad
E_4^+ = \begin{pmatrix} 0_{4\times4} & \\ & \begin{matrix} 0 & & \\ & 0 & -1 \\ & 0 & 0 \end{matrix} \\ & & 0 \end{pmatrix}
$$

$$
E_5^+ = \begin{pmatrix} 0_{4\times4} & 0_{4\times4} \\ \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} & 0_{4\times4} \end{pmatrix}, \quad
E_7^+ = \begin{pmatrix} \begin{matrix} 0 & \\ & 0 \\ & & \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} \end{matrix} \\ & 0_{4\times4} \end{pmatrix}
$$

$$
E_1^+ = \begin{pmatrix} 0_{4\times4} & 0_{4\times4} \\ \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} & 0_{4\times4} \end{pmatrix}, \quad
E_7^+ = \begin{pmatrix} 0_{4\times4} & \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{matrix} \\ 0_{4\times4} & 0_{4\times4} \end{pmatrix}
$$

$$(3.2.12)$$

We end this section by introducing the dilatation generator D, hypercharge B and the central charge C. They are defined as:

$$
D = -(H_3 + H_4 + H_5) - \frac{1}{2}(H_2 + H_6) = \frac{1}{2}\begin{pmatrix} \begin{matrix} 1 & \\ & 1 \\ & & -1 \\ & & & -1 \end{matrix} \\ & & 0_{4\times4} \end{pmatrix}
$$

$$
B = \frac{1}{2}\begin{pmatrix} \begin{matrix} 1 & \\ & 1 \\ & & 1 \\ & & & 1 \end{matrix} \\ & & 0_{4\times4} \end{pmatrix}
$$

$$
C = \frac{1}{2}(H_1 - H_3 + H_5 - H_7) = \frac{1}{2}\mathbb{I}_{8\times8}
$$

$$(3.2.13)$$

Notice that B does not belong to $\mathfrak{su}(2,2|4)$ and that C is the generator of the subalgebra $\mathfrak{u}(1)$ which means that C does not belong to $\mathfrak{psu}(2,2|4)$. Finally, these generators satisfy the following commutation relations:

$$
\begin{aligned}
[D, E_1^-] &= \frac{1}{2}E_1^-, & [B, E_1^-] &= \frac{1}{2}E_1^- \\
[D, E_3^-] &= -\frac{1}{2}E_3^-, & [B, E_3^-] &= -\frac{1}{2}E_3^- \\
[D, E_5^-] &= -\frac{1}{2}E_5^-, & [B, E_5^-] &= \frac{1}{2}E_5^- \\
[D, E_7^-] &= \frac{1}{2}E_7^-, & [B, E_7^-] &= -\frac{1}{2}E_7^- \\
[C, X] &= 0 \quad \text{for all } X \in \mathfrak{su}(2,2|4)
\end{aligned}
$$

$$(3.2.14)$$

3.2.2   Chevalley Basis II

Before we move on to the representation theory of $\mathfrak{su}(2,2|4)$, we will describe an alternative Chevalley basis $\{\tilde{H}_i, \tilde{E}_i^+, \tilde{E}_i^- | i = 1, \ldots, 7\}$ which is sometimes called the "Beauty" basis [6].

In this basis the Cartan generators are

$$
\tilde{H}_3 = \begin{pmatrix} 0_{4\times4} & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 & \\ & & & & 0 \end{pmatrix}, \quad
\tilde{H}_4 = \begin{pmatrix} 0_{4\times4} & & & \\ & 0 & & \\ & & 1 & \\ & & & -1 & \\ & & & & 0 \end{pmatrix}, \quad
\tilde{H}_5 = \begin{pmatrix} 0_{4\times4} & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 & \\ & & & & -1 \end{pmatrix}
$$

$$
\tilde{H}_1 = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 & \\ & & & & 0_{4\times4} \end{pmatrix}, \quad
\tilde{H}_7 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & -1 & \\ & & & 1 & \\ & & & & 0_{4\times4} \end{pmatrix}
$$

$$
\tilde{H}_2 = \begin{pmatrix} 0 & & & & \\ & -1 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & -1 & \\ & & & & & 0 \\ & & & & & & 0 \\ & & & & & & & 0 \end{pmatrix}, \quad
\tilde{H}_6 = \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 0 & \\ & & & & & 0 \\ & & & & & & 0 \\ & & & & & & & 1 \end{pmatrix}
\tag{3.2.15}
$$

and we see that they are related to the Cartan generators (3.2.7) by:

$$
\begin{aligned}
&\tilde{H}_1 = H_2, \quad \tilde{H}_2 = -(H_1 + H_2), \quad \tilde{H}_3 = H_1 + H_2 + H_3, \quad \tilde{H}_4 = H_4 \\
&\tilde{H}_7 = H_6, \quad \tilde{H}_6 = -(H_6 + H_7), \quad \tilde{H}_5 = H_5 + H_6 + H_7
\end{aligned}
\tag{3.2.16}
$$

Notice that $\{\tilde{H}_3, \tilde{H}_4, \tilde{H}_5\}$ form a Cartan subalgebra of $\mathfrak{su}(4)$ and that $\{\tilde{H}_1, \tilde{H}_7\}$ generate $\mathfrak{su}(2)\oplus\mathfrak{su}(2) \subset \mathfrak{su}(2,2)$. The simple root system $\tilde{\Pi}$ consists of even simple roots $\{\tilde{\alpha}_1, \tilde{\alpha}_3, \tilde{\alpha}_4, \tilde{\alpha}_5, \tilde{\alpha}_7\}$ and odd simple roots $\{\tilde{\alpha}_2, \tilde{\alpha}_6\}$ defined by:

$$
\begin{aligned}
&\alpha_1 = \epsilon_1 - \epsilon_2, \quad \alpha_2 = \epsilon_2 - \delta_1, \quad \alpha_3 = \delta_1 - \delta_2, \quad \alpha_4 = \delta_2 - \delta_3 \\
&\alpha_5 = \delta_3 - \delta_4, \quad \alpha_6 = \delta_4 - \epsilon_3, \quad \alpha_7 = \epsilon_3 - \epsilon_4
\end{aligned}
\tag{3.2.17}
$$

The (symmetric) Cartan matrix $\tilde{a}$ is given by

$$
\tilde{a} = \begin{pmatrix}
-2 & +1 & 0 & 0 & 0 & 0 & 0 \\
+1 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & -1 & +2 & -1 & 0 & 0 & 0 \\
0 & 0 & -1 & +2 & -1 & 0 & 0 \\
0 & 0 & 0 & -1 & +2 & -1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & +1 \\
0 & 0 & 0 & 0 & 0 & +1 & -2
\end{pmatrix}
\tag{3.2.18}
$$

and it is graphically represented by the following Dynkin diagram:

$$\tag{3.2.19}$$

We complete our Chevalley basis by giving the corresponding simple root generators. The $\{\tilde{E}_i^-\}$ are given by:

$$
\tilde{E}_3^- = \begin{pmatrix} 0_{4\times4} & & \\ & \begin{matrix} 0 & 0 \\ -1 & 0 \end{matrix} & \\ & & \begin{matrix} 0 & \\ & 0 \end{matrix} \end{pmatrix}, \quad
\tilde{E}_4^- = \begin{pmatrix} 0_{4\times4} & & \\ & 0 & \\ & & \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix} \\ & & & 0 \end{pmatrix}, \quad
\tilde{E}_5^- = \begin{pmatrix} 0_{4\times4} & & \\ & 0 & \\ & & 0 \\ & & & \begin{matrix} 0 & 0 \\ -1 & 0 \end{matrix} \end{pmatrix}
$$

$$
\tilde{E}_1^- = \begin{pmatrix} \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix} & & \\ & 0 & \\ & & 0 \\ & & & 0_{4\times4} \end{pmatrix}, \quad
\tilde{E}_7^- = \begin{pmatrix} 0 & & \\ & 0 & \\ & & \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix} \\ & & & 0_{4\times4} \end{pmatrix}
$$

$$
\tilde{E}_2^- = \begin{pmatrix} 0_{4\times4} & \begin{matrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \\ & 0_{4\times4} \end{pmatrix}, \quad
\tilde{E}_6^- = \begin{pmatrix} 0_{4\times4} & \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{matrix} \\ 0_{4\times4} & 0_{4\times4} \end{pmatrix} \tag{3.2.20}
$$

and the $\{\tilde{E}_i^+\}$ by

$$
\tilde{E}_3^+ = \begin{pmatrix} 0_{4\times4} & & \\ & \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} & \\ & & \begin{matrix} 0 & \\ & 0 \end{matrix} \end{pmatrix}, \quad
\tilde{E}_4^+ = \begin{pmatrix} 0_{4\times4} & & \\ & 0 & \\ & & \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \\ & & & 0 \end{pmatrix}, \quad
\tilde{E}_5^+ = \begin{pmatrix} 0_{4\times4} & & \\ & 0 & \\ & & 0 \\ & & & \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} \end{pmatrix}
$$

$$
\tilde{E}_1^+ = \begin{pmatrix} \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} & & \\ & 0 & \\ & & 0 \\ & & & 0_{4\times4} \end{pmatrix}, \quad
\tilde{E}_7^+ = \begin{pmatrix} 0 & & \\ & 0 & \\ & & \begin{matrix} 0 & -1 \\ 0 & 0 \end{matrix} \\ & & & 0_{4\times4} \end{pmatrix}
$$

$$
\tilde{E}_2^+ = \begin{pmatrix} 0_{4\times4} & \begin{matrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \\ 0_{4\times4} & 0_{4\times4} \end{pmatrix}, \quad
\tilde{E}_6^+ = \begin{pmatrix} 0_{4\times4} & 0_{4\times4} \\ \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{matrix} & 0_{4\times4} \end{pmatrix} \tag{3.2.21}
$$

In terms of the "Beauty" basis, the dilatation generator D, and the central charge C are written as:

$$
D = -\sum_{i=2}^{6} \tilde{H}_i - \frac{1}{2}(\tilde{H}_1 + \tilde{H}_7) = \frac{1}{2} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \\ & & & & 0_{4\times4} \end{pmatrix} \tag{3.2.22}
$$

$$
C = -\frac{1}{2}(\tilde{H}_1 + 2\tilde{H}_2 + \tilde{H}_3 - \tilde{H}_5 - 2\tilde{H}_6 - \tilde{H}_7) = \frac{1}{2}\mathbb{I}_{8\times8}
$$

These generators satisfy the commutation relations

$$
[D, \tilde{E}_2^-] = -\frac{1}{2}\tilde{E}_2^-, \qquad [D, \tilde{E}_6^-] = -\frac{1}{2}\tilde{E}_6^-
$$

$$
[B, \tilde{E}_2^-] = -\frac{1}{2}\tilde{E}_2^-, \qquad [B, \tilde{E}_6^-] = \frac{1}{2}\tilde{E}_6^- \tag{3.2.23}
$$

$$
[C, X] = 0 \quad \text{for all } X \in \mathfrak{su}(2,2|4)
$$

Furthermore, since

$$\tilde{H}_2 = -\left(\frac{3}{4}\tilde{H}_3 + \frac{1}{2}\tilde{H}_4 + \frac{1}{4}\tilde{H}_5 + \frac{1}{2}\tilde{H}_1 + \frac{1}{2}D + \frac{1}{2}C\right)$$

$$\tilde{H}_6 = -\left(\frac{1}{4}\tilde{H}_3 + \frac{1}{2}\tilde{H}_4 + \frac{3}{4}\tilde{H}_5 + \frac{1}{2}\tilde{H}_7 + \frac{1}{2}D - \frac{1}{2}C\right)$$

(3.2.24)

we can choose to work with D and C instead of $\tilde{H}_2$ and $\tilde{H}_6$.

3.3 Representation Theory of $\mathfrak{psu}(2,2|4)$

In this section we will consider highest weight representations constructed using an oscillator algebra.

So let $\{H_i, E_i^{\pm}\}$ be a Chevalley basis (not necessarily one from the previous section) of $\mathfrak{su}(2,2|4)$ and let $|\text{hwv}\rangle$ be the highest weight vector of a highest weight representation of $\mathfrak{su}(2,2|4)$. Then an oscillator algebra $\mathcal{A}$ is an algebra which acts on the highest weight vector, together with a set of commutation relations between the generators of the algebra and the Chevalley basis. One of the most famous examples of an oscillator algebra is the algebra generated by Chevalley elements $\{E_i^-\}$, and the constructed module is the Verma module. As a way of an example, we will work this out for the two Chevalley bases of the previous section.

3.3.1 The Verma Module for the Chevalley Basis I

Consider the Chevalley basis of section 3.2.1, and let us denote the highest weight vector $|\text{hwv}\rangle$ by $|\text{hwv}\rangle = |r_1, r_2, r_3, r_4, r_5, r_6, r_7\rangle$ where the (Dynkin) labels $[r_1, r_2, r_3, r_4, r_5, r_6, r_7]$ are defined by:

$$H_i|\text{hwv}\rangle = r_i|\text{hwv}\rangle \quad \text{for all} \quad i = 1, \ldots, 7 \tag{3.3.1}$$

Additionally, $|\text{hwv}\rangle$ will sometimes also contain the labels b, Δ or c which are defined as

$$B|\text{hwv}\rangle = b|\text{hwv}\rangle$$

$$D|\text{hwv}\rangle = \Delta|\text{hwv}\rangle \tag{3.3.2}$$

$$C|\text{hwv}\rangle = c|\text{hwv}\rangle$$

The labels Δ and c do not give any new information, since D and C can be expressed in terms of the Chevalley generators $\{H_i\}$. The element B however, is not contained in $\mathfrak{su}(2,2|4)$, so the label b does contains information about $|\text{hwv}\rangle$ which is not present in the r_i's.

Note that since $\{H_2, H_3, H_4\}$ and $\{H_4, H_5, H_6\}$ form two $\mathfrak{su}(2|2)$ subalgebras, the values $[r_2, r_3, r_4]$ and $[r_4, r_5, r_6]$ form the Dynkin labels of highest weight representations of $\mathfrak{su}(2|2)$. Reminding ourselves that $\mathfrak{psu}(2,2|4) = \mathfrak{su}(2,2|4)/\mathfrak{u}(1)$, we can make $|\text{hwv}\rangle$ a highest weight representation of $\mathfrak{psu}(2,2|4)$ by fixing $c = 0$. Using (3.2.13), this implies that the r_i's are restricted by the condition

$$r_1 - r_3 + r_5 - r_7 = 0 \tag{3.3.3}$$

Our Verma module is defined as the linear span of vectors of the form

$$|n_i\rangle = (E_1^-)^{n_1} \ldots (E_7^-)^{n_7} |\text{hwv}\rangle \tag{3.3.4}$$

Using relations (3.2.5), (3.2.9), (3.2.14), (3.3.1) and (3.3.2) we see that

$$H_j|n_i\rangle = (r_j - \sum_{k=1}^{7} a_{jk}n_k)|n_i\rangle \quad \text{for} \quad j = 1, \ldots, 7$$

$$B|n_i\rangle = \frac{1}{2}(2b + n_1 - n_3 + n_5 - n_7)|n_i\rangle \tag{3.3.5}$$

$$D|n_i\rangle = \frac{1}{2}(2\Delta + n_1 - n_3 - n_5 + n_7)|n_i\rangle$$

$$C|n_i\rangle = c|n_i\rangle$$

When we talk about the weight of vector $|n_i\rangle$, we will refer to $[r_1', r_2', r_3', r_4', r_5', r_6', r_7']$, where the label r_j' is defined as $H_j|n_i\rangle = r_j'|n_i\rangle$. Now notice that there are many orderings of the generators $\{E_i^-\}$ possible in $|n_i\rangle$ resulting in different vectors. However, all these vectors have the same weight.

We end this section by giving an example of a highest weight vector of $\mathfrak{psu}(2,2|4)$ which has significant physical meaning. Consider the primary field $\mathcal{Z}$ from the $\mathcal{N} = 4$ Super Yang-Mills theory [4] and suppose that we have a spin chain of length L for which $\mathcal{Z}$ is attached to each lattice site. The field $\mathcal{Z}$ is modeled by the highest weight vector $|\mathcal{Z}\rangle = |0, 0, 0, 1, 0, 0, 0; b = 0\rangle$ (notice the additional label b) and the spin chain is represented by the highest weight vector

$$|\text{hwv}\rangle = |0, 0, 0, L, 0, 0, 0; 0\rangle \tag{3.3.6}$$

(which is just the tensor product of $|\mathcal{Z}\rangle$). The vector $|n_i\rangle = (E_1^-)^{n_1} \ldots (E_7^-)^{n_7}|\text{hwv}\rangle$ then represents an excitation of this spin chain, and using (3.3.3) and (3.3.5) we see that $|n_i\rangle$ has weight $[r_1', r_2', r_3', r_4', r_5', r_6'; b']$ if

$$
\begin{aligned}
n_1 &= \frac{1}{4}(2b' + 2L - 3r_2' - 6r_3' - 2r_4' + 2r_5' - r_6' - 4r_7') \\
n_2 &= -r_3' + r_5' - r_7' \\
n_3 &= \frac{1}{4}(-2b' - 2L - r_2' - 2r_3' + 2r_4' + 6r_5' + r_6' - 4r_7') \\
n_4 &= r_5' - r_7' \\
n_5 &= \frac{1}{4}(2b' - 2L + r_2' + 2r_3' + 2r_4' + 2r_5' - r_6' - 4r_7') \\
n_6 &= -r_7' \\
n_7 &= \frac{1}{4}(-2b' + 2L - r_2' - 2r_3' - 2r_4' - 2r_5' - 3r_6' - 4r_7')
\end{aligned}
\tag{3.3.7}
$$

3.3.2 The Verma Module for the Chevalley Basis II

Now consider the Chevalley basis of section 3.2.2, and let us denote the highest weight vector $|\text{hwv}\rangle$ by $|\text{hwv}\rangle = |\Delta, c; s_1, s_2; q_1, p, q_2; b\rangle$ where the labels $[\Delta, c; s_1, s_2; q_1, p, q_2; b]$ are defined by (3.3.2) and:

$$
\begin{aligned}
\tilde{H}_1|\text{hwv}\rangle &= -s_1|\text{hwv}\rangle \\
\tilde{H}_3|\text{hwv}\rangle &= q_1|\text{hwv}\rangle \\
\tilde{H}_4|\text{hwv}\rangle &= p|\text{hwv}\rangle \\
\tilde{H}_5|\text{hwv}\rangle &= q_2|\text{hwv}\rangle \\
\tilde{H}_7|\text{hwv}\rangle &= -s_2|\text{hwv}\rangle
\end{aligned}
\tag{3.3.8}
$$

which, using (3.2.24), implies

$$
\begin{aligned}
\tilde{H}_2|\text{hwv}\rangle &= \frac{1}{2}\left(s_1 - \frac{3}{2}q_1 - p - \frac{3}{2}q_2 - \Delta - c\right)|\text{hwv}\rangle \\
\tilde{H}_6|\text{hwv}\rangle &= \frac{1}{2}\left(s_2 - \frac{1}{2}q_1 - p - \frac{3}{2}q_2 - \Delta + c\right)|\text{hwv}\rangle
\end{aligned}
\tag{3.3.9}
$$

Note that since $\{\tilde{H}_3, \tilde{H}_4, \tilde{H}_5\}$ form a $\mathfrak{su}(4)$ subalgebra, the values $[q_1, p, q_2]$ form the Dynkin labels of highest weight representations for $\mathfrak{su}(4)$. In the same way, s_1 and s_2 are Dynkin labels of highest weight representations for the $\mathfrak{su}(2)$ subalgebras. Furthermore, $|\text{hwv}\rangle$ is a highest weight vector of $\mathfrak{psu}(2, 2|4)$ if we fix $c = 0$.

The Verma module is defined as the linear span of vectors of the form

$$
|n_i\rangle = (\tilde{E}_1^-)^{n_1} \dots (\tilde{E}_7^-)^{n_7}|\text{hwv}\rangle
\tag{3.3.10}
$$

and using relations (3.2.5), (3.2.23), (3.3.8) and (3.3.9) we see that

$$
\begin{aligned}
\tilde{H}_1|n_i\rangle &= (2n_1 - n_2 - s_1)|n_i\rangle \\
\tilde{H}_2|n_i\rangle &= \frac{1}{2}\left(-2n_1 + 2n_3 + s_1 - \frac{3}{2}q_1 - p - \frac{1}{2}q_2 - \Delta - c\right)|n_i\rangle \\
\tilde{H}_3|n_i\rangle &= (n_2 - 2n_3 + n_4 + q_1)|n_i\rangle \\
\tilde{H}_4|n_i\rangle &= (n_3 - 2n_4 + n_5 + p)|n_i\rangle \\
\tilde{H}_5|n_i\rangle &= (n_4 - 2n_5 + n_6 + q_2)|n_i\rangle \\
\tilde{H}_6|n_i\rangle &= \frac{1}{2}\left(2n_5 - 2n_7 + s_2 - \frac{1}{2}q_1 - p - \frac{3}{2}q_2 - \Delta + c\right)|n_i\rangle \\
\tilde{H}_7|n_i\rangle &= (-n_6 + 2n_7 - s_2)|n_i\rangle \\
B|n_i\rangle &= \frac{1}{2}\left(2b - n_2 + n_6\right)|n_i\rangle \\
D|n_i\rangle &= \frac{1}{2}\left(2\Delta - n_2 - n_6\right)|n_i\rangle \\
C|n_i\rangle &= c|n_i\rangle
\end{aligned}
\tag{3.3.11}
$$

Finally, using (3.2.16) we see that (3.3.6) corresponds with $|\text{hwv}\rangle = |-L, 0; 0, 0; 0, L, 0; 0\rangle$. Because c is always zero for representations of $\mathfrak{psu}(2,2|4)$, we will omit the label for c, so $|\text{hwv}\rangle$ will be denoted by $|-L; 0, 0; 0, L, 0; 0\rangle$. Using (3.3.11) we then see that the vector $|n_i\rangle = (\tilde{E}_1^-)^{n_1} \ldots (\tilde{E}_7^-)^{n_7} |\text{hwv}\rangle$ has weight $[\Delta'; s_1', s_2'; q_1', p', q_2'; B']$ if

$$
\begin{aligned}
n_1 &= -\frac{1}{2}\left(s_1' + L + B' + \Delta'\right) \\
n_2 &= -(L + B' + \Delta') \\
n_3 &= -\frac{3}{4}q_1' - \frac{1}{2}p' - \frac{1}{4}q_2' - \frac{1}{2}L - \frac{1}{2}B' + \Delta' \\
n_4 &= -\frac{1}{2}q_1' - p' - \frac{1}{2}q_2' - \Delta' \\
n_5 &= -\frac{1}{4}q_1' - \frac{1}{2}p' - \frac{3}{4}q_2' - \frac{1}{2}L + \frac{1}{2}B' - \Delta' \\
n_6 &= B' - \Delta' - L \\
n_7 &= \frac{1}{2}\left(B' - s_2' - \Delta' - L\right)
\end{aligned}
\tag{3.3.12}
$$

Chapter 4

The Bethe Ansatz for the Centrally Extended $\mathfrak{su}(2|2)$ invariant S-matrix

The Superstring action on $\text{AdS}_5 \times \text{S}^5$ has the global symmetry $\mathfrak{psu}(2,2|4)$. However, because this action contains non-physical bosonic and fermionic degrees of freedom it is difficult to extract useful information out of it. Fortunately, this problem can be solved by fixing the gauge. A suitable gauge is the light-cone gauge and it has been shown in [2] that the gauge fixed action has the symmetry $\mathfrak{J} = \mathfrak{psu}(2|2) \oplus \mathfrak{psu}(2|2) \oplus \mathbf{H} \oplus \mathbf{C} \oplus \mathbf{C}^\dagger$, where $\mathbf{H}$ is the Hamiltonian of the system and $\mathbf{C}, \mathbf{C}^\dagger$ are central elements (they commute with every element $X \in \mathfrak{psu}(2,2|4)$) which are not contained in $\mathfrak{psu}(2,2|4)$. So $\mathfrak{J}$ is not a subalgebra of $\mathfrak{psu}(2,2|4)$. However, we do have that $\mathfrak{psu}(2|2) \oplus \mathbf{H} = \mathfrak{su}(2|2) \subseteq \mathfrak{psu}(2,2|4)$ so we will call $\mathfrak{su}(2|2)_{C,C^\dagger} \equiv \mathfrak{su}(2|2) \oplus \mathbf{C} \oplus \mathbf{C}^\dagger$ the centrally extended $\mathfrak{su}(2|2)$ Lie superalgebra. With a little abuse of notation, we will write $\mathfrak{J}$ as $\mathfrak{J} = \mathfrak{su}(2|2)_{C,C^\dagger} \oplus \mathfrak{su}(2|2)_{C,C^\dagger}$ where it is understood that the generators $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{C}^\dagger$ are shared by each copy of $\mathfrak{su}(2|2)_{C,C^\dagger}$.

Now that we have identified the symmetry algebra $\mathfrak{J}$ of the gauge fixed superstring action, we can construct the corresponding S-matrix using the invariance condition as outlined in section 2.7. We will call this S-matrix the $\mathfrak{J}$-invariant S-matrix. Because of the nice form of $\mathfrak{J}$, this S-matrix will be build as the tensor product of two smaller $\mathfrak{su}(2|2)_{C,C^\dagger}$-invariant S-matrices. But before we do that, we will first give a detailed description of $\mathfrak{su}(2|2)_{C,C^\dagger}$.

4.1 The Centrally Extended $\mathfrak{su}(2|2)$ Lie Superalgebra

The Lie superalgebra $\mathfrak{su}(2|2)_{C,C^\dagger}$ consists of the generators $\mathbf{L}_a{}^b$, $\mathbf{R}_\alpha{}^\beta$ of the $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ subalgebra, the generators $\mathbf{Q}_\alpha{}^a$, $\mathbf{Q}^\dagger{}_a{}^\alpha$ of the odd part of $\mathfrak{su}(2|2)$ and the central elements $\mathbf{H}$, $\mathbf{C}$ and $\mathbf{C}^\dagger$. Here, the Roman indices take the values $\{1, 2\}$ and the Greek indices the values $\{3, 4\}$. Furthermore, $\mathbf{L}_1{}^1 = -\mathbf{L}_2{}^2$ and $\mathbf{R}_1{}^1 = -\mathbf{R}_2{}^2$. The non-trivial commutation relations

between these generators are

$$
\begin{aligned}
&[\mathbf{L}_1{}^2, \mathbf{L}_2{}^1] = 2\mathbf{L}_1{}^1 , && [\mathbf{L}_1{}^1, \mathbf{L}_2{}^1] = -\mathbf{L}_2{}^1 , && [\mathbf{L}_1{}^1, \mathbf{L}_1{}^2] = \mathbf{L}_1{}^2 \\
&[\mathbf{R}_3{}^4, \mathbf{R}_4{}^3] = 2\mathbf{R}_3{}^3 , && [\mathbf{R}_3{}^3, \mathbf{R}_4{}^3] = -\mathbf{R}_4{}^3 , && [\mathbf{R}_3{}^3, \mathbf{R}_3{}^4] = \mathbf{R}_3{}^4 \\
&[\mathbf{L}_a{}^b, \mathbf{Q}^\dagger{}_c{}^\alpha] = \delta_{bc}\mathbf{Q}^\dagger{}_a{}^\alpha - \frac{1}{2}\delta_{ab}\mathbf{Q}^\dagger{}_c{}^\alpha , && [\mathbf{R}_\alpha{}^\beta, \mathbf{Q}_\gamma{}^a] = \delta_{\beta\gamma}\mathbf{Q}_\alpha{}^a - \frac{1}{2}\delta_{\alpha\beta}\mathbf{Q}_\gamma{}^a \\
&[\mathbf{L}_a{}^b, \mathbf{Q}_\gamma{}^c] = -\delta_{ac}\mathbf{Q}_\gamma{}^b + \frac{1}{2}\delta_{ab}\mathbf{Q}_\gamma{}^c , && [\mathbf{R}_\alpha{}^\beta, \mathbf{Q}^\dagger{}_a{}^\gamma] = -\delta_{\alpha\gamma}\mathbf{Q}^\dagger{}_a{}^\beta + \frac{1}{2}\delta_{\alpha\beta}\mathbf{Q}^\dagger{}_a{}^\gamma \\
&\{\mathbf{Q}_\alpha{}^a, \mathbf{Q}^\dagger{}_b{}^\beta\} = \delta_{ab}\mathbf{R}_\alpha{}^\beta + \delta_{\alpha\beta}\mathbf{L}_b{}^a + \frac{1}{2}\delta_{ab}\delta_{\alpha\beta}\mathbf{H} \\
&\{\mathbf{Q}_\alpha{}^a, \mathbf{Q}_\beta{}^b\} = \varepsilon_{\alpha\beta}\varepsilon^{ab}\mathbf{C} , && \{\mathbf{Q}^\dagger{}_a{}^\alpha, \mathbf{Q}^\dagger{}_b{}^\beta\} = \varepsilon_{ab}\varepsilon^{\alpha\beta}\mathbf{C}^\dagger
\end{aligned}
$$

$$(4.1.1)$$

where [,] and { , } denote the commutator and anti-commutator respectively.

Let us consider the Lie superalgebra $\mathfrak{su}(2|2)$ for a moment (so there is no central extension by $\mathbf{C}$ and $\mathbf{C}^\dagger$). As with the $\mathfrak{psu}(2,2|4)$ Lie superalgebra, we can describe $\mathfrak{su}(2|2)$ with a Chevalley basis $\{M_i, F_i^+, F_i^-\}$ which satisfy the relations

$$
[\![M_i, M_j]\!] = 0 , \qquad [\![F_i^+, F_j^-]\!] = \delta_{ij}M_j , \qquad [\![M_i, F_j^\pm]\!] = \pm a_{ij}F_j^\pm
\qquad (4.1.2)
$$

In this case, we will take the Cartan matrix as

$$
a = \begin{pmatrix} -2 & +1 & 0 \\ +1 & 0 & -1 \\ 0 & -1 & +2 \end{pmatrix}
\qquad (4.1.3)
$$

which admits the Dynkin diagram

$$(4.1.4)$$

and we will take the Chevalley generators as

$$
\begin{aligned}
M_1 &= -2\mathbf{L}_1{}^1 , & F_1^+ &= -\mathbf{L}_1{}^2 , & F_1^- &= \mathbf{L}_2{}^1 \\
M_2 &= \mathbf{L}_1{}^1 - \mathbf{R}_3{}^3 - \frac{1}{2}\mathbf{H} , & F_2^+ &= -\mathbf{Q}^\dagger{}_2{}^3 , & F_2^- &= \mathbf{Q}_3{}^2 \\
M_3 &= 2\mathbf{R}_3{}^3 , & F_3^+ &= -\mathbf{R}_3{}^4 , & F_3^- &= -\mathbf{R}_4{}^3
\end{aligned}
\qquad (4.1.5)
$$

Note that if we compare (4.1.3) with (3.2.9), we see that (4.1.3) is exactly the sub-matrix of (3.2.9) which starts at the $(2,2)$ position and ends at the $(4,4)$ position. So $(M_i, F_i^\pm)$ corresponds with $(H_i, E_i^\pm)$, $i = 2,3,4$ from 3.2.7, 3.2.11 and 3.2.12. This will become more clearly in the next paragraph when we write these generators out in matrix representation. It is also interesting to note that we can describe the second copy of $\mathfrak{su}(2|2) \subseteq \mathfrak{J}$ by a Chevalley basis with Cartan matrix

$$
\begin{pmatrix} +2 & -1 & 0 \\ -1 & 0 & +1 \\ 0 & +1 & -2 \end{pmatrix}
\qquad (4.1.6)
$$

In this case, (4.1.6) will be the sub-matrix of (3.2.9) which starts at the $(4, 4)$ position and ends at the $(6, 6)$ position, and its Chevalley basis will corresponds to $\{(H_i, E_i^{\pm})|i = 4, 5, 6\}$. So we see that the two copies of $\mathfrak{su}(2|2)$ have a natural embedding in $\mathfrak{psu}(2, 2|4)$.

Of course, we can also compare Cartan matrix (4.1.3) with (3.2.18). In this case, (4.1.3) coincides with the sub-matrix of (3.2.18) which starts at the $(1, 1)$ position and ends at the $(3, 3)$ position, and its Chevalley basis corresponds with $\{(\tilde{H}_i, \tilde{E}_i^{\pm})|i = 1, 2, 3\}$. Analogously, (4.1.6) coincides with the sub-matrix of (3.2.18) which starts at the $(5, 5)$ position and ends at the $(7, 7)$ position, and its Chevalley basis corresponds with $\{(\tilde{H}_i, \tilde{E}_i^{\pm})|i = 5, 6, 7\}$.

Let us now return to $\mathfrak{su}(2|2)_{C,C^{\dagger}}$. The representation of $\mathfrak{su}(2|2)_{C,C^{\dagger}}$ which we will be interested in is the fundamental representation. We will denote the representation space by $V \equiv V(p, \zeta)$ and the basis vectors as $|e_M\rangle \equiv |e_M(p, \zeta)\rangle$ where $M = \{a, \alpha\}$, $a = 1, 2$ and $\alpha = 3, 4$. The vectors $\{|e_a\rangle\}_{a=1,2}$ span $V_{\bar{0}}$, while the vectors $\{|e_{\alpha}\rangle\}_{a=3,4}$ span $V_{\bar{1}}$. Furthermore, the parameters p and ζ are complex numbers which parameterize the values of the central elements:

$$\mathbf{H}|e_M\rangle = H(p, \zeta)|e_M\rangle, \qquad \mathbf{C}|e_M\rangle = C(p, \zeta)|e_M\rangle, \qquad \mathbf{C}^{\dagger}|e_M\rangle = \bar{C}(p, \zeta)|e_M\rangle \qquad (4.1.7)$$

The other generators act on V as

$$
\begin{aligned}
\mathbf{L}_a{}^b|e_c\rangle &= \delta_c^b|e_a\rangle - \frac{1}{2}\delta_a^b|e_c\rangle, & \mathbf{R}_{\alpha}{}^{\beta}|e_{\gamma}\rangle &= \delta_{\gamma}^{\beta}|e_{\alpha}\rangle - \frac{1}{2}\delta_{\alpha}^{\beta}|e_{\gamma}\rangle \\
\mathbf{Q}_{\alpha}{}^a|e_b\rangle &= a\delta_b^a|e_{\alpha}\rangle, & \mathbf{Q}_{\alpha}{}^a|e_{\beta}\rangle &= b\varepsilon_{\alpha\beta}\varepsilon^{ab}|e_b\rangle & (4.1.8) \\
\mathbf{Q}^{\dagger}{}_a{}^{\alpha}|e_{\beta}\rangle &= d\delta_{\beta}^{\alpha}|e_a\rangle, & \mathbf{Q}^{\dagger}{}_a{}^{\alpha}|e_b\rangle &= c\varepsilon_{ab}\varepsilon^{\alpha\beta}|e_{\beta}\rangle
\end{aligned}
$$

Here, a, b, c, d are complex numbers which depend on p and ζ, and they satisfy the consistency condition $ad - bc = 1$ due to the commutation relations (4.1.1). Furthermore, the eigenvalues of the central elements are expressed in terms of a, b, c, d as

$$H(p, \zeta) = ad + bc, \qquad C(p, \zeta) = ab, \qquad \bar{C}(p, \zeta) = cd \qquad (4.1.9)$$

If we take

$$|e_1\rangle = (1, 0, 0, 0), \quad |e_2\rangle = (0, 1, 0, 0), \quad |e_3\rangle = (0, 0, 1, 0), \quad |e_4\rangle = (0, 0, 0, 1) \quad (4.1.10)$$

then the action of a generator $\mathbf{J} \in \mathfrak{su}(2|2)_{C,C^{\dagger}}$ can be written as

$$\mathbf{J}|e_M(p, \zeta)\rangle = J_{MN}(p, \zeta)|e_N(p, \zeta)\rangle \qquad (4.1.11)$$

where $J(p, \zeta)$ is the matrix form of the action. In particular, we have:

$$
L_2{}^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad
L_1{}^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad
L_1{}^1 = \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
R_4{}^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad
R_3{}^4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad
R_3{}^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & -1/2 \end{pmatrix} \tag{4.1.12}
$$

$$
Q^\dagger{}_1{}^3 = \begin{pmatrix} 0 & 0 & d & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & c & 0 & 0 \end{pmatrix}, \quad
Q^\dagger{}_2{}^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & 0 \\ -c & 0 & 0 & 0 \end{pmatrix}, \quad
Q^\dagger{}_1{}^4 = \begin{pmatrix} 0 & 0 & 0 & d \\ 0 & 0 & 0 & 0 \\ 0 & -c & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad
Q^\dagger{}_2{}^4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d \\ c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
Q_3{}^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b \\ a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad
Q_4{}^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -b & 0 \\ 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 \end{pmatrix}, \quad
Q_3{}^2 = \begin{pmatrix} 0 & 0 & 0 & -b \\ 0 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad
Q_4{}^2 = \begin{pmatrix} 0 & 0 & b & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \end{pmatrix}
$$

Note that these matrices reduce to the matrix form of $\mathfrak{su}(2|2)$ generators when $ab = cd = 0$. In particular, when we choose $b = c = 0$ and $a = d = 1$ we see that the matrix form of the Chevalley generators $\{M_i, F_i^\pm | i = 1, 2, 3\}$ can be obtained from the $\mathfrak{psu}(2, 2|4)$ generators $\{H_i, E_i^\pm | i = 2, 3, 4\}$ by taking the "star" entries of $(H_i, E_i^\pm)$ as indicated in the left matrix of:

$$
\begin{pmatrix} * & * & \cdot & \cdot & \cdot & * & * & \cdot \\ * & * & \cdot & \cdot & \cdot & * & * & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ * & * & \cdot & \cdot & \cdot & * & * & \cdot \\ * & * & \cdot & \cdot & \cdot & * & * & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad
\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \circ & \circ & \cdot & \circ & \circ & \cdot \\ \cdot & \cdot & \circ & \circ & \cdot & \circ & \circ & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \circ & \circ & \cdot & \circ & \circ & \cdot \\ \cdot & \cdot & \circ & \circ & \cdot & \circ & \circ & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad
\begin{pmatrix} * & * & \cdot & \cdot & * & * & \cdot & \cdot \\ * & * & \cdot & \cdot & * & * & \cdot & \cdot \\ \cdot & \cdot & \circ & \circ & \cdot & \cdot & \circ & \circ \\ \cdot & \cdot & \circ & \circ & \cdot & \cdot & \circ & \circ \\ * & * & \cdot & \cdot & * & * & \cdot & \cdot \\ * & * & \cdot & \cdot & * & * & \cdot & \cdot \\ \cdot & \cdot & \circ & \circ & \cdot & \cdot & \circ & \circ \\ \cdot & \cdot & \circ & \circ & \cdot & \cdot & \circ & \circ \end{pmatrix} \tag{4.1.13}
$$

Analogously, $\{M_i, F_i^\pm | i = 1, 2, 3\}$ can also be obtained from $\{\tilde{H}_i, \tilde{E}_i^\pm | i = 1, 2, 3\}$ by taking its "star" entries as indicated in the right matrix of (4.1.13).

We also note that the matrix form of the Chevalley generators corresponding to Cartan matrix (4.1.6) can be obtained from $\{H_i, E_i^\pm | i = 2, 3, 4\}$ and $\{\tilde{H}_i, E_i^\pm | i = 2, 3, 4\}$ respectively, by taking the "circle" entries of the middle and the right matrix of (4.1.13).

To relate the above discussion with the string sigma model, we will express a, b, c, d in terms of the string theory parameters g, x^+, x^-, ζ, η as follows:

$$
a = \sqrt{g}\eta, \qquad b = \sqrt{g}\frac{i\zeta}{\eta}\left(\frac{x^+}{x^-} - 1\right), \qquad c = -\sqrt{g}\frac{\eta}{\zeta x^+}, \qquad d = \sqrt{g}\frac{x^+}{i\eta}\left(1 - \frac{x^-}{x^+}\right) \tag{4.1.14}
$$

Here, the parameter g is the string sigma model coupling constant, p is the world-sheet momenta and η is a free parameter which reflects a freedom in the choice of the basis vectors $|e_M\rangle$. Next, the consistency condition $ad - bc = 1$ translates into

$$
x^+ + \frac{1}{x^+} - x^- - \frac{1}{x^-} = \frac{i}{g} \tag{4.1.15}
$$

and the parameters $x^\pm$ are related to p by

$$\frac{x^+}{x^-} = e^{ip} \qquad (4.1.16)$$

In string theory we will be interested in fundamental unitary representations which are characterized by the conditions $c = \bar{b}$, $d = \bar{a}$ and $p \in \mathbb{R}$. To achieve this we will set

$$\zeta = e^{2i\xi}, \qquad \eta = \sqrt{ix^- - ix^+}\, e^{i\xi} \qquad (4.1.17)$$

where $\xi \in \mathbb{R}$. Then a and b become

$$a = \sqrt{g}\,\sqrt{ix^- - ix^+}\, e^{i\xi} = \sqrt{g}\eta, \qquad b = -\sqrt{g}\,\frac{\sqrt{ix^- - ix^+}}{x^-}\, e^{i\xi} \qquad (4.1.18)$$

which implies that

$$H(p) = \sqrt{1 + 16g^2 \sin^2\left(\frac{1}{2}p\right)}, \qquad C(p,\zeta) = ig\zeta(e^{ip} - 1) \qquad (4.1.19)$$

Finally, let us note that this representation simplifies to a $\mathfrak{su}(2|2)$ representation when we take $p = 0$.

4.2 The $\mathfrak{su}(2|2)_{C,C^\dagger}$ invariant S-matrix

As in section 2.7, we will describe an n-particle state of the string theory by

$$|A^\dagger_{i_1}(p_1)\ldots A^\dagger_{i_n}(p_n)\rangle \equiv A^\dagger_{i_1}(p_1)\ldots A^\dagger_{i_n}(p_n)|0\rangle \qquad (4.2.1)$$

Reminding ourselves that the symmetry algebra of the light-cone string theory consist of two copies of $\mathfrak{su}(2|2)_{C,C^\dagger}$ we can write $A^\dagger_i(p) = A^\dagger_{M,\dot{M}}(p)$, where the indices M, $\dot{M}$ correspond to the first and second copy of $\mathfrak{su}(2|2)_{C,C^\dagger}$ respectively. In fact, we can think of $A^\dagger_{M,\dot{M}}(p)$ being the tensor product $A^\dagger_{M,\dot{M}}(p) = A^\dagger_M(p) \otimes A^\dagger_{\dot{M}}(p)$ which act on the vacuum $|0\rangle = |0\rangle_1 \otimes |0\rangle_2$. With this notation it becomes clear that the complete S-matrix is build as the tensor product of two $\mathfrak{su}(2|2)_{C,C^\dagger}$-invariant S-matrices. Therefore, we will continue our discussion by restricting to the first tensor component $|A^\dagger_{M_1}(p_1)\ldots A^\dagger_{M_n}(p_n)\rangle$, of the states

$$A^\dagger_{M_1,\dot{M}_1}(p_1)\ldots A^\dagger_{M_n,\dot{M}_n}(p_n)|0\rangle \equiv |A^\dagger_{M_1}(p_1)\ldots A^\dagger_{M_n}(p_n)\rangle \otimes |A^\dagger_{\dot{M}_1}(p_1)\ldots A^\dagger_{\dot{M}_n}(p_n)\rangle \qquad (4.2.2)$$

To construct the $\mathfrak{su}(2|2)_{C,C^\dagger}$-invariant S-matrix we will view $|A^\dagger_{M_1}(p_1)\ldots A^\dagger_{M_n}(p_n)\rangle$ as an element of the tensor product of fundamental unitary representations. This identification is not trivial and we will have to impose certain conditions on our parameter ζ in order to have a consistent identification. The reason for this is because the operator $\mathbf{C}$ acts differently on

$A^\dagger_{M_1}(p_1)\ldots A^\dagger_{M_n}(p_n)|0\rangle$ than on $V(p_1, \zeta_1)\bar{\otimes}\ldots\bar{\otimes}V(p_n, \zeta_n)$. In the first case, it was shown in [2] that the central element $\mathbf{C}$ is written as

$$\mathbf{C} = ig\left(e^{i\mathbf{P}} - 1\right) \tag{4.2.3}$$

where $\mathbf{P}$ is the total momentum operator defined as

$$\mathbf{P}|0\rangle = 0\,, \qquad \mathbf{P}A^\dagger_M(p) = A^\dagger_M(p)(\mathbf{P} + p) \tag{4.2.4}$$

In the second case, $\mathbf{C}$ acts with the usual tensor product action (3.1.7).

   With these definitions, it is clear that the one particle state $|A^\dagger_M(p)\rangle$ is identified with the basis vector $|e_M\rangle$ of the fundamental unitary representation $V(p, 1)$. It becomes a bit more complicated when we try to identify the two-particle state $|A^\dagger_{M_1}(p_1)A^\dagger_{M_2}(p_2)\rangle$ with an element of $V(p_1, \zeta_1) \bar{\otimes} V(p_2, \zeta_2)$. Because

$$\mathbf{C}|A^\dagger_{M_1}(p_1)A^\dagger_{M_2}(p_2)\rangle = ig(e^{i(p_1+p_2)} - 1)|A^\dagger_{M_1}(p_1)A^\dagger_{M_2}(p_2)\rangle$$
$$\mathbf{C}\left(V(p_1, \zeta_1)\bar{\otimes} V(p_2, \zeta_2)\right) = ig\left(\zeta_1(e^{ip_1} - 1) + \zeta_2(e^{ip_2} - 1)\right)V(p_1, \zeta_1)\bar{\otimes} V(p_2, \zeta_2) \tag{4.2.5}$$

we must choose ζ_1 and ζ_2 such that

$$ig(e^{i(p_1+p_2)} - 1) = ig\left(\zeta_1(e^{ip_1} - 1) + \zeta_2(e^{ip_2} - 1)\right) \tag{4.2.6}$$

To satisfy this equation we will choose $\zeta_1 = e^{ip_2}$, $\zeta_2 = 1$. Then, the two-particle state $|A^\dagger_{M_1}(p_1)A^\dagger_{M_2}(p_2)\rangle$ will be identified with $|e_{M_1}(p_1, e^{ip_1})\rangle \bar{\otimes} |e_{M_1}(p_2, 1)\rangle$. Furthermore, we will also have the commutation relation

$$\mathbf{C}A^\dagger_M(p) = C(p)A^\dagger_M(p)\,e^{i\mathbf{P}} + A^\dagger_M(p)\mathbf{C} \qquad \text{where} \qquad C(p) \equiv C(p, 1) \tag{4.2.7}$$

   We can easily generalize our construction to a n-particle state. In this case we want to identify $|A^\dagger_{M_1}(p_1)\ldots A^\dagger_{M_n}(p_n)\rangle$ with an element of $V(p_1, \zeta_1)\bar{\otimes}\ldots\bar{\otimes}V(p_n, \zeta_n)$. Letting $\mathbf{C}$ act on the particle states and the tensor product representation yields the constraint

$$e^{i(p_1+\ldots+p_n)} - 1 = \sum_{k=1}^n \zeta_k(e^{ip_k} - 1) \tag{4.2.8}$$

Because we also want (4.2.7) to be satisfied, we will choose the ζ's as

$$\zeta_1 = e^{i(p_2+\ldots+p_n)}\,, \quad \zeta_2 = e^{i(p_3+\ldots+p_n)}\,, \quad \ldots \quad, \quad \zeta_{n-1} = e^{ip_n}\,, \quad \zeta_n = 1 \tag{4.2.9}$$

and we obtain the identification

$$|A^\dagger_{M_1}(p_1)\ldots A^\dagger_{M_n}(p_n)\rangle \sim |e_{M_1}(p_1, e^{i(p_2+\ldots+p_n)})\rangle \bar{\otimes} |e_{M_1}(p_2, e^{i(p_3+\ldots+p_n)})\rangle \bar{\otimes} \ldots \bar{\otimes} V(p_n, 1) \tag{4.2.10}$$

   Now that we have identified the particle states with the tensor product of fundamental unitary representations, we will now look at how the action of $\mathfrak{su}(2|2)_{C,C^\dagger}$ generators is

carried over to the particle states. First we note that because of the grading, a generator $\mathbf{J} \in \mathfrak{su}(2|2)_{C,C^\dagger}$ will act in a graded way on the tensor product representation (see (3.1.7)). Next, we will write

$$A^\dagger(p) \equiv \left(A_1^\dagger(p), A_2^\dagger(p), A_3^\dagger(p), A_4^\dagger(p)\right) \quad \text{and} \quad A(p) \equiv \left(A_1(p), A_2(p), A_3(p), A_4(p)\right)^T$$
$$(4.2.11)$$

(so that $A^\dagger(p)$ is a row vector and $A(p)$ is a column vector). Furthermore, we will define the action of $\mathfrak{su}(2|2)_{C,C^\dagger}$ on the vacuum vector $|0\rangle$ to be the trivial action, i.e. $\mathbf{J}|0\rangle = 0$ for all $\mathbf{J} \in \mathfrak{su}(2|2)_{C,C^\dagger}$. Then, from (4.2.10), (4.2.9) and (4.1.8) we will get the commutation relations:

$$\begin{aligned}
\mathbf{L}_a{}^b A^\dagger(p) &= A^\dagger(p) L_a{}^b + A^\dagger(p) \mathbf{L}_a{}^b \\
\mathbf{R}_\alpha{}^\beta A^\dagger(p) &= A^\dagger(p) R_\alpha{}^\beta + A^\dagger(p) \mathbf{R}_\alpha{}^\beta \\
\mathbf{Q}_\alpha{}^a A^\dagger(p) &= A^\dagger(p) Q_\alpha{}^a(p, 1) e^{i\mathbf{P}/2} + A^\dagger(p) \Sigma \mathbf{Q}_\alpha{}^a \\
\mathbf{Q}^\dagger{}_a{}^\alpha A^\dagger(p) &= A^\dagger(p) Q^\dagger{}_a{}^\alpha(p, 1) e^{-i\mathbf{P}/2} + A^\dagger(p) \Sigma \mathbf{Q}^\dagger{}_a{}^\alpha
\end{aligned}$$
$$(4.2.12)$$

where Σ, defined in (3.2.3), is to account for the grading.

We can now construct the desired S-matrix. Using the identification between the particle states and the tensor product of fundamental unitary representations, we can view the S-matrix as a map

$$\mathbf{S} : V(p_1, \zeta_1) \,\bar\otimes\, V(p_2, \zeta_2) \to V(p_2, \zeta_2) \,\bar\otimes\, V(p_1, \zeta_1) \qquad (4.2.13)$$

where $\zeta_1 = e^{ip_2}$ and $\zeta_2 = 1$. Furthermore, the invariance condition (2.7.6) can be written as

$$S_{1,2}(p_1, p_2) \left(J \otimes \mathbb{I} + \mathbb{I} \otimes J\right) = \left(J \otimes \mathbb{I} + \mathbb{I} \otimes J\right) S_{1,2}(p_1, p_2) \qquad (4.2.14)$$

when $J = L_a{}^b$, $R_\alpha{}^\beta$, and

$$\begin{aligned}
S_{1,2}(p_1, p_2) \left(J(p_1, e^{ip_2}) \,\bar\otimes\, \mathbb{I} + \mathbb{I} \,\bar\otimes\, J(p_2, 1)\right) &= \left(J(p_1, 1) \,\bar\otimes\, \mathbb{I} + \mathbb{I} \,\bar\otimes\, J(p_2, e^{ip_1})\right) S_{1,2}(p_1, p_2) \quad \text{or} \\
S_{1,2}(p_1, p_2) \left(J(p_1, e^{ip_2}) \otimes \mathbb{I} + \Sigma \otimes J(p_2, 1)\right) &= \left(J(p_1, 1) \otimes \Sigma + \mathbb{I} \otimes J(p_2, e^{ip_1})\right) S_{1,2}(p_1, p_2)
\end{aligned}$$
$$(4.2.15)$$

when $J = Q_\alpha{}^a$, $Q^\dagger{}_a{}^\alpha$. This condition together with the Yang-Baxter equation determines the S-matrix up to a constant scalar factor which we will denote by $S_0(p_1, p_2)$. This calcu-

lation has been done in [3] and the resulting S-matrix $S_{1,2}(p_1, p_2)$ is given by

$$
S_{1,2}(p_1, p_2) \equiv \left(
\begin{array}{cccc|cccc|cccc|cccc}
a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & a_1+a_2 & 0 & 0 & -a_2 & 0 & 0 & 0 & 0 & 0 & 0 & a_7 & 0 & 0 & -a_7 & 0 \\
0 & 0 & a_5 & 0 & 0 & 0 & 0 & 0 & a_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & a_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_9 & 0 & 0 & 0 \\
\hline
0 & -a_2 & 0 & 0 & a_1+a_2 & 0 & 0 & 0 & 0 & 0 & 0 & -a_7 & 0 & 0 & a_7 & 0 \\
0 & 0 & 0 & 0 & 0 & a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & a_5 & 0 & 0 & a_9 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & a_5 & 0 & 0 & 0 & 0 & 0 & a_9 & 0 & 0 \\
\hline
0 & 0 & a_{10} & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & a_{10} & 0 & 0 & a_6 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_3 & 0 & 0 & 0 & 0 & 0 \\
0 & -a_8 & 0 & 0 & a_8 & 0 & 0 & 0 & 0 & 0 & 0 & a_3+a_4 & 0 & 0 & -a_4 & 0 \\
\hline
0 & 0 & 0 & a_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{10} & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 \\
0 & a_8 & 0 & 0 & -a_8 & 0 & 0 & 0 & 0 & 0 & 0 & -a_4 & 0 & 0 & a_3+a_4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_3
\end{array}
\right)
$$

$$(4.2.16)$$

where $a_i = S_0(p_1, p_2)\,\tilde{a}_i$

$$
\tilde{a}_1 = \frac{x_2^- - x_1^+}{x_2^+ - x_1^-}\frac{e^{\frac{i}{2}p_2}}{e^{\frac{i}{2}p_1}}
$$

$$
\tilde{a}_2 = \frac{(x_1^- - x_1^+)(x_2^- - x_2^+)(x_2^- + x_1^+)}{(x_1^- - x_2^+)(x_1^- x_2^- - x_1^+ x_2^+)}\frac{e^{\frac{i}{2}p_2}}{e^{\frac{i}{2}p_1}}
$$

$$
\tilde{a}_3 = -1
$$

$$
\tilde{a}_4 = \frac{(x_1^- - x_1^+)(x_2^- - x_2^+)(x_1^- + x_2^+)}{(x_1^- - x_2^+)(x_1^- x_2^- - x_1^+ x_2^+)}
$$

$$
\tilde{a}_5 = \frac{x_2^- - x_1^-}{x_2^+ - x_1^-}e^{\frac{i}{2}p_2}
$$

$$
\tilde{a}_6 = \frac{x_1^+ - x_2^+}{x_1^- - x_2^+}\frac{1}{e^{\frac{i}{2}p_1}}
$$

$$
\tilde{a}_7 = \frac{\sqrt{(x_1^- - x_1^+)(x_2^- - x_2^+)}(x_1^+ - x_2^+)}{(x_1^- - x_2^+)(1 - x_1^- x_2^-)e^{\frac{i}{2}p_1}}
$$

$$
\tilde{a}_8 = \frac{(x_1^+ - x_2^+)\sqrt{(x_1^- - x_1^+)(x_2^- - x_2^+)}}{(x_1^- - x_2^+)(x_1^- x_2^- - 1)e^{ip_1}e^{\frac{i}{2}p_2}}
$$

$$(4.2.17)$$

$$
\tilde{a}_9 = \frac{\sqrt{(x_1^- - x_1^+)(x_2^- - x_2^+)}}{x_2^+ - x_1^-}
$$

$$
\tilde{a}_{10} = \frac{\sqrt{(x_1^- - x_1^+)(x_2^- - x_2^+)}}{x_2^+ - x_1^-}\frac{e^{\frac{i}{2}p_2}}{e^{\frac{i}{2}p_1}}
$$

with $x_j^+/x_j^- = x^+(p_j)/x^-(p_j) = e^{ip_j}$. Now that we have obtained the $\mathfrak{su}(2|2)_{C,C^\dagger}$ invariant S-matrix $S_{1,2}(p_1, p_2)$, we can simply obtain the complete S-matrix $S_{1,2}^*(p_1, p_2)$ of the light-cone string theory by setting

$$
S_{1,2}^*(p_1, p_2) = S_{1,2}(p_1, p_2) \otimes S_{1,2}(p_1, p_2) \tag{4.2.18}
$$

4.3　The Bethe Ansatz Equations of the $\mathfrak{su}(2|2)_{C,C^\dagger}$ invariant S-matrix

As we have seen in the section on the XXX model and the Hubbard model, the Bethe ansatz equations follow from imposing boundary conditions on the integrable model. In

this section we will derive the Bethe equations using the nested algebraic Bethe ansatz and the nested coordinate Bethe ansatz.

4.3.1 Deriving the Bethe Ansatz Equations using the Nested Algebraic Bethe Ansatz

Before we derive the Bethe equations, we first define the matrix $\bar{S}$ which is just the matrix S under a different basis:

$$\bar{S}_{1,2}(p_1, p_2) = [G(p_1) \otimes G(p_2)] \, S_{1,2}(p_1, p_2) \, [G^{-1}(p_1) \otimes G^{-1}(p_2)]$$

$$G(p) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & t(p) \\ 0 & 0 & t(p) & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{4.3.1}$$

Furthermore, we will introduce the R-matrix $\bar{R}_{1,2}(p_1, p_2) \equiv P_{1,2}\bar{S}_{1,2}(p_1, p_2)$ where $P_{1,2}$ is the permutation operator defined as $P_{1,2} = \sum_{a,b=1}^{4} e_a^b \otimes e_b^a$, with $e_i^j \equiv \delta_{ij}$. So $\bar{R}_{1,2}(p_1, p_2)$ can be viewed as the 16×16 matrix:

$$\bar{R}_{1,2}(p_1, p_2) \equiv \left(\begin{array}{cccc|cccc|cccc|cccc} a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a'_{10} & 0 & 0 & a_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a'_{10} & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_2 & 0 & 0 & -a'_7 & 0 & 0 & a'_7 & 0 & 0 & a_1+a_2 & 0 & 0 & 0 \\ \hline 0 & a_5 & 0 & 0 & a'_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a'_8 & 0 & 0 & -a_4 & 0 & 0 & a_3+a_4 & 0 & 0 & a'_8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a'_9 & 0 & 0 & 0 & 0 & 0 & a_5 & 0 & 0 \\ \hline 0 & 0 & a_5 & 0 & 0 & 0 & 0 & 0 & a'_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a'_8 & 0 & 0 & a_3+a_4 & 0 & 0 & -a_4 & 0 & 0 & -a'_8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a'_9 & 0 & 0 & a_5 & 0 \\ \hline 0 & 0 & 0 & a_1+a_2 & 0 & 0 & a'_7 & 0 & 0 & -a'_7 & 0 & 0 & -a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 & 0 & 0 & 0 & a'_{10} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 & a'_{10} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_1 \end{array}\right)$$

where $\quad a'_7 = \dfrac{a_7}{t(p_1)t(p_2)}, \quad a'_8 = a_8\, t(p_1)t(p_2), \quad a'_9 = a_9\dfrac{t(p_2)}{t(p_1)}, \quad a'_{10} = a_{10}\dfrac{t(p_1)}{t(p_2)}$ (4.3.2)

Note that the basis transformation due to $G(p)$ caused a change in the gradings for the basis vectors; $|e_1\rangle$, $|e_4\rangle$ are now even and $|e_2\rangle$, $|e_3\rangle$ are now odd. We also note that $\bar{S}_{1,2}(p_1, p_2)$ satisfies the Yang Baxter equation (A.2.3).

If we now take

$$x^+(p) = i\frac{a(p)}{b(p)}e^{2h(p)}, \qquad x^-(p) = -i\frac{b(p)}{a(p)}e^{2h(p)}, \qquad t(p) = \left(\frac{x^+(p)}{x^-(p)}\right)^{1/4} \tag{4.3.3}$$

where the functions $a(p)$, $b(p)$ and $h(p)$ satisfy the constraints

$$a^2(p) + b^2(p) = 1, \qquad \sinh[2h(p)] = \frac{a(p)b(p)}{2g} \tag{4.3.4}$$

due to the constraints (4.1.15) and (4.1.16), we will obtain the relation

$$\frac{\bar{R}_{1,2}(p_1, p_2)}{\bar{R}_{1,2}(p_1, p_2)_{1,1}^{1,1}} = \frac{\bar{R}_{1,2}^{(s)}(p_1, p_2)}{\bar{R}_{1,2}^{(s)}(p_1, p_2)_{1,1}^{1,1}} \tag{4.3.5}$$

where $\bar{R}_{1,2}^{(s)}(p_1, p_2)$ is Shastry's graded R-matrix (2.6.8) with $c = -g/2$. Therefore, following from the similarity between $\bar{R}_{1,2}(p_1, p_2)$ and $\bar{R}_{1,2}^{(s)}(p_1, p_2)$ we can use the NABA of the Hubbard model to obtain the desired Bethe equations.

To see how this works, we first define the graded S-matrix $\tilde{S}_{1,2}(p_1, p_2) = P_{1,2}^g \bar{R}_{1,2}(p_1, p_2)$ where $P_{1,2}^g = \sum_{a,b=1}^4 (-1)^{p(a)p(b)} e_a^b \otimes e_b^a$. Next, we define the monodromy matrix $T_a(\lambda, \{p_i\})$ and transfer matrix $\tau(\lambda, \{p_i\})$ as

$$T_a(\lambda, \{p_i\}) = \tilde{S}_{a,N}(\lambda, p_N)\, \tilde{S}_{a,N-1}(\lambda, p_{N-1}) \ldots \tilde{S}_{a,1}(\lambda, p_1)$$
$$\tau(\lambda, \{p_i\}) = \mathrm{str}_a[T_a(\lambda, \{p_i\})] \tag{4.3.6}$$

The monodromy matrix satisfies the fundamental commutation relation

$$\bar{R}_{a,b}(\lambda, \mu)\, [T_a(\lambda, \{p_i\} \hat{\otimes} T_b(\lambda, \{p_i\})] = [T_a(\mu, \{p_i\}) \hat{\otimes} T_b(\lambda, \{p_i\}]\, \bar{R}_{a,b}(\lambda, \mu) \tag{4.3.7}$$

where $\hat{\otimes}$ is the graded tensor product.

We now remind ourselves that the Bethe equations are obtained by solving the auxiliary eigenvalue problem and imposing a condition on the eigenvalue. For the model of section 2.5 these were (2.5.13) and (2.5.17), and for the Hubbard model (in the two-excitation case) these were (2.6.37) and (2.6.43). Here, we will do the same. The auxiliary eigenvalue problem is

$$\tau(\lambda, \{p_i\}_{i=1}^N)|\Psi\rangle = \Lambda(\lambda, \{p_i\}_{i=1}^N)|\Psi\rangle \tag{4.3.8}$$

and the eigenvalue condition will be

$$\left(\frac{x_k^+}{x_k^-}\right)^L = \Lambda(p_k, \{p_i\}_{i=1}^N) \tag{4.3.9}$$

which, in the spin chain picture, has the physical interpretation that when you shift the k-th particle around the spin chain it will acquire a free phase factor $e^{ip_k L} = (x_k^+/x_k^-)^L$.

We will now focus our attention on the eigenvalue problem (4.3.9). To solve this eigenvalue equation we will generalize the results of the NABA of the Hubbard model by stating the expressions in terms of matrix elements $\bar{R}(\lambda, \mu)_{a,c}^{b,d}$.

First we note that the monodromy matrix is a 4×4 matrix in auxiliary space of the form (2.6.11). Secondly, note that the $\tilde{S}_{a,i}(\lambda, p_i)$ plays the role of the Lax operator. The vacuum vector we will use is

$$|0\rangle = \overbrace{\bigotimes}^N_{j=1} |0\rangle_j \quad \text{with} \quad |0\rangle_i \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}_i \tag{4.3.10}$$

Therefore

$$
\tilde{S}_{a,i}(\lambda,\mu)|0\rangle_i =
\begin{pmatrix}
\tilde{S}(\lambda,\mu)_{1,1}^{1,1}|0\rangle_i & * & * & * \\
0 & \tilde{S}(\lambda,\mu)_{2,1}^{2,1}|0\rangle_i & 0 & * \\
0 & 0 & \tilde{S}(\lambda,\mu)_{3,1}^{3,1}|0\rangle_i & * \\
0 & 0 & 0 & \tilde{S}(\lambda,\mu)_{4,1}^{4,1}|0\rangle_i
\end{pmatrix}
\tag{4.3.11}
$$

where $\tilde{S}(\lambda,\mu)_{2,1}^{2,1} = \tilde{S}(\lambda,\mu)_{3,1}^{3,1}$. Now because $\tilde{S}_{1,2}(p_1,p_2) = P_{1,2}^g \bar{R}_{1,2}(p_1,p_2)$ we can also write

$$
\bar{R}(\lambda,\mu)_{1,1}^{1,1} = \tilde{S}(\lambda,\mu)_{1,1}^{1,1}, \quad \bar{R}(\lambda,\mu)_{1,2}^{2,1} = \tilde{S}(\lambda,\mu)_{2,1}^{2,1}, \quad \bar{R}(\lambda,\mu)_{1,4}^{4,1} = \tilde{S}(\lambda,\mu)_{4,1}^{4,1}
\tag{4.3.12}
$$

Therefore

$$
T_a(\lambda,\{p_i\})|0\rangle =
\begin{pmatrix}
\prod_{i=1}^{N} \tilde{R}(\lambda,p_i)_{1,1}^{1,1} & * & * & * \\
0 & \prod_{i=1}^{N} \tilde{R}(\lambda,p_i)_{1,2}^{2,1} & 0 & * \\
0 & 0 & \prod_{i=1}^{N} \tilde{R}(\lambda,p_i)_{1,2}^{2,1} & * \\
0 & 0 & 0 & \prod_{i=1}^{N} \tilde{R}(\lambda,p_i)_{1,4}^{4,1}
\end{pmatrix}|0\rangle
\tag{4.3.13}
$$

So $|0\rangle$ is an eigenvector of (4.3.8) with eigenvalue

$$
\Lambda(\lambda) = \prod_{i=1}^{N} \bar{R}(\lambda,p_i)_{1,1}^{1,1} - 2\prod_{i=1}^{N} \bar{R}(\lambda,p_i)_{1,2}^{2,1} + \prod_{i=1}^{N} \bar{R}(\lambda,p_i)_{1,4}^{4,1}
\tag{4.3.14}
$$

For the K_1-particle eigenvector we will again use the form (2.6.23):

$$
|\Phi_{K_1}(\lambda_1,\ldots,\lambda_{K_1})\rangle = \boldsymbol{\Phi}_{K_1}(\lambda_1,\ldots,\lambda_{K_1})\cdot\mathbf{F}|0\rangle
\tag{4.3.15}
$$

The expressions for the one and two-particle wave vectors are analogous to (2.6.24) and (2.6.34):

$$
\boldsymbol{\Phi}_1(\lambda_1) = \mathbf{B}(\lambda_1), \qquad \boldsymbol{\Phi}_2(\lambda_1,\lambda_2) = \mathbf{B}(\lambda_1)\otimes\mathbf{B}(\lambda_2) + \boldsymbol{\xi} F(\lambda_1) B(\lambda_2)\frac{\bar{R}(\lambda_1,\lambda_j)_{1,4}^{2,3}}{\bar{R}(\lambda_1,\lambda_j)_{1,4}^{4,1}}
\tag{4.3.16}
$$

Furthermore, the K_1-particle wave vector $\boldsymbol{\Phi}_{K_1}(\lambda_1,\ldots,\lambda_{K_1})$ is of the form

$$
\boldsymbol{\Phi}_{K_1}(\lambda_1,\ldots,\lambda_{K_1}) = \mathbf{B}(\lambda_1)\otimes\boldsymbol{\Phi}_{K_1-1}(\lambda_2,\ldots,\lambda_{K_1}) + \sum_{j=2}^{K_1}\frac{\bar{R}(\lambda_1,\lambda_j)_{1,4}^{2,3}}{\bar{R}(\lambda_1,\lambda_j)_{1,4}^{4,1}}\prod_{\substack{k=2\\k\neq j}}^{K_1}\frac{\bar{R}(\lambda_k,\lambda_j)_{1,1}^{1,1}}{\bar{R}(\lambda_k,\lambda_j)_{1,2}^{2,1}}
$$

$$
\times\left[\boldsymbol{\xi}\otimes F(\lambda_1)\boldsymbol{\Phi}_{K_1-2}(\lambda_2,\ldots,\lambda_{j-1},\lambda_{j+1},\ldots\lambda_{K_1})B(\lambda_j)\right]\prod_{l=2}^{j-1}\frac{\bar{R}(\lambda_l,\lambda_j)_{2,2}^{2,2}}{\bar{R}(\lambda_l,\lambda_j)_{1,1}^{1,1}}\hat{r}_{l,l+1}(\lambda_l,\lambda_j)
$$

$$
\tag{4.3.17}
$$

$$\text{where}\quad \boldsymbol{\xi} = (0,1,-1,0)\quad\text{and}\quad \hat{r}(\lambda,\mu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \bar{a}(\lambda,\mu) & \bar{b}(\lambda,\mu) & 0 \\ 0 & \bar{b}(\lambda,\mu) & \bar{a}(\lambda,\mu) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{with}\quad \bar{a}(\lambda,\mu) = \frac{\bar{R}(\lambda,\mu)^{2,3}_{2,3}\bar{R}(\lambda,\mu)^{4,1}_{1,4} + \bar{R}(\lambda,\mu)^{2,3}_{1,4}\bar{R}(\lambda,\mu)^{4,1}_{2,3}}{\bar{R}(\lambda,\mu)^{2,2}_{2,2}\bar{R}(\lambda,\mu)^{4,1}_{1,4}} \tag{4.3.18}$$

$$\text{and}\quad \bar{b}(\lambda,\mu) = \frac{\bar{R}(\lambda,\mu)^{3,2}_{2,3}\bar{R}(\lambda,\mu)^{4,1}_{1,4} + \bar{R}(\lambda,\mu)^{3,2}_{1,4}\bar{R}(\lambda,\mu)^{4,1}_{2,3}}{\bar{R}(\lambda,\mu)^{2,2}_{2,2}\bar{R}(\lambda,\mu)^{4,1}_{1,4}}$$

Note that these expressions reduce exactly to the Hubbard model expressions (2.6.24), (2.6.34), (2.6.48) and (2.6.18) when we take $\bar{R}(\lambda,\mu)$ as $\bar{R}^s(\lambda,\mu)$.

If we now carry out the NABA analogously to the hubbard model case, we will obtain the eigenvalue $\Lambda(\lambda,\{p_i\}) = \Lambda(\lambda;\{p_i\}_{i=1}^N, \{\lambda_j\}_{j=1}^{K_1}, \{\mu_k\}_{k=1}^{K_2})$:

$$\Lambda(\lambda,\{p_i\}) = \prod_{i=1}^N \bar{R}(\lambda,p_i)^{1,1}_{1,1} \prod_{j=1}^{K_1} \frac{\bar{R}(\lambda_j,\lambda)^{1,1}_{1,1}}{\bar{R}(\lambda_j,\lambda)^{2,1}_{1,2}} + \prod_{i=1}^N \bar{R}(\lambda,p_i)^{4,1}_{1,4} \prod_{j=1}^{K_1} \frac{\bar{R}(\lambda,\lambda_j)^{4,2}_{2,4}}{\bar{R}(\lambda,\lambda_j)^{4,1}_{1,4}}$$

$$- \Lambda^{(1)}(\lambda;\{\lambda_i\},\{\mu_j\}) \prod_{i=1}^N \bar{R}(\lambda,p_i)^{2,1}_{1,2} \prod_{j=1}^{K_1} -\frac{\bar{R}(\lambda,\lambda_j)^{2,2}_{2,2}}{\bar{R}(\lambda,\lambda_j)^{2,1}_{1,2}}$$

$$\text{where}\quad \Lambda^{(1)}(\lambda;\{\lambda_i\},\{\mu_j\}) = \prod_{j=1}^{K_2} \frac{1}{\bar{b}(\mu_j,\lambda)} + \frac{\prod_{j=1}^{K_1}\bar{b}(\lambda,\lambda_j)}{\prod_{l=1}^{K_2}\bar{b}(\lambda,\mu_l)}\quad\text{with}\quad 1 \le K_2 \le K_1 \tag{4.3.19}$$

together with the Bethe equations

$$\prod_{l=1}^{K_1} \bar{b}(\mu_j,\lambda_l) = \prod_{\substack{l=1\\l\ne j}}^{K_2} \frac{\bar{b}(\mu_j,\mu_l)}{\bar{b}(\mu_l,\mu_j)}\quad\text{for all}\quad j = 1,\ldots,K_2 \tag{4.3.20}$$

$$\prod_{i=1}^N \frac{\bar{R}(\lambda_l,p_i)^{1,1}_{1,1}}{\bar{R}(\lambda_l,p_i)^{2,1}_{2,1}} = \Lambda^{(1)}(\lambda_l;\{\lambda_i\},\{\mu_j\}) \prod_{\substack{k=1\\k\ne l}}^{K_1} -\frac{\bar{R}(\lambda_l,\lambda_k)^{2,2}_{2,2}\,\bar{R}(\lambda_k,\lambda_l)^{2,1}_{2,1}}{\bar{R}(\lambda_l,\lambda_k)^{2,1}_{2,1}\,\bar{R}(\lambda_k,\lambda_l)^{1,1}_{1,1}}\quad\text{for all}\quad l = 1,\ldots,K_1$$

Let us note that the parameter $t(\lambda)$ is canceled out in all these expressions. If we now plug in the expressions for $\bar{R}(\lambda,\mu)$ explicitly and use (4.1.15), then after some algebra, (4.3.20) becomes

$$\prod_{l=1}^{K_1} \frac{w_j - y_l - \frac{1}{y_l} + \frac{i}{2g}}{w_j - y_l - \frac{1}{y_l} - \frac{i}{2g}} = \prod_{\substack{l=1\\l\ne j}}^{K_2} \frac{w_j - w_l + \frac{i}{g}}{w_j - w_l - \frac{i}{g}}\quad\text{for all}\quad j = 1,\ldots,K_2 \tag{4.3.21}$$

$$\prod_{i=1}^N \frac{y_l - x_i^-}{y_l - x_i^+}e^{i\frac{p_i}{2}} = \prod_{k=1}^{K_2} \frac{y_l + \frac{1}{y_l} - w_k + \frac{i}{2g}}{y_l + \frac{1}{y_l} - w_k - \frac{i}{2g}}\quad\text{for all}\quad l = 1,\ldots,K_1$$

$$\text{where}\quad x_i^\pm \equiv x^\pm(p_i),\quad y_i \equiv x^+(\lambda_i),\quad w_i \equiv x^+(\mu_i) + \frac{1}{x^+(\mu_i)} - \frac{i}{2g} = x^-(\mu_i) + \frac{1}{x^-(\mu_i)} + \frac{i}{2g}$$

Furthermore, let us also note that all these expressions reduce exactly to the Hubbard model expressions when we take $\bar{R}(\lambda,\mu)$ as $\bar{R}^s(\lambda,\mu)$. In particular

$$\frac{\bar{R}^{(s)}(\lambda_l,\lambda_k)_{2,2}^{2,2}\,\bar{R}^{(s)}(\lambda_k,\lambda_l)_{2,1}^{2,1}}{\bar{R}^{(s)}(\lambda_l,\lambda_k)_{2,1}^{2,1}\,\bar{R}^{(s)}(\lambda_k,\lambda_l)_{1,1}^{1,1}} = 1 \tag{4.3.22}$$

Returning to our eigenvalue condition (4.3.9) we see that

$$\left(\frac{x_k^+}{x_k^-}\right)^L = \prod_{i=1}^{N} S_0(p_k,p_i)\frac{x_i^- - x_k^+}{x_i^+ - x_k^-}\frac{e^{\frac{i}{2}p_i}}{e^{\frac{i}{2}p_k}} \cdot \prod_{j=1}^{K_1} e^{\frac{i}{2}p_k}\frac{x_k^- - y_j}{x_k^+ - y_j} \tag{4.3.23}$$

where we have used (4.3.19), (4.3.2), (4.2.17) and $\bar{R}(p_k,p_k)_{1,2}^{2,1} = \bar{R}(p_k,p_k)_{1,4}^{4,1} = 0$. By using $x_j^+/x_j^- = e^{ip_j}$ and $P \equiv \sum_{i=1}^{N} p_i$ we can simplify (4.3.23) to

$$e^{ip_k\left(L+\frac{N}{2}-\frac{K_1}{2}\right)} = e^{\frac{i}{2}P}\prod_{i=1}^{N} S_0(p_k,p_i)\frac{x_i^- - x_k^+}{x_i^+ - x_k^-} \cdot \prod_{j=1}^{K_1}\frac{x_k^- - y_j}{x_k^+ - y_j} \tag{4.3.24}$$

Equation (4.3.24) together with (4.3.21) form the complete Bethe equations of the $\mathfrak{su}(2|2)_{C,C^\dagger}$ invariant S-matrix. If we additionally impose the constraint $e^{iP/2} = 1$ we will obtain the Bethe equations of the $\mathfrak{su}(2|2)$ invariant S-matrix, since this constraint forces $\mathbf{C} = \mathbf{C}^\dagger = 0$ (see the remark at the end of section 4.1). Furthermore, this constraint also singles out the physical Bethe vectors.

4.3.2 Deriving the Bethe Ansatz Equations using the Nested Coordinate Bethe Ansatz

In this section, we will derive the Bethe ansatz equations using the alternative view of the nested coordinate Bethe ansatz [9] as sketched in section 2.4. The central object to be considered here is the S-matrix (4.2.16) and we will construct solutions of the equation

$$S_{i,j}(p_i,p_j)|\Psi\rangle = \lambda\,|\Psi\rangle_{(i,j)}, \qquad \lambda \in \mathbb{C} \tag{4.3.25}$$

to obtain the Bethe ansatz equations. Here, $|\Psi\rangle$ is a particle state, and $|\Psi\rangle_{(i,j)}$ is $|\Psi\rangle$ with the pseudo-momenta p_i and p_j interchanged.

Let us start the NCBA by defining the first level system. The first level system is the original spin chain of length L and the pseudo-vacuum just the empty spin chain. We will denote this by:

$$|0\rangle^{\mathrm{I}} \equiv |\bullet\ldots\bullet\rangle^0 \tag{4.3.26}$$

where "$\bullet$" stands for an empty lattice point. A general level I excitation will be of the form

$$|A_{M_1}^\dagger(p_1)_{i_1}\ldots A_{M_N}^\dagger(p_N)_{i_N}\rangle^{\mathrm{I}} \equiv |\ldots\bullet A_{M_1}^\dagger(p_1)_{i_1}\bullet\ldots\bullet A_{M_N}^\dagger(p_N)_{i_N}\bullet\ldots\rangle^0 \tag{4.3.27}$$

where $1 \le i_j \le L$ denotes the position of $A^\dagger_{M_j}$ and N is the total number of excitations out of the level I vacuum. So we see that with this notation, a general state $|A^\dagger_{M_1}(p_1) \dots A^\dagger_{M_N}(p_N)\rangle$ is viewed as the level I excitation $|A^\dagger_{M_1}(p_1)_{i_1} \dots A^\dagger_{M_N}(p_N)_{i_N}\rangle^{\mathrm{I}}$.

For the second level system we have a spin chain of length N, and the pseudo-vacuum $|0\rangle^{\mathrm{II}}$ is just the empty spin chain which is given by:

$$|0\rangle^{\mathrm{II}} = |A^\dagger_1(p_1) \dots A^\dagger_1(p_N)\rangle^{\mathrm{I}} \tag{4.3.28}$$

and from (4.2.16) we see that

$$S_{i,j}|0\rangle^{\mathrm{II}} = s^{\mathrm{I,I}}(p_i, p_j)|0\rangle^{\mathrm{II}}_{(i,j)} \quad \text{where} \quad s^{\mathrm{I,I}}(p_i, p_j) = a_1(p_i, p_j) = S_0(p_i, p_j)\frac{x_i^+ - x_j^-}{x_i^- - x_j^+}\frac{e^{\frac{i}{2}p_j}}{e^{\frac{i}{2}p_i}} \tag{4.3.29}$$

Let us consider the excitations on this level. In general, a vector on this spin chain will contain $A^\dagger_\alpha$ and $A^\dagger_2$. However, if we consider the one-excitation states only, we will have to be a bit more careful; it will only consist of $A^\dagger_\alpha$ since from (4.2.16) we have that

$$S_{1,2}(p_1, p_2)|A^\dagger_1(p_1)A^\dagger_2(p_2)\rangle = -a_8(p_1, p_2)\left(|A^\dagger_3(p_2)A^\dagger_4(p_1)\rangle - |A^\dagger_4(p_2)A^\dagger_3(p_1)\rangle\right)$$
$$\left(a_1(p_1, p_2) + a_2(p_1, p_2)\right)|A^\dagger_1(p_2)A^\dagger_2(p_1)\rangle - a_2(p_1, p_2)|A^\dagger_2(p_2)A^\dagger_1(p_1)\rangle \tag{4.3.30}$$

So $A^\dagger_2$ behaves like a double excitation since it scatters into a $A^\dagger_3$ and $A^\dagger_4$. Therefore, a general one-excitation solution is of the form

$$|A^\dagger_\alpha\rangle^{\mathrm{II}} = \sum_{k=1}^{N} \psi_k^{(1)}|A^\dagger_1(p_1) \dots A^\dagger_1(p_{k-1})A^\dagger_\alpha(p_k)A^\dagger_1(p_{k+1}) \dots A^\dagger_1(p_N)\rangle^{\mathrm{I}} \tag{4.3.31}$$

where it is understood that $A^\dagger_M(p_i)$ sits on position $1 \le i \le N$ of the second level spin chain. For the wave function $\psi_k^{(1)}$ we use the ansatz

$$\psi_k^{(1)} = f(x_k) \prod_{l=1}^{k-1} s^{\mathrm{II,I}}(x_l) \tag{4.3.32}$$

The wave function $\psi_k^{(1)}$ is determined by the compatibility condition

$$S_{i,j}|A^\dagger_\alpha\rangle^{\mathrm{II}} = s^{\mathrm{I,I}}(p_i, p_j)|A^\dagger_\alpha(y)\rangle^{\mathrm{II}}_{(i,j)} \tag{4.3.33}$$

where $|A^\dagger_\alpha\rangle^{\mathrm{II}}_{(i,j)}$ is $|A^\dagger_\alpha\rangle^{\mathrm{II}}$ with p_i and p_j interchanged. Considering this on a spin chain with only two sites gives:

$$|A^\dagger_\alpha\rangle^{\mathrm{II}} = f(x_1)|A^\dagger_\alpha(p_1)A^\dagger_1(p_2)\rangle^{\mathrm{I}} + f(x_2)s^{\mathrm{II,I}}(x_1)|A^\dagger_1(p_1)A^\dagger_\alpha(p_2)\rangle^{\mathrm{I}}$$
$$|A^\dagger_\alpha\rangle^{\mathrm{II}}_{(1,2)} = f(x_1)|A^\dagger_\alpha(p_1)A^\dagger_1(p_2)\rangle^{\mathrm{I}} + f(x_2)s^{\mathrm{II,I}}(x_1)|A^\dagger_1(p_1)A^\dagger_\alpha(p_2)\rangle^{\mathrm{I}} \tag{4.3.34}$$

Writing out the compatibility condition then gives:

$$
\begin{aligned}
f(x_1)\,a_9(p_1,p_2) + f(x_2)\,s^{\mathrm{II,I}}(x_1)\,a_5(p_1,p_2) &= f(x_2)\,a_1(p_1,p_2) \\
f(x_1)\,a_6(p_1,p_2) + f(x_2)\,s^{\mathrm{II,I}}(x_1)\,a_{10}(p_1,p_2) &= f(x_1)\,s^{\mathrm{II,I}}(x_2)\,a_1(p_1,p_2)
\end{aligned}
\tag{4.3.35}
$$

From the first equation we can get an expression for $s^{\mathrm{II,I}}(x_1)$ in terms of $x_1^\pm$ and f when we fix $x_2^+ = c$ (and therefore x_2^- is also fixed since $x^\pm$ are related with each other). Plugging this expression into the second equation enables us to solve $f(x_1)$ in terms of x_1^-, x_2^- and $f(y,x_2)$. By differentiating this equation to x_2^- we get a differential equation which we can solve. The resulting expression is

$$
f(x_j) \equiv f(y,x_j) = e^{-\frac{i}{2}p_j}\,\frac{y\,\sqrt{i(x_j^- - x_j^+)}}{y - x_j^-}
\tag{4.3.36}
$$

where y is an integration constant. From this we obtain

$$
s^{\mathrm{II,I}}(x_j) \equiv s^{\mathrm{II,I}}(y,x_j) = e^{-\frac{i}{2}p_j}\,\frac{y - x_j^+}{y - x_j^-}
\tag{4.3.37}
$$

and we will also write

$$
|A_\alpha^\dagger\rangle^{\mathrm{II}} = |A_\alpha^\dagger(y)\rangle^{\mathrm{II}}
\tag{4.3.38}
$$

to emphasize the y-dependence. Note that as in section 2.4.2, we can view $|A_\alpha^\dagger(y)\rangle^{\mathrm{II}}$ as an excitation out of the level II pseudo-vacuum with pseudo-momentum y.

We will now move on to the two excitation case. Here however, we have to be a bit careful since $|A_2^\dagger\rangle^{\mathrm{II}}$ causes a two-excitation. Therefore, we will use the following general ansatz for the two-excitation solution

$$
\begin{aligned}
|A_\alpha^\dagger(y_1)A_\beta^\dagger(y_2)\rangle^{\mathrm{II}} &= |\alpha,y_1;\beta,y_2\rangle_*^{\mathrm{II}} + \varepsilon^{\alpha\beta}|y_1,y_1\rangle_*^{\mathrm{II}} + S_{1,2}^{\mathrm{II}}|\alpha,y_1;\beta,y_2\rangle_*^{\mathrm{II}} \quad \text{where} \\
|\alpha,y_1;\beta,y_2\rangle_*^{\mathrm{II}} &= \sum_{1\le k<l\le N} \psi_k^{(1)}(y_1)\psi_l^{(1)}(y_2)|\ldots A_1^\dagger A_\alpha^\dagger(p_k)A_1^\dagger\ldots A_1^\dagger A_\beta^\dagger(p_l)A_1^\dagger\ldots\rangle^{\mathrm{I}} \\
S_{1,2}^{\mathrm{II}}|\alpha,y_1;\beta,y_2\rangle_*^{\mathrm{II}} &= M(y_1,y_2)|\alpha,y_2;\beta,y_1\rangle_*^{\mathrm{II}} + N(y_1,y_2)|\beta,y_2;\alpha,y_1\rangle_*^{\mathrm{II}} \\
\text{and} \quad |y_1,y_1\rangle_*^{\mathrm{II}} &= \sum_{k=1}^{N} \psi_k^{(1)}(y_1)\psi_k^{(1)}(y_2)g(y_1,y_2,x_k)|\ldots A_1^\dagger A_2^\dagger(p_k)A_1^\dagger\ldots\rangle^{\mathrm{I}}
\end{aligned}
\tag{4.3.39}
$$

and we recognize $S_{1,2}^{\mathrm{II}}$ as the second level S-matrix. Considering this again on a spin chain of two sites and using the compatibility condition

$$
S_{i,j}|A_\alpha^\dagger(y_1)A_\beta^\dagger(y_2)\rangle^{\mathrm{II}} = s^{\mathrm{I,I}}(p_i,p_j)|A_\alpha^\dagger(y_1)A_\beta^\dagger(y_2)\rangle_{(i,j)}^{\mathrm{II}}
\tag{4.3.40}
$$

will give us [9]:

$$
\begin{aligned}
g(y_1,y_2,x_k) &= \frac{y_1 y_2 - x_k^+ x_k^-}{y_1 y_2 x_k^-}\,\frac{y_1 - y_2}{v_1 - v_2 - \frac{i}{g}} \quad \text{where} \quad v_k \equiv y_k + \frac{1}{y_k} \\
M(y_1,y_2) &= \frac{\frac{i}{g}}{v_1 - v_2 - \frac{i}{g}}, \qquad N(y_1,y_2) = -\frac{v_1 - v_2}{v_1 - v_2 - \frac{i}{g}}
\end{aligned}
\tag{4.3.41}
$$

To generalize this construction to more than two excitations, we will make use of the generators of $\mathfrak{su}(2|2)_{C,C^\dagger}$ as discussed in [5] and [9]. First, let $\mathbf{J} \in \mathfrak{su}(2|2)_{C,C^\dagger}$ and define $\mathbf{J}_k$ as

$$\mathbf{J}_k \, V(p_1, \zeta_1) \bar{\otimes} \ldots \bar{\otimes} V(p_n, \zeta_n) = \tag{4.3.42}$$

$$= V(p_1, \zeta_1) \bar{\otimes} \ldots \bar{\otimes} \mathbf{J} \, V(p_k, \zeta_k) \bar{\otimes} \ldots \bar{\otimes} V(p_n, \zeta_n) \tag{4.3.43}$$

Then, using (4.1.12), (4.2.10), (4.2.9) and reminding ourselves that $|A_M^\dagger(p)\rangle$ is identified with the basis vector $|e_M\rangle$, we see that

$$\left(\mathbf{Q}_\alpha{}^1\right)_k |0\rangle^{\mathrm{II}} = a_k |\ldots A_1^\dagger(p_{k-1}) A_\alpha^\dagger(p_k) A_1^\dagger(p_{k+1}) \ldots\rangle^{\mathrm{I}}$$

$$\left(\mathbf{Q}_\alpha{}^1\right)_k \left(\mathbf{Q}_\beta{}^1\right)_l |0\rangle^{\mathrm{II}} = a_k a_l |\ldots A_1^\dagger(p_{k-1}) A_\alpha^\dagger(p_k) A_1^\dagger(p_{k+1}) \ldots A_1^\dagger(p_{l-1}) A_\beta^\dagger(p_l) A_1^\dagger(p_{l+1}) \ldots\rangle^{\mathrm{I}}$$

$$\left(\mathbf{Q}_\alpha{}^1\right)_k \left(\mathbf{Q}_\beta{}^1\right)_k |0\rangle^{\mathrm{II}} = a_k b_k \varepsilon^{\alpha\beta} |\ldots A_1^\dagger(p_{k-1}) A_2^\dagger(p_k) A_1^\dagger(p_{k+1}) \ldots\rangle^{\mathrm{I}}$$

where $\quad \varepsilon^{34} = 1, \quad a_k = \sqrt{g} \sqrt{i x_k^- - i x_k^+} \, e^{\frac{i}{2}(p_{k+1} + \ldots + p_N)} \quad$ and $\quad b_k = -\dfrac{a_k}{x_k^-} \tag{4.3.44}$

Let us now introduce the dressed generators

$$\left(\mathbf{Q}_\alpha{}^1\right)_k^\pm = e^{i\frac{\mathbf{P}}{2}} \frac{x_k^\pm}{x_k^\pm - x_k^\mp} \left(\mathbf{Q}_\alpha{}^1\right)_k \tag{4.3.45}$$

Then, by using the identity

$$\frac{y}{y - x_k^-} = \frac{x_k^+}{x_k^+ - x_k^-} + \frac{x_k^-}{x_k^- - x_k^+} \frac{y - x_k^+}{y - x_k^-} \tag{4.3.46}$$

we see that the one excitation solution (4.3.31) can be written as

$$|A_\alpha^\dagger\rangle^{\mathrm{II}} = \sum_{k=1}^N \left(\Phi_k(y) \left(\mathbf{Q}_\alpha{}^1\right)_k^- + \Phi_{k-1}(y) \left(\mathbf{Q}_\alpha{}^1\right)_k^+ \right) |0\rangle^{\mathrm{II}} \qquad \text{where}$$

$$\Phi_k(y) \equiv \prod_{l=1}^k \frac{y - x_l^+}{y - x_l^-} \qquad \text{and} \qquad \Phi_0(y) \equiv 1 \tag{4.3.47}$$

So let us define

$$\mathfrak{Q}_{\alpha,k}(y) \equiv \Phi_k(y) \left(\mathbf{Q}_\alpha{}^1\right)_k^- + \Phi_{k-1}(y) \left(\mathbf{Q}_\alpha{}^1\right)_k^+ \tag{4.3.48}$$

The two-excitation state (4.3.39) can then be written as:

$$|A^\dagger_\alpha(y_1)A^\dagger_\beta(y_2)\rangle^{\mathrm{II}} = |\alpha, y_1; \beta, y_2\rangle^{\mathrm{II}} + S^{\mathrm{II}}_{1,2}|\alpha, y_1; \beta, y_2\rangle^{\mathrm{II}} \quad \text{where}$$

$$|\alpha, y_1; \beta, y_2\rangle^{\mathrm{II}} = \Big\| \sum_{k,l=1}^{N} \mathfrak{Q}_{\alpha,k}(y_1)\mathfrak{Q}_{\beta,l}(y_2) \Big\| |0\rangle^{\mathrm{II}} \equiv$$

$$\equiv \sum_{1\le k<l\le N} \left(\Phi_k(y_1)\left(\mathbf{Q}_\alpha{}^1\right)^-_k + \Phi_{k-1}(y_1)\left(\mathbf{Q}_\alpha{}^1\right)^+_k\right)\left(\Phi_l(y_2)\left(\mathbf{Q}_\beta{}^1\right)^-_l + \Phi_{l-1}(y_2)\left(\mathbf{Q}_\beta{}^1\right)^+_l\right)|0\rangle^{\mathrm{II}}$$

$$+ \sum_{k=1}^{N}\frac{1}{2}\left(\Phi_k(y_1)\Phi_k(y_2)\left(\mathbf{Q}_\alpha{}^1\right)^-_k\left(\mathbf{Q}_\beta{}^1\right)^-_k + \Phi_{k-1}(y_1)\Phi_{k-1}(y_2)\left(\mathbf{Q}_\alpha{}^1\right)^+_k\left(\mathbf{Q}_\beta{}^1\right)^+_k\right)|0\rangle^{\mathrm{II}}$$

$$+ \sum_{k=1}^{N}\Phi_{k-1}(y_1)\left(\mathbf{Q}_\alpha{}^1\right)^+_k\Phi_k(y_2)\left(\mathbf{Q}_\beta{}^1\right)^-_k|0\rangle^{\mathrm{II}} \qquad \text{and}$$

$$S^{\mathrm{II}}_{1,2}|\alpha, y_1; \beta, y_2\rangle^{\mathrm{II}} = M(y_1, y_2)|\alpha, y_2; \beta, y_1\rangle^{\mathrm{II}} + N(y_1, y_2)|\beta, y_2; \alpha, y_1\rangle^{\mathrm{II}} \tag{4.3.49}$$

Here, the notation $\|\Psi\|$ is to denote the "ordered" version of Ψ. In the case of (4.3.49), we have a ordering in the pseudo-momenta y_i; the terms are arranged in such a way that y_1 stays left of y_2. Furthermore, if we compare (4.3.49) with (4.3.39), we see that the term $\varepsilon^{\alpha\beta}|y_1, y_1\rangle^{\mathrm{II}}_*$ has been distributed between the terms $|\alpha, y_1; \beta, y_2\rangle^{\mathrm{II}}$ and $S^{\mathrm{II}}_{1,2}|\alpha, y_1; \beta, y_2\rangle^{\mathrm{II}}$.

With these notations, we can easily see how to generalize this construction to K_1-excitations analogous to (2.4.110):

$$|A^\dagger_\alpha(y_1)\ldots A^\dagger_\beta(y_{K_1})\rangle^{\mathrm{II}} = |\alpha_1, y_1; \ldots; \alpha_{K_1}, y_{K_1}\rangle^{\mathrm{II}} + \sum_{P\in S_{K_1}} \mathbf{S}^{\mathrm{II}}_P |\alpha_1, y_1; \ldots; \alpha_{K_1}, y_{K_1}\rangle^{\mathrm{II}}$$

$$\text{where} \quad |\alpha_1, y_1; \ldots; \alpha_{K_1}, y_{K_1}\rangle^{\mathrm{II}} \equiv \Big\| \sum_{\{l_i\}} \mathfrak{Q}_{\alpha_1,l_1}(y_1)\ldots\mathfrak{Q}_{\alpha_{K_1},l_{K_1}}(y_{K_1}) \Big\| |0\rangle^{\mathrm{II}} \tag{4.3.50}$$

We now move on to the third level. The third level system is a spin chain of length K_1 whose pseudo-vacuum $|0\rangle^{\mathrm{III}}$ is the empty spin chain, which is given by

$$|0\rangle^{\mathrm{III}} = |A^\dagger_3(y_1)\ldots A^\dagger_3(y_{K_1})\rangle^{\mathrm{II}} \tag{4.3.51}$$

and using (4.3.39) we see that

$$S^{\mathrm{II}}_{i,j}|0\rangle^{\mathrm{III}} = s^{\mathrm{II,II}}(y_i, y_j)|0\rangle^{\mathrm{III}}_{(i,j)} \quad \text{with} \quad s^{\mathrm{II,II}}(y_i, y_j) = 1 \tag{4.3.52}$$

where $|0\rangle^{\mathrm{II}}_{(i,j)}$ is $|0\rangle^{\mathrm{II}}$ with y_i and y_j interchanged. Here (and in the following calculations) a minus sign is added to M and N to account for the fermionic nature of the particles.

Let us consider the excitations on this level. Analogous to the calculations for the second level system, a one-excitation solution is of the form

$$|A^\dagger_4(w)\rangle^{\mathrm{III}} = \sum_{k=1}^{K_1} \psi^{(2)}_k(w)|A^\dagger_3(y_1)\ldots A^\dagger_3(y_{k-1})A^\dagger_4(y_k)A^\dagger_3(y_{k+1})\ldots A^\dagger_3(y_{K_1})\rangle^{\mathrm{II}}$$

$$\tag{4.3.53}$$

$$\text{where} \quad \psi^{(2)}_k(w) = \tilde{f}(w, y_k)\prod_{l=1}^{k-1} s^{\mathrm{III,II}}(w, y_l)$$

Using the compatibility condition

$$S^{\mathrm{II}}_{i,j}|A^{\dagger}_4(w)\rangle^{\mathrm{II}} = s^{\mathrm{II},\mathrm{II}}(y_i, y_j)|A^{\dagger}_4(w)\rangle^{\mathrm{II}}_{(i,j)} \tag{4.3.54}$$

we can solve f and $s^{\mathrm{III},\mathrm{II}}$:

$$\tilde{f}(w, y_k) = \frac{w - \frac{i}{2g}}{w - v_k - \frac{i}{2g}}, \qquad s^{\mathrm{III},\mathrm{II}}(w, y_k) = \frac{w - v_k + \frac{i}{2g}}{w - v_k - \frac{i}{2g}} \tag{4.3.55}$$

For the two-excitation solution, we will use the ansatz

$$|A^{\dagger}_4(w_1)A^{\dagger}_4(w_2)\rangle^{\mathrm{III}} = |w_1, w_2\rangle^{\mathrm{III}} + S^{\mathrm{III}}_{1,2}|w_1, w_2\rangle^{\mathrm{III}} \quad \text{where}$$

$$S^{\mathrm{III}}_{1,2}|w_1, w_2\rangle^{\mathrm{III}} = s^{\mathrm{III},\mathrm{III}}(w_1, w_2)|w_2, w_1\rangle^{\mathrm{III}} \qquad \text{and} \tag{4.3.56}$$

$$|w_1, w_2\rangle^{\mathrm{III}} = \sum_{1 \le l_1 < l_2 \le K_1} \psi^{(2)}_{l_1}(w_1)\psi^{(2)}_{l_2}(w_2)|\ldots A^{\dagger}_3 A^{\dagger}_4(y_{l_1})A^{\dagger}_3 \ldots A^{\dagger}_3 A^{\dagger}_4(y_{l_2})A^{\dagger}_3 \ldots\rangle^{\mathrm{II}}$$

and the (1×1) matrix $S^{\mathrm{III}}_{1,2}$ is of course the third level S-matrix. If we now consider the compatibility condition

$$S^{\mathrm{II}}_{i,j}|A^{\dagger}_4(w_1)A^{\dagger}_4(w_2)\rangle^{\mathrm{III}} = s^{\mathrm{II},\mathrm{II}}(y_i, y_j)|A^{\dagger}_4(w_1)A^{\dagger}_4(w_2)\rangle^{\mathrm{III}}_{(i,j)} \tag{4.3.57}$$

on a spin chain of length two we will get

$$s^{\mathrm{III},\mathrm{III}}(w_1, w_2) = \frac{w_1 - w_2 - \frac{i}{g}}{w_1 - w_2 + \frac{i}{g}} \tag{4.3.58}$$

Analogous to (2.4.110) and (4.3.50) these results can easily be generalized to the K_2-excitation case:

$$|A^{\dagger}_4(w_1)\ldots A^{\dagger}_4(w_{K_2})\rangle^{\mathrm{III}} = |w_1, \ldots, w_{K_2}\rangle^{\mathrm{II}} + \sum_{P \in S_{K_2}} \mathbf{S}^{\mathrm{III}}_P|w_1, \ldots, w_{K_2}\rangle^{\mathrm{II}}$$

$$\text{where} \quad |w_1, \ldots, w_{K_2}\rangle^{\mathrm{II}} \equiv \sum_{1 \le l_1 < \ldots l_{K_2} \le K_1} \prod_{j=1}^{K_2} \psi^{(2)}_{l_j}(w_j)|\ldots A^{\dagger}_4(y_{l_1})\ldots A^{\dagger}_4(y_{l_{K_2}})\ldots\rangle^{\mathrm{II}} \tag{4.3.59}$$

We have now solved (4.3.25); a solution of (4.3.25) can be obtained by writing out a K_2-excitation state $|A^{\dagger}_4(w_1)\ldots A^{\dagger}_4(w_{K_1})\rangle^{\mathrm{III}}$ in terms of level I excitations by using (4.3.59), (4.3.50) and (4.3.28):

$$|A^{\dagger}_4(w_1)\ldots A^{\dagger}_4(w_{K_2})\rangle^{\mathrm{III}} = \left(1 + \sum_{P \in S_{K_2}} \mathbf{S}^{\mathrm{III}}_P\right) \sum_{1 \le n_1 < \ldots n_{K_2} \le K_1} \prod_{j=1}^{K_2} \psi^{(2)}_{n_j}(w_j) \cdot$$

$$\cdot \left(1 + \sum_{P' \in S_{K_1}} \mathbf{S}^{\mathrm{II}}_{P'}\right) \bigg\| \sum_{\{l_i\}} \mathfrak{Q}_{\alpha_1, l_1}(y_1)\ldots \mathfrak{Q}_{\alpha_{K_1}, l_{K_1}}(y_{K_1}) \bigg\| |A^{\dagger}_1(p_1)\ldots A^{\dagger}_1(p_N)\rangle^{\mathrm{I}} \tag{4.3.60}$$

From this equation we can also see that

$$
\begin{aligned}
N &= N(A_1^\dagger) + N(\mathfrak{Q}_3) + N(\mathfrak{Q}_4) \\
K_1 &= N(\mathfrak{Q}_3) + N(\mathfrak{Q}_4) \\
K_2 &= N(\mathfrak{Q}_4)
\end{aligned}
\tag{4.3.61}
$$

where $N(\mathfrak{Q}_\alpha)$ and $N(A_1^\dagger)$ stands for the number of $\mathfrak{Q}_\alpha$ and $A_1^\dagger$ respectively in (4.3.60).

We can now derive the Bethe equations analogous to section 2.4.2. First let us introduce the phase factor $s^{\mathrm{I},0} = e^{-ip}$. Then, by imposing the periodicity condition on the auxiliary spin chains we get:

$$
\prod_{i=1}^{L} s^{\mathrm{I},0}(p_k) \prod_{\substack{l=1 \\ l \neq k}}^{N} s^{\mathrm{I},\mathrm{I}}(p_k, p_l) \prod_{m=1}^{K_1} s^{\mathrm{I},\mathrm{II}}(p_k, y_m) = 1 \quad \text{for all} \quad k = 1, \ldots, N
$$

$$
\prod_{i=1}^{N} s^{\mathrm{II},\mathrm{I}}(y_l, p_i) \prod_{\substack{j=1 \\ j \neq l}}^{K_1} s^{\mathrm{II},\mathrm{II}}(y_l, y_j) \prod_{k=1}^{K_2} s^{\mathrm{II},\mathrm{III}}(y_l, w_k) = 1 \quad \text{for all} \quad l = 1, \ldots, K_1
\tag{4.3.62}
$$

$$
\prod_{i=1}^{K_1} s^{\mathrm{III},\mathrm{II}}(w_j, y_i) \prod_{\substack{l=1 \\ l \neq k}}^{K_2} s^{\mathrm{III},\mathrm{III}}(w_j, w_l) = 1 \quad \text{for all} \quad j = 1, \ldots, K_2
$$

For example, the second Bethe ansatz equation can be found by considering $|\alpha_1, y_1; \ldots; \alpha_{K_1}, y_{K_1}\rangle^{\mathrm{II}}$ (4.3.50). When we shift an excitation ($\alpha = 3, y_l$) around this spin chain it can scatter with an excitation of the same type, with an excitation of type $\alpha = 4$ or it can scatter with a lattice site of the level II spin chain (i.e. $A_1^\dagger$). In each of these cases we will get the phase factors $s^{\mathrm{II},\mathrm{I}}(y_l, p_i)$, $s^{\mathrm{II},\mathrm{II}}(y_l, y_j)$ or $s^{\mathrm{II},\mathrm{III}}(y_l, w_k)$ respectively. The Periodicity condition then translates into the second Bethe equation. The first and third Bethe equation are obtained with a similar reasoning.

If we write out (4.3.62) explicitly we get

$$
e^{-ip_k L} \prod_{\substack{l=1 \\ l \neq k}}^{N} S_0(p_k, p_l) \frac{x_k^+ - x_l^-}{x_k^- - x_l^+} \frac{e^{\frac{i}{2} p_l}}{e^{\frac{i}{2} p_k}} \prod_{m=1}^{K_1} e^{\frac{i}{2} p_k} \frac{y_m - x_k^-}{y_m - x_k^+} = 1 \quad \text{for all} \quad k = 1, \ldots, N
$$

$$
\prod_{i=1}^{N} e^{-\frac{i}{2} p_i} \frac{y_l - x_i^+}{y_l - x_i^-} \prod_{k=1}^{K_2} \frac{w_k - v_l - \frac{i}{2g}}{w_k - v_l + \frac{i}{2g}} = 1 \quad \text{for all} \quad l = 1, \ldots, K_1
\tag{4.3.63}
$$

$$
\prod_{i=1}^{K_1} \frac{w_j - v_i + \frac{i}{2g}}{w_j - v_i - \frac{i}{2g}} \prod_{\substack{l=1 \\ l \neq k}}^{K_2} \frac{w_j - w_l - \frac{i}{g}}{w_j - w_l + \frac{i}{g}} = 1 \quad \text{for all} \quad j = 1, \ldots, K_2
$$

and this coincides exactly with (4.3.21) and (4.3.23) as expected. Finally, to obtain physical relevant solutions, we will impose the condition

$$
P = \sum_{i=1}^{N} p_i = 0 \quad \text{or in other words} \quad \prod_{i=1}^{N} \frac{x_i^+}{x_i^-} = 1
\tag{4.3.64}
$$

4.4   Representation Theory of the Bethe ansatz vectors

Now that we have derived the Bethe vectors, we will describe some of its weight properties. First, let us remind ourselves of the discussion in section 2.6.3, where it was shown that the Bethe vectors of the Hubbard model are highest weight vectors of the $\mathfrak{su}(2) \oplus \mathfrak{su}(2) \subseteq \mathfrak{su}(2|2)$ subalgebra. Since the matrix $\bar{R}_{1,2}$ is related to Shastry's R-matrix $\bar{R}_{1,2}^{(s)}$, it is expected that the Bethe vectors we found in the previous section will have some kind of highest weight property with respect to the two $\mathfrak{su}(2)$ copies which are generated by $\{M_1^+, F_1^+, F_1^-\}$ and $\{M_3^+, F_3^+, F_3^-\}$ respectively. This will become more clearly when we consider a Bethe vector constructed from the NABA.

First we note that the whole NABA construction of section 4.3.1 is performed after a basis transformation (4.3.1). Therefore, let us define

$$
\tilde{F}_1^+ = G(p)\, F_1^+\, G^{-1}(p) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and}
$$

$$
\tilde{F}_3^- = G(p)\, F_3^-\, G^{-1}(p) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

(4.4.1)

We immediate see from (2.6.52) that $\tilde{F}_1^+ = -\Sigma^+$ and $\tilde{F}_3^- = -S^+$. Therefore, the Bethe vectors are highest weight vectors with respect to the algebra generated by $\{M_1^+, F_1^+, F_1^-\}$ and they are lowest weight vectors with respect to the algebra generated by $\{M_3^+, F_3^+, F_3^-\}$. In the same way, we see that

$$
G(p)\, M_1^+\, G^{-1}(p) = -\Sigma^z \quad \text{and} \quad G(p)\, M_1^+\, G^{-1}(p) = -S^z \tag{4.4.2}
$$

which means that our Bethe vector has $(\mathfrak{su}(2) \oplus \mathfrak{su}(2))$ weight $[K_1 - N, 2K_2 - K_1]$.

     To calculate the weight with respect to $\mathfrak{su}(2|2)$, it will be convenient to consider the Bethe vector (4.3.60). If we use (4.1.1) and (4.1.8) we get:

$$
\begin{aligned}
M_1 |A_1^\dagger(p)\rangle &= -|A_1^\dagger(p)\rangle, & [M_1, \mathbf{Q}_\alpha{}^1] &= \mathbf{Q}_\alpha{}^1 \\
M_2 |A_1^\dagger(p)\rangle &= 0, & [M_2, \mathbf{Q}_\alpha{}^1] &= -\delta_{3\,\alpha} \mathbf{Q}_3{}^1 \\[4pt]
M_3 |A_1^\dagger(p)\rangle &= 0, & [M_3, \mathbf{Q}_\alpha{}^1] &= \begin{cases} \mathbf{Q}_3{}^1 & \text{if } \alpha = 3 \\ -\mathbf{Q}_4{}^1 & \text{if } \alpha = 4 \end{cases}
\end{aligned}
\tag{4.4.3}
$$

Then, using (4.3.61), we can easily see that (4.3.60) has $\mathfrak{su}(2|2)$ weight $[K_1 - N, K_2 - K_1, K_1 - 2K_2]$. At this point you may wonder why the third component of the weight is $K_1 - 2K_2$ instead of $2K_2 - K_1$. This can be explained by the nesting procedure of the NCBA. There we see that the third level pseudo-vacuum is given by $|0\rangle^{\text{III}} = |A_3^\dagger(y_1) \dots A_3^\dagger(y_{K_1})\rangle^{\text{II}}$ where $A_3^\dagger$ is viewed as an excitation of this vacuum. However, we could perfectly define this pseudo-vacuum as $|0\rangle^{\text{III}} = |A_4^\dagger(y_1) \dots A_4^\dagger(y_{K_1})\rangle^{\text{II}}$ and view $A_3^\dagger$ as an excitation of this vacuum. In that

case the Bethe vector would have weight $[K_1 - N, K_2 - K_1, 2K_2 - K_1]$ as expected.

We have now calculated the weights of the Bethe vectors and shown that they have highest/lowest weight properties with respect to $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$.

4.5   The Complete Bethe Ansatz Equations

As we have remarked earlier, the $\mathfrak{su}(2|2)_{C,C^\dagger} \oplus \mathfrak{su}(2|2)_{C,C^\dagger}$-invariant S-matrix $S^*_{1,2}$ is given by (4.2.18). So let us define

$$\tilde{S}^*_{1,2}(p_1, p_2) = \tilde{S}_{1,2}(p_1, p_2) \otimes \tilde{S}_{1,2}(p_1, p_2) \tag{4.5.1}$$

Then, the Bethe equations of the light-cone gauge string theory will follow from solving the auxiliary eigenvalue problem

$$\Lambda^*(\lambda, \{p_i\})|\Psi\rangle = \tau^*(p_k, \{p_i\}_{i=1}^N)|\Psi\rangle, \qquad \text{where}$$
$$\tau^*(p_k, \{p_i\}_{i=1}^N) \equiv \text{str}_a[\tilde{S}^*_{a,N}(\lambda, p_N)\, \tilde{S}^*_{a,N-1}(\lambda, p_{N-1}) \dots \tilde{S}^*_{a,1}(\lambda, p_1)] \tag{4.5.2}$$

and imposing the eigenvalue condition

$$\Lambda^*(\lambda, \{p_i\}) = \left(\frac{x_k^+}{x_k^-}\right)^L \tag{4.5.3}$$

analogous to (4.3.8) and (4.3.9). But because of (4.5.1), it follows directly that

$$\Lambda^*(\lambda, \{p_i\}) = \Lambda(\lambda; \{p_i\}_{i=1}^N, \{\lambda_j\}_{j=1}^{K_1^{(1)}}, \{\mu_k\}_{k=1}^{K_2^{(1)}})\, \Lambda(\lambda; \{p_i\}_{i=1}^N, \{\lambda_j\}_{j=1}^{K_1^{(2)}}, \{\mu_k\}_{k=1}^{K_2^{(2)}})$$
$$\text{where} \quad 0 \le K_1^{(\alpha)} \le K_2^{(\alpha)} \le N \le L \quad \text{for} \quad \alpha = 1, 2 \tag{4.5.4}$$

Therefore, the complete Bethe ansatz equations are

$$e^{iP} \prod_{i=1}^N \left(S_0(p_k, p_i)\frac{x_i^- - x_k^+}{x_i^+ - x_k^-}\right)^2 \cdot \prod_{\alpha=1}^2 \prod_{j=1}^{K_1} \frac{x_k^- - y_j^{(\alpha)}}{x_k^+ - y_j^{(\alpha)}} = e^{ip_k\left(L+N-\frac{K_1^1}{2}-\frac{K_1^2}{2}\right)} \tag{4.5.5}$$

$$\prod_{i=1}^N e^{-\frac{i}{2}p_i}\frac{y_l^{(\alpha)} - x_i^+}{y_l^{(\alpha)} - x_i^-} \prod_{k=1}^{K_2} \frac{w_k^{(\alpha)} - v_l^{(\alpha)} - \frac{i}{2g}}{w_k^{(\alpha)} - v_l^{(\alpha)} + \frac{i}{2g}} = 1 \quad \text{for all} \quad l = 1, \dots, K_1 \quad \text{and} \quad \alpha = 1, 2$$

$$\prod_{i=1}^{K_1} \frac{w_j^{(\alpha)} - v_i^{(\alpha)} + \frac{i}{2g}}{w_j^{(\alpha)} - v_i^{(\alpha)} - \frac{i}{2g}} \prod_{\substack{l=1 \\ l \neq k}}^{K_2} \frac{w_j^{(\alpha)} - w_l^{(\alpha)} - \frac{i}{g}}{w_j^{(\alpha)} - w_l^{(\alpha)} + \frac{i}{g}} = 1 \quad \text{for all} \quad j = 1, \dots, K_2 \quad \text{and} \quad \alpha = 1, 2$$

We have now derived the Bethe equations for the full string sigma model in the light-cone gauge.

Chapter 5

Summary and Conclusions

In this thesis we have reviewed various Bethe ansatz techniques. We started with the spin-1/2 XXX Heisenberg model and solved it with the coordinate Bethe ansatz. Next we presented the algebraic Bethe ansatz using, again, the XXX model and derived the integrability of the model and demonstrated the highest weight property of the Bethe vectors. We then introduced the Hubbard model and solved it using two characterizations of the nested coordinate Bethe ansatz. We also presented the nested algebraic Bethe ansatz for graded models and solved the Hubbard model with it. Finally, we ended the review with a short exposition on the S-matrix and its connection with the Bethe ansatz and integrability. The main idea is that the S-matrix of an integrable model is restricted by the symmetries of the model and that, depending on the symmetry algebra, these restrictions are sometimes so strong that they determine the S-matrix uniquely up to a constant.

Next we introduced basic Lie superalgebra theory and we described the $\mathfrak{psu}(2,2|4)$ and the (extended) $\mathfrak{su}(2|2)$ Lie superalgebra. The (extended) $\mathfrak{su}(2|2)$-invariant S-matrix is then derived and the corresponding Bethe ansatz equations are calculated. Finally, we generalized these results to obtain Bethe equations for the string sigma model in the light-cone gauge.

Appendix A

Tensors, The Yang Baxter Equation and Graded Linear Algebra

A.1 Tensor Notations

Let V and W be finite dimensional vector spaces. Then we can form the tensor product $V \otimes W$, which is a vector space of dimension $\dim(V) \cdot \dim(W)$. Now let $A : V \to V'$ and $B : W \to W'$ be linear maps. Then we can define a linear map $A \otimes B : V \otimes W \to V' \otimes W'$ by

$$(A \otimes B)(v \otimes w) = A(v) \otimes B(w) \qquad \text{for all} \quad v \in V, w \in W \tag{A.1.1}$$

If we additionally have linear maps $C : \bar{V} \to V$ and $D : \bar{W} \to W$, then

$$(A \otimes B)(C \otimes D) = AC \otimes BD \qquad \text{with} \quad AC \otimes BD : \bar{V} \otimes \bar{W} \to V' \otimes W' \tag{A.1.2}$$

In this thesis, we will often use an explicit matrix representation of the tensor products. The conventions used for this matrix representation are addressed below.

First, let V_a be a n-dimensional vector space with basis $\{f_i\}$ and V_i be a m-dimensional vector space with basis $\{g_i\}$. For $End(V_a)$ and $End(V_i)$ we will use the standard basis $\{f_i^j | 1 \leq i, j \leq n\}$ and $\{g_i^j | 1 \leq i, j \leq m\}$ respectively, which is characterized by the properties

$$f_i^j f_k = f_i \delta_{jk}, \qquad f_i^j f_k^l = f_i^l \delta_{jk}, \qquad g_i^j g_k = g_i \delta_{jk}, \qquad g_i^j g_k^l = g_i^l \delta_{jk} \tag{A.1.3}$$

Then, the space $V_a \otimes V_i$ is a nm-dimensional vector space for which we will use the basis $\{e_i^j = f_i \otimes g_j | 1 \leq i \leq n$ and $1 \leq j \leq m\}$, and the standard basis for $End(V_a \otimes V_i)$ is $\{e_{ik}^{jl} = f_i^j \otimes g_k^l | 1 \leq i, j \leq n$ and $1 \leq k, l \leq m\}$ which has the properties

$$
\begin{aligned}
e_{ik}^{jl} e_p^q &= (f_i^j \otimes g_k^l)(f_p \otimes g_q) = f_i^j f_p \otimes g_k^l g_q = \delta_{jp} \delta_{lq} \left(f_i \otimes g_k \right) = \delta_{jp} \delta_{lq} e_i^k \\
e_{ik}^{jl} e_{pr}^{qs} &= (f_i^j \otimes g_k^l)(f_p^q \otimes g_r^s) = f_i^j f_p^q \otimes g_k^l g_r^s = \delta_{jp} \delta_{lr} \left(f_i^q \otimes g_k^s \right) = \delta_{jp} \delta_{lr} e_{ik}^{qs}
\end{aligned}
\tag{A.1.4}
$$

Now let $v \in V_a$ and $w \in V_i$, then we denote the e_{ij} component of the vector $(v \otimes w)$ with $(v \otimes w)_{i,j}$. Analogously, if we have mappings $A \in End(V_a)$ and $B \in End(V_i)$, then the e_{ik}^{jl} component of matrix $S_{a,i} \equiv A \otimes B : V_a \otimes V_i \rightarrow V_a \otimes V_i$ is denoted by

$$(S_{a,i})_{i,k}^{j,l} = (A \otimes B)_{i,k}^{j,l} = A_i^j B_k^l \quad \text{where} \quad A = \sum_{i,j} A_i^j f_i^j, \quad B = \sum_{i,j} B_i^j g_i^j \tag{A.1.5}$$

In general, we can view $S_{a,i}$ as a $(nm \times nm)$ matrix. However, sometimes we will refer to V_a as the auxiliary space. In that case $S_{a,i}$ will be viewed as a $n \times n$ matrix were each matrix element is a $m \times m$ matrix.

Now suppose that we have another matrix $R_{a,j} \equiv A' \otimes B' : V_a \otimes V_j \rightarrow V_a \otimes V_j$, where V_j is a r-dimensional vector space and $A' \in End(V_a)$, $B' \in End(V_j)$. Then we will define the operator $S_{a,i} R_{a,j}$ by

$$S_{a,i} R_{a,j} \equiv (AA') \otimes B \otimes B' : V_a \otimes V_i \otimes V_j \rightarrow V_a \otimes V_i \otimes V_j \tag{A.1.6}$$

where AA' is the usual matrix multiplication, and $S_{a,i} R_{a,j}$ can be viewed as a $(nmr \times nmr)$ matrix. If we use V_a as an auxiliary space, then $S_{a,i} R_{a,j}$ can be viewed as a $m \times m$ matrix were each matrix element is a $(mr \times mr)$ matrix acting on the space $V_i \otimes V_j$.

To illustrate how the matrices look like, let us consider the following example. Suppose that V_a is a three-dimensional vector space with standard basis $\{f_1 = (1,0,0), f_2 = (0,1,0), f_3 = (0,0,1)\}$ and that V_i is a two-dimensional vector space with standard basis $\{g_1 = (1,0,), g_2 = (0,1)\}$. Then we will use the convention that the basis of $V_a \otimes V_i$ is represented as

$$\begin{aligned} e_1^1 &= (1,0,0,0,0,0), & e_1^2 &= (0,1,0,0,0,0), & e_2^1 &= (0,0,1,0,0,0) \\ e_2^2 &= (0,0,0,1,0,0), & e_3^1 &= (0,0,0,0,1,0), & e_3^2 &= (0,0,0,0,0,1) \end{aligned} \tag{A.1.7}$$

Now suppose that we have matrices $A \in End(V_a)$ and $B \in End(V_i)$. Then we can view $S_{a,i} = A \otimes B$ as an 6×6 matrix with the following form:

$$S_{a,i} = \begin{pmatrix} S_{1,1}^{1,1} & S_{1,1}^{1,2} & S_{1,1}^{2,1} & S_{1,1}^{2,2} & S_{1,1}^{3,1} & S_{1,1}^{3,2} \\ S_{1,2}^{1,1} & S_{1,2}^{1,2} & S_{1,2}^{2,1} & S_{1,2}^{2,2} & S_{1,2}^{3,1} & S_{1,2}^{3,2} \\ S_{2,1}^{1,1} & S_{2,1}^{1,2} & S_{2,1}^{2,1} & S_{2,1}^{2,2} & S_{2,1}^{3,1} & S_{2,1}^{3,2} \\ S_{2,2}^{1,1} & S_{2,2}^{1,2} & S_{2,2}^{2,1} & S_{2,2}^{2,2} & S_{2,2}^{3,1} & S_{2,2}^{3,2} \\ S_{3,1}^{1,1} & S_{3,1}^{1,2} & S_{3,1}^{2,1} & S_{3,1}^{2,2} & S_{3,1}^{3,1} & S_{3,1}^{3,2} \\ S_{3,2}^{1,1} & S_{3,2}^{1,2} & S_{3,2}^{2,1} & S_{3,2}^{2,2} & S_{3,2}^{3,1} & S_{3,2}^{3,2} \end{pmatrix} \tag{A.1.8}$$

If we view V_a as an auxiliary space, then $S_{a,i}$ will be viewed as a 3×3 matrix were each matrix element is a 2×2 matrix:

$$S_{a,i} = \begin{pmatrix} A_1^1 B & A_1^2 B & A_1^3 B \\ A_2^1 B & A_2^2 B & A_2^3 B \\ A_3^1 B & A_3^2 B & A_3^3 B \end{pmatrix} \tag{A.1.9}$$

Now suppose that we have another matrix $R_{a,j} = A' \otimes B' : V_a \otimes V_j \to V_a \otimes V_j$, with V_j being a r-dimensional vector space, $A' \in End(V_a)$ and $B' \in End(V_j)$. Then using V_a as an auxiliary space, $S_{a,i}R_{a,j}$ will be of the form

$$S_{a,i}R_{a,j} = \begin{pmatrix} (AA')_1^1 B \otimes B' & (AA')_1^2 B \otimes B' & (AA')_1^3 B \otimes B' \\ (AA')_2^1 B \otimes B' & (AA')_2^2 B \otimes B' & (AA')_2^3 B \otimes B' \\ (AA')_3^1 B \otimes B' & (AA')_3^2 B \otimes B' & (AA')_3^3 B \otimes B' \end{pmatrix} \quad \text{with } (BB')_i^j = \sum_{k=1}^{2} B_i^k B'^{j}_{k}$$

(A.1.10)

An often used operator on tensor spaces is the permutation operator. We will derive two simple but useful identities for the permutation operator which will be used occasionally in this thesis. First, let V be a n-dimensional vector space with basis $\{f_i\}$. The permutation operator P is then defined as

$$P : V \otimes V \to V \otimes V : v_1 \otimes v_2 \to v_2 \otimes v_1 \quad \text{where} \quad P = \sum_{i,j=1}^{n} f_i^j \otimes f_j^i \qquad \text{(A.1.11)}$$

Now suppose that we have a linear mapping $S \equiv A \otimes B : V \otimes V \to V \otimes V$ defined by

$$S_{a,c}^{b,d} = (A \otimes B)_{a,c}^{b,d} = A_a^b B_c^d \quad \text{where} \quad A = \sum_{i,j} A_i^j f_i^j, \quad B = \sum_{i,j} B_i^j f_i^j \qquad \text{(A.1.12)}$$

then it is easy to see that we have the identities

$$(PS)_{a,c}^{b,d} e_{a,c}^{b,d} = \sum_{i,j} (f_a^c \otimes f_c^a)(A_i^b f_i^b \otimes B_j^d f_j^d) = A_c^b f_a^b \otimes A_a^d f_c^d = S_{c,a}^{b,d} e_{a,c}^{b,d}$$

$$(SP)_{a,c}^{b,d} e_{a,c}^{b,d} = \sum_{i,j} (A_a^i f_a^i \otimes B_c^j f_c^j)(f_d^b \otimes f_b^d) = A_a^d f_a^b \otimes A_c^b f_c^d = S_{a,c}^{d,b} e_{a,c}^{b,d}$$

(A.1.13)

We will end this section by introducing the partial trace. Suppose that we have finite dimensional vector spaces V_a, V_i and linear maps $A \in End(V_a)$ and $B \in End(V_i)$. Then the partial trace $tr_a \equiv tr_{V_a}$ over the space V_a of the linear map $A \otimes B$ is defined as

$$tr_a(A \otimes B) = tr(A) \otimes B \qquad \text{(A.1.14)}$$

where tr is the ordinary trace.

A.2   The Yang Baxter Equation

Let $\mathcal{A}$ be an algebra. Then an operator $\mathcal{R} : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}$ is called an universal R-matrix if it satisfies the abstract Yang-Baxter relation

$$\mathcal{R}_{1,2}\mathcal{R}_{1,3}\mathcal{R}_{2,3} = \mathcal{R}_{2,3}\mathcal{R}_{1,3}\mathcal{R}_{1,2} \qquad \text{(A.2.1)}$$

which is an equation in $\mathcal{A} \otimes \mathcal{A} \otimes \mathcal{A}$. Here, the operator $\mathcal{R}_{i,j}$ is the canonically induced operator on $\mathcal{A} \otimes \mathcal{A} \otimes \mathcal{A}$; for example, $\mathcal{R}_{1,2} = \mathcal{R} \otimes \mathbb{I}$ with $\mathbb{I}$ being the identity operator.

Now let $\{\rho(a, \lambda)\}$ be a family of representations parameterized by a discrete label a and continuous parameter λ and let us define

$$
\begin{aligned}
R_{1,3}(\lambda, \zeta) &= \left[\rho(a_1, \lambda) \otimes \rho(a_3, \zeta)\right] \mathcal{R} \\
R_{2,3}(\mu, \zeta) &= \left[\rho(a_2, \mu) \otimes \rho(a_3, \zeta)\right] \mathcal{R} \\
R_{1,2}(\lambda, \mu) &= \left[\rho(a_1, \lambda) \otimes \rho(a_2, \mu)\right] \mathcal{R}
\end{aligned}
\tag{A.2.2}
$$

Then the abstract Yang-Baxter relation becomes the ordinary Yang-Baxter equation (YBE)

$$
R_{1,2}(\lambda, \mu) R_{1,3}(\lambda, \zeta) R_{2,3}(\mu, \zeta) = R_{2,3}(\mu, \zeta) R_{1,3}(\lambda, \zeta) R_{1,2}(\lambda, \mu)
\tag{A.2.3}
$$

in $V_1 \otimes V_2 \otimes V_3$ where V_i is the representation space of $\rho(a_i, \alpha)$. The matrix $R_{i,j}(\alpha, \beta)$ in the YBE acts on $V_1 \otimes V_2 \otimes V_3$ using the canonically induced action; it acts trivially on the third space V_k, $k \neq i, j$. We will call a solution of the YBE a R-matrix.

If we take $\mu = \zeta = 0$ and assume that the family of representations $\{\rho(a, \lambda)\}$ has some homogeneity such that

$$
[\rho(a_1, \lambda) \otimes \rho(a_2, \mu)]\mathcal{R} = R_{1,2}(\lambda - \mu)
\tag{A.2.4}
$$

then (A.2.3) simplifies to the form

$$
R_{1,2}(\lambda - \mu) R_{1,3}(\lambda) R_{2,3}(\nu) = R_{2,3}(\nu) R_{1,3}(\lambda) R_{1,2}(\lambda - \mu)
\tag{A.2.5}
$$

which is more widely used in the literature. In this case the variable α occurring in the argument of the R-matrix is called the spectral parameter.

A.3 An Alternate Form of the FCR

Let us consider the fundamental commutation relation (2.3.5):

$$
R_{a,b}(\lambda - \mu)[T_a(\lambda) T_b(\mu)] = [T_b(\mu) T_a(\lambda)] R_{a,b}(\lambda - \mu)
\tag{A.3.1}
$$

Then $T_a(\lambda)$ is an (induced) operator on $V_a \otimes V_b \otimes V$ where the action on V_b is trivial (and $V \equiv V_1 \otimes \ldots \otimes V_n$). Therefore

$$
T_a(\lambda) T_b(\mu) = [T(\lambda) \otimes \mathbb{I}_b] \cdot [\mathbb{I}_a \otimes T(\mu)] = T(\lambda) \otimes T(\mu)
\tag{A.3.2}
$$

With some abuse of notation, we will write $T(\lambda) \otimes T(\mu)$ as $T_a(\lambda) \otimes T_b(\mu)$ where it is understood that $T_a(\lambda)$ is an operator on $V_a \otimes V$. Furthermore, note that

$$
(T(\lambda) \otimes \mathbb{I}_b) \cdot (\mathbb{I}_a \otimes T(\mu)) \neq (\mathbb{I}_a \otimes T(\mu)) \cdot (T(\lambda) \otimes \mathbb{I}_b)
\tag{A.3.3}
$$

since $T_a(\lambda)$ and $T_b(\mu)$ act on V in a noncommutative way. Now let P be the permutation operator on $V_a \otimes V_b$. Then it is easy to see that

$$P\left[T_b(\mu)T_a(\lambda)\right]P = T_a(\mu)T_b(\lambda) = T_a(\mu) \otimes T_b(\lambda) \tag{A.3.4}$$

So if we define $\bar{R}_{a,b} = PR_{a,b}$, then

$$\bar{R}_{a,b}\left[T_a(\lambda) \otimes T_b(\mu)\right] = \left[T_a(\mu) \otimes T_b(\lambda)\right]\bar{R}_{a,b} \tag{A.3.5}$$

This is a form of the FCR that will also be used in this thesis. Finally, we note that $\bar{R}$ satisfies the relation

$$\bar{R}_{2,3}(\lambda,\mu)\bar{R}_{1,2}(\lambda,\nu)\bar{R}_{2,3}(\mu,\nu) = \bar{R}_{1,2}(\mu,\nu)\bar{R}_{2,3}(\lambda,\nu)\bar{R}_{1,2}(\lambda,\mu) \tag{A.3.6}$$

In some literature, this equation is sometimes referred to as being the Yang-Baxter equation instead of (A.2.3). Therefore, with some abuse of definition, we will also call $\bar{R}$ a R-matrix.

A.4 Graded Linear Algebra

In this section we shall introduce the basic concepts of graded vector spaces and graded linear mappings. More specifically, we will restrict our discussion to $\mathbb{Z}_2$-graded linear algebra although the whole construction can easily be generalized to $\mathbb{Z}_n$, $n \in \mathbb{Z}$.

Let V be a vector space over the field k. Then V is called $\mathbb{Z}_2$ graded if V can be written as $V = V_{\bar{0}} \oplus V_{\bar{1}}$. In this case we sometimes write $V \equiv V^{(m|n)}$ where m and n denote the dimensions of $V_{\bar{0}}$ and $V_{\bar{1}}$ respectively.

An element $v \in V$ is called homogeneous of degree $i \in \mathbb{Z}_2$, if $v \in V_i$. When v is a homogeneous vector, then i is also called the parity of v, and we will write $p(v) = i$. Furthermore, v will be called even or odd when $v \in V_{\bar{0}}$ or $v \in V_{\bar{1}}$ respectively.

Now let W be a second $\mathbb{Z}_2$ graded vector space. A linear mapping $A : V \to W$ is called homogenous of degree $j \in \mathbb{Z}_2$ if

$$A(V_i) \subseteq W_{i+j} \quad \text{for all} \quad i \in \mathbb{Z}_2 \tag{A.4.1}$$

Furthermore, the notions of parity, even and odd apply to A as well.

Given $\mathbb{Z}_2$ graded vector spaces V and W, we can form the tensor product between V and W which we will denote with $V \bar{\otimes} W$ to underline the fact that we are dealing with graded spaces. The tensor product $V \bar{\otimes} W$ has a natural $\mathbb{Z}_2$ gradation defined by

$$(V \bar{\otimes} W)_i = \bigoplus_{k+l=i} (V_k \bar{\otimes} W_l) \quad \text{where} \quad i, k, l \in \mathbb{Z}_2 \tag{A.4.2}$$

Now suppose that V' and W' are $\mathbb{Z}_2$ graded vector spaces, and let $A : V \to V'$ and $B : W \to W'$ be two linear mappings which are homogeneous of degrees i and j respectively. Then we can define a linear mapping $A \bar{\otimes} B : V \bar{\otimes} W \to V' \bar{\otimes} W'$ by

$$(A \bar{\otimes} B)(v \bar{\otimes} w) = (-1)^{jk} A(v) \bar{\otimes} B(w) \quad \text{for all} \quad v \in V_k, w \in W \tag{A.4.3}$$

Note that this implies that if we additionally have linear maps $C : \bar{V} \to V$ and $D : \bar{W} \to W$ which are homogeneous of degrees r and s respectively, then

$$(A \,\bar{\otimes}\, B)(C \,\bar{\otimes}\, D) = (-1)^{jr} AC \,\bar{\otimes}\, BD \qquad \text{with} \quad AC \,\bar{\otimes}\, BD : \bar{V} \,\bar{\otimes}\, \bar{W} \to V' \,\bar{\otimes}\, W' \qquad (A.4.4)$$

Furthermore, we see that all these definitions reduce to the ordinary tensor definitions of section A.1 when $\dim V_{\bar{1}} = \dim W_{\bar{1}} = 0$.

As mentioned before, we will often use the explicit matrix representation of the tensor product. Therefore, let us fix a basis for the graded vector spaces. So suppose that $V^{(m|n)} = V_{\bar{0}} \oplus V_{\bar{1}}$ is a graded vector space with basis $\{f_1, \ldots, f_m\}$ and $\{f_{m+1}, \ldots, f_{m+n}\}$ for $V_{\bar{0}}$ and $V_{\bar{1}}$ respectively, and let $W^{(r|s)}$ be a second graded vector space with basis $\{g_1, \ldots, g_r, g_{r+1}, \ldots, g_{r+s}\}$ defined in the same manner. For $End(V^{(m|n)})$ we will use the standard basis $\{f_i^j | 1 \leq i, j \leq m+n\}$, which is characterized by (A.1.3), and the parity of f_i^j is given by $p(f_i^j) = p(f_i) + p(f_j)$. To lighten up the notation, we will write $p(k)$ for $p(f_k)$ when there is no chance for confusion. For $End(W^{(r|s)})$ we will denote the standard basis by $\{g_i^j | 1 \leq i, j \leq r+s\}$ and it has similar properties as $\{f_i^j\}$. Then, as in section A.1, the space $V \,\bar{\otimes}\, W$ is a $(n+m)(r+s)$-dimensional vector space for which we will use the basis $\{e_i^j = f_i \,\bar{\otimes}\, g_j | 1 \leq i \leq n+m \text{ and } 1 \leq j \leq r+s\}$, and the standard basis for $End(V \,\bar{\otimes}\, W)$ is $\{e_{ik}^{jl} = f_i^j \,\bar{\otimes}\, g_k^l | 1 \leq i, j \leq n+m \text{ and } 1 \leq k, l \leq r+s\}$. The parity of e_i^j is $p(i) + p(j)$ and the parity of e_{ik}^{jl} is $p(i) + p(j) + p(k) + p(l)$. Furthermore, we have the properties

$$
\begin{aligned}
e_{ik}^{jl}(e_a^b) &= (f_i^j \,\bar{\otimes}\, g_k^l)(f_a \,\bar{\otimes}\, g_b) = (-1)^{[p(k)+p(l)]p(a)} f_i^j f_a \,\bar{\otimes}\, g_k^l g_b \\
&= (-1)^{[p(k)+p(l)]p(a)} \delta_{ja}\, \delta_{lb} \left(f_i \,\bar{\otimes}\, g_k \right) = (-1)^{[p(k)+p(l)]p(a)} \delta_{ja}\, \delta_{lb}\, e_i^k \\
e_{ik}^{jl} e_{ac}^{bd} &= (f_i^j \,\bar{\otimes}\, g_k^l)(f_a^b \,\bar{\otimes}\, g_c^d) = (-1)^{[p(k)+p(l)][p(a)+p(b)]}(f_i^j f_a^b \,\bar{\otimes}\, g_k^l g_c^d) \\
&= (-1)^{[p(k)+p(l)][p(a)+p(b)]} \delta_{ja}\, \delta_{lc}\, (f_i^b \,\bar{\otimes}\, g_k^d) = (-1)^{[p(k)+p(l)][p(a)+p(b)]} \delta_{ja}\, \delta_{lc}\, e_{ik}^{bd}
\end{aligned}
\qquad (A.4.5)
$$

Because of all the (-1) factors, it is hard to get an explicit matrix realization with these definitions. Therefore, let us introduce the so called "graded tensor product" $\hat{\otimes}$ which is defined as

$$f_i \,\hat{\otimes}\, g_j \equiv (-1)^{p(i)p(j)}(f_i \,\bar{\otimes}\, g_j), \qquad (f_i^j \,\hat{\otimes}\, g_k^l) \equiv (-1)^{[p(i)+p(j)]p(k)}(f_i^j \,\bar{\otimes}\, g_k^l) \qquad (A.4.6)$$

With these definitions we see that

$$(f_i^j \,\hat{\otimes}\, g_k^l)(f_a^b \,\hat{\otimes}\, g_c^d) = f_i^j f_a^b \,\hat{\otimes}\, g_k^l g_c^d \quad \text{and} \quad (f_i^j \,\hat{\otimes}\, g_k^l)(f_a \,\hat{\otimes}\, g_b) = f_i^j f_a \,\hat{\otimes}\, g_k^l g_b \qquad (A.4.7)$$

Therefore, by using the basis $\{\hat{e}_i^j \equiv f_i \,\hat{\otimes}\, g_j\}$ and $\{\hat{e}_{ik}^{jl} = f_i^j \,\hat{\otimes}\, g_k^l\}$ we can get an explicit matrix realization following the conventions of section A.1.

Note that the graded tensor product is extended to multiple vector spaces in the following way. Suppose that $\{V_i\}_{i=1}^n$ are a collection of graded vector spaces with basis $\left\{e_{m_i}^{(i)}\right\}$ for V_i, and suppose that $\left\{e_{k_i}^{l_i\,(i)}\right\}$ form a basis of $End(V_i)$. Then

$$e_{m_1}^{(1)} \,\hat{\otimes}\, e_{m_2}^{(2)} \ldots \hat{\otimes} \ldots \hat{\otimes}\, e_{m_n}^{(n)} = (-1)^{\sum_{i=1}^n p(m_i) \sum_{j=i+1}^n p(m_j)} e_{m_1}^{(1)} \,\bar{\otimes}\, e_{m_2}^{(2)} \ldots \bar{\otimes} \ldots \bar{\otimes}\, e_{m_n}^{(n)} \qquad (A.4.8)$$

$$e_{k_1}^{l_1\,(1)} \,\hat{\otimes}\, e_{k_2}^{l_2\,(2)} \ldots \hat{\otimes} \ldots \hat{\otimes}\, e_{k_n}^{l_n\,(n)} = (-1)^{\sum_{i=1}^n (p(k_i)+p(l_i)) \sum_{j=i+1}^n p(k_j)} e_{k_1}^{l_1\,(1)} \,\bar{\otimes}\, e_{k_2}^{l_2\,(2)} \ldots \bar{\otimes} \ldots \bar{\otimes}\, e_{k_n}^{l_n\,(n)}$$

Now that we have treated the basic concepts of graded vector spaces, we will introduce the notions of "supertrace" and "partial supertrace". Suppose that V is a finite dimensional graded vector space and $A \in End(V)$ is a linear map written as $A = \sum_{i,j} A_i^j f_i^j$ in basis $\{f_i^j\}$. Then the supertrace of A is defined as

$$\text{str}(A) = \sum_i (-1)^{p(f_i^j)} A_i^i \tag{A.4.9}$$

Now suppose that we have finite dimensional graded vector spaces V_a, V_i and linear maps $A \in End(V_a)$ and $B \in End(V_i)$. Then the partial supertrace $\text{str}_a \equiv \text{str}_{V_a}$ over the space V_a of the linear map $A \hat{\otimes} B$ is defined as

$$\text{str}_a(A \hat{\otimes} B) = \text{str}(A) \hat{\otimes} B \tag{A.4.10}$$

We will end this section by making two (notational) remarks. First of all, when X and Y are graded objects (e.g. vector spaces or linear mappings), then the tensor product $X \otimes Y$ is to be understood as the ordinary tensor product between X and Y where the grading is ignored. So $X \otimes Y$ is viewed as a non graded object and the conventions of section A.1 are of use.

Secondly, let us make a remark on induced actions of graded operators. Suppose that V, W are graded vector spaces and that $A \in End(W)$. Then because of the occurrence of minus signs of the tensor product action we have to be careful when we talk about the induced action of A on $V \bar{\otimes} W$. This is because we will sometimes interpret the induced action as $\mathbb{I}_V \hat{\otimes} B$ and sometimes as $\mathbb{I}_V \bar{\otimes} B$. To avoid confusion we will always state in the text which induced action we use.

Appendix B

Lie Algebras

B.1 Definitions and Notations

In this section we will recall some definitions and properties of Lie algebras. It is by no means intended as a short introduction into Lie algebra theory, but merely as a short reminder for the reader who has already encountered the subject before. The material in this section is largely based on [19].

Let $\mathfrak{g}$ be a semi-simple Lie algebra over $\mathbb{C}$, $\mathfrak{k}$ a compact real form of $\mathfrak{g}$, $\mathfrak{t}$ a maximal commutative subalgebra of $\mathfrak{k}$ and $\mathfrak{h} \equiv \mathfrak{t} + i\mathfrak{t}$ a Cartan subalgebra of $\mathfrak{g}$. A nonzero linear functional α on $\mathfrak{h}$ is called a root of $\mathfrak{g}$ (relative to the Cartan subalgebra $\mathfrak{h}$), if there exists a nonzero element X of $\mathfrak{g}$ such that

$$[H, X] = \alpha(H)X \qquad \text{for all } H \in \mathfrak{h} \tag{B.1.1}$$

The space of all X in $\mathfrak{g}$ for which this relation holds is called the root space $\mathfrak{g}_\alpha$ and an element of $\mathfrak{g}_\alpha$ is called a root vector (for the root α). The set of all roots is denoted by $R = R(\mathfrak{g}, \mathfrak{t})$ and it has the property that R spans $\mathfrak{h}^*$ and that $R \subseteq i\mathfrak{t}^*$. Furthermore, if α is a root, then $-\alpha$ is also a root and we can find nonzero elements $X_\alpha \in \mathfrak{g}_\alpha$, $Y_\alpha \in \mathfrak{g}_{-\alpha}$ and $H_\alpha \in \mathfrak{h}$ such that

$$[H_\alpha, X_\alpha] = 2X_\alpha, \qquad [H_\alpha, Y_\alpha] = -2Y_\alpha, \qquad [X_\alpha, Y_\alpha] = H_\alpha \tag{B.1.2}$$

The element H_α is unique and is called the co-root of α. It satisfies the relation $\alpha(H_\alpha) = 2$.

Now let us fix an inner product $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ on $\mathfrak{g}$ (note that each inner product is a multiple of the Killing form) and consider its restriction to $\mathfrak{h}$, which is an inner product on $\mathfrak{h}$. Then, for each $\alpha \in \mathfrak{h}^*$ (not necessarily a root) there exists a unique element $H^\alpha \in \mathfrak{h}$ such that $\alpha(H) = \langle H^\alpha, H \rangle$ for all H in $\mathfrak{h}$. Therefore, we can identify the element $\alpha \in \mathfrak{h}^*$ with $H^\alpha \in \mathfrak{h}$ which enables us to define an inner product on $\mathfrak{h}^*$, which we again denote with $\langle \cdot, \cdot \rangle$, by $\langle \alpha, \beta \rangle \equiv \langle H^\alpha, H^\beta \rangle = \alpha(H^\beta)$. From this we see that

$$H_\alpha = 2\frac{H^\alpha}{\langle \alpha, \alpha \rangle} = 2\frac{H^\alpha}{\langle H^\alpha, H^\alpha \rangle} \qquad \text{and} \qquad H^\alpha = 2\frac{H_\alpha}{\langle H_\alpha, H_\alpha \rangle} \tag{B.1.3}$$

Now let (E, R) be a root system. Then a simple root system for R is a subset $P = \{\alpha_1, \dots, \alpha_r\}$ such that P forms a basis for E as a vector space and such that for each $\alpha \in R$ we have

$$\alpha = n_1\alpha_1 + n_2\alpha_2 + \dots + n_r\alpha_r \qquad \text{where} \quad n_i \geq 0 \; \forall i \quad \text{or} \quad n_i < 0 \; \forall i \qquad \text{(B.1.4)}$$

Once a base P is chosen, the set of α's for which $n_i \geq 0$ are called positive roots (with respect to P) and the the set of α's for which $n_i < 0$ are called negative roots. The elements of P are called (positive) simple roots. For each simple root α_i there exists a unique element $\omega_i \in \mathfrak{h}^*$ such that $\omega_i(H_{\alpha_j}) = \delta_{ij}$ for all j, where H_{α_j} is the co-root of α_j. We will call this element a fundamental weight and the set of fundamental weights form a basis of $\mathfrak{h}^*$. Note that a general weight $\mu \in \mathfrak{h}^*$ can be written as

$$\mu = \sum_{i=1}^{r} \mu_i \omega_i \quad \text{where} \quad \mu_i = \mu(H_{\alpha_i}) = 2\frac{\langle \mu, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \qquad \text{(B.1.5)}$$

which implies that for all $\alpha_j \in P$ we have

$$\alpha_j = \sum_{i=1}^{r} n_{ij} \omega_i \quad \text{where} \quad n_{ij} = \alpha_j(H_{\alpha_i}) = 2\frac{\langle \alpha_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \qquad \text{(B.1.6)}$$

The $r \times r$ matrix n with entries n_{ij} is called the Cartan matrix associated with P. Finally, if μ is the highest weight of a highest weight representation, then the numbers μ_i are called Dynkin coefficients or Dynkin labels. Note that the Cartan matrix and the Dynkin coefficients are independent of the choice of the inner product.

We will end this section by mentioning some uniqueness results regarding the choices we made for the Lie algebra $\mathfrak{g}$ of Lie group G:

- If $\mathfrak{k}_1$ and $\mathfrak{k}_2$ are two compact real forms of $\mathfrak{g}$, then there exists an element A of G such that $Ad_A(\mathfrak{k}_1) = \mathfrak{k}_2$ where Ad is the adjoint representation of G in $\mathfrak{g}$.

- Suppose that $\mathfrak{k}$ is a compact real form of $\mathfrak{g}$ and that K is the compact subgroup of G whose Lie algebra is $\mathfrak{k}$. If $\mathfrak{t}_1$ and $\mathfrak{t}_2$ are two maximal commutative subalgebras of $\mathfrak{g}$, then there exists an element A of K such that $Ad_A(\mathfrak{t}_1) = \mathfrak{t}_2$.

- If $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two Cartan subalgebras of $\mathfrak{g}$, then there exists an element A of G such that $Ad_A(\mathfrak{h}_1) = \mathfrak{h}_2$

- Any two bases P_1 and P_2 of a root system (E, R) can be mapped into one another by the action of the Weyl group.

Bibliography

[1] O. Aharony, S. S. Gruber, J. M. Maldacena, H. Ooguri, and Y. Oz, *Large N Field Theories, String Theory and Gravity*, `hep-th/9905111`.

[2] G. Arutyunov, S. Frolov, J. Plefka, and M. Zamaklar, *The Off-shell Symmetry Algebra of the Light-cone $AdS_5 \times S_5$ Superstring*, `hep-th/0609157`.

[3] G. Arutyunov, S. Frolov, and M. Zamaklar, *The Zamolodchikov-Faddeev Algebra for $AdS_5 \times S_5$ Superstring*, `hep-th/0612229`.

[4] N. Beisert, *The Dilatation Operator of $\mathcal{N}$=4 Super Yang-Mills Theory and Integrability*.

[5] ———, *The $\mathfrak{su}(2|2)$ Dynamic S-Matrix*, `hep-th/0511082`.

[6] N. Beisert and M. Staudacher, *The $\mathcal{N}$=4 SYM Integrable Super Spin Chain*, `hep-th/0307042`.

[7] H. Bethe, *Zur Theorie der Metalle. Eigenwerte und Eigenfunktionen der linearen Atomkette*, Z. Phys. **71** (1931), 205–226.

[8] A. R. Chowdhury and A. G. Choudhury, *Quantum Integrable Systems*, Chapman & Hall/CRC, 2003.

[9] M. de Leeuw, *Coordinate Bethe Ansatz for the String S-Matrix*, `hep-th/07052369`.

[10] V. G. Drinfeld, *Quantum Groups*, Proc. Int. Cong. Mathematicians **1** (1986), 798.

[11] F. H. L. Essler and V. E. Korepin, *Exact Solution of an Electronic Model of Superconductivity in 1+1 Dimensions I*, `cond-mat/9211001`.

[12] Fabian H. L. Essler, Holger Frahm, Frank Gohmann, Andreas Klumper, and Vladimir E. Korepin, *The One-Dimensional Hubbard Model*, Cambridge University Press, 2005.

[13] L. D. Faddeev, *Algebraic Aspects of Bethe Ansatz*, `hep-th/9404013`.

[14] ———, *How Algebraic Bethe Ansatz works for integrable model*, `hep-th/9605187`.

[15] L. D. Faddeev and L. A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons*, Springer Series in Soviet Mathematics, Springer-Verlag, 1987.

[16] M. K. Fung, *Validity of the Bethe-Yang Hypothesis in the Delta Function Interaction*, J. Math. Phys. **22** (1981), 2017.

[17] C. S. Gardner, J. M. Green, M. D. Kruskal, and R. M. Miura, *Method for Solving the Korteweg-deVries Equation*, Phys. Rev. Lett. **19** (1967), 1095.

[18] Zachary N. C. Ha, *Quantum Many-Body Systems in One Dimension*, Series on Advances in Statistical Mechanics, vol. 12, World Scientific, 1996.

[19] Brian C. Hall, *Lie Groups, Lie Algebras and Representations*, Springer-Verlag, 2003.

[20] M. Karbach and G. Muller, *Introduction to the Bethe Ansatz I*, `cond-mat/9809162`.

[21] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, 1993.

[22] M. J. Martins and P. B. Ramos, *The quantum inverse scattering method for Hubbard-like models*, Nuclear Physics B **522** (1998), 413–470.

[23] D. C. Mattis, *The Many-Body Problem, An Encyclopedia of Exactly Solved Models in One Dimension*, World Scientific, 1993.

[24] J. A. Minahan and K. Zarembo, *The Bethe Ansatz for N=4 Super Yang Mills*.

[25] B. Sutherland, *Further Results for the Many-Body Problem in One Dimension*, Phys. Rev. Lett. **20** (1968), no. 3, 98.

[26] ———, *An Introduction to the Bethe Ansatz*, Exactly Solvable Problems in Condensed Matter and Relativistic Field Theory (B. S. Shastry, S. S. Jha, and V. Singh, eds.), Springer-Verlag, 1986, Lecture Notes in Physics 242, pp. 1–95.

[27] C. N. Yang, *Some Exact Results for the Many-Body Problem in One Dimension with Repulsive Delta-function Interaction*, Phys. Rev. Lett. **19** (1967), no. 23, 1312.

[28] Alexander B. Zamolodchikov and Alexey B. Zamolodchikov, *Factorized S-Matrices in Two Dimensions as the Exact Solutions of Certain Relativistic Quantum Field Theory Models*, Annals of Physics **120** (1979), 253–291.