The Algebraic Bethe Ansatz and Tensor Networks

V. Murg,[1] V. E. Korepin,[2] and F. Verstraete[1]

[1] *Vienna Center for Quantum Science and Technology, Faculty of Physics, University of Vienna, Vienna, Austria*

[2] *C. N. Yang Institute for Theoretical Physics, State University of New York at Stony Brook, NY 11794-3840, USA*

(Dated: January 27, 2012)

We describe the Algebraic Bethe Ansatz for the spin 1/2 XXX and XXZ Heisenberg chains with open and periodic boundary conditions in terms of tensor networks. These Bethe eigenstates have the structure of Matrix Product States with a conserved number of down-spins. The tensor network formulation suggestes possible extensions of the Algebraic Bethe Ansatz to two dimensions.

I. INTRODUCTION

The coordinate Bethe ansatz[1] is an extremely successful method for solving one-dimensional problems exactly. It reduces the complex problem of diagonalizing the Hamiltonian to finding the solutions of a set of algebraic equations. Once solutions to these algebraic equations are found – numerical approaches to find them efficiently exist in many cases – the eigenvalues are known exactly. However, the eigenstates are available only as a complex mathematical expressions the structure of which is not evident. This makes it insuperable, in general, to get interesting properties out of the states – like their entanglement characteristics or their correlations. The algebraic Bethe ansatz[2] reveals more about the structure of the eigenstate and offers new perspectives to obtain scalar products[3], norms[2] and correlations[2].

In this paper, we point out this structure by formulating the algebraic Bethe ansatz in the pictoresque tensor network language. In addition to making the ansatz more vivid, the tensor network formulation might bear the potential of extending the ansatz to higher dimensions.

The description of states in terms of tensor networks has been very successful in the recent past. The one-dimensional matrix product states (MPS)[4,5] form the basis for the extremely successful density matrix normalization group (DMRG)[6,7]. Also, they have attracted considerable interest in the interdisciplinary field of quantum information and condensed matter physics[8–11]. For describing the ground state of systems on higher-dimensional lattices, the projected entangled pair states (PEPS)[12] were introduced and proved to be useful for the numerical study of ground states of two-dimensional systems[13,14]. The Multiscale Entanglement Renormalization Ansatz (MERA)[15,16] allows the description and numerical study of critical systems.

From the tensor network desription of the Bethe eigenstates it is immediately obvious that eigenstates can be described as MPS: see also Katsura and Maruyama [17]. Katsura and Maruyama also show that the alternative formulation of the Bethe Ansatz by Alcaraz and Lazo [18–20] is equivalent to the algebraic Bethe ansatz.

In Sec. II, we describe the tensor network form of the Bethe eigenstates and the structure of the obtained MPS. In Sec. III, we formulate the algebraic Bethe ansatz in the tensor network language. In Sec.IV, we give a pic-
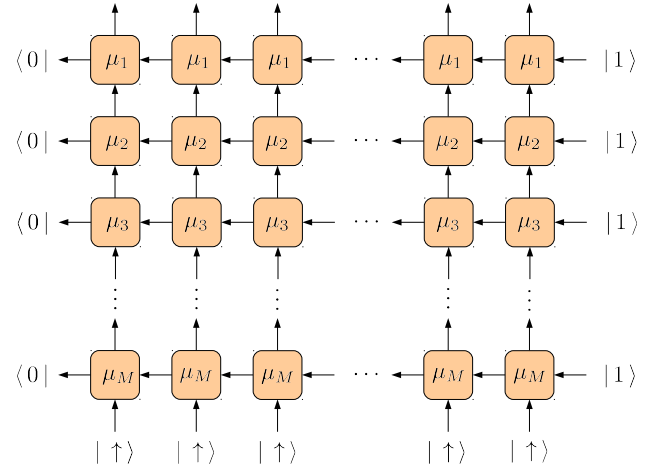

FIG. 1. Tensor network constituting the Bethe eigenstate of the Heisenberg model or XXZ model with periodic boundary conditions. $S_z = \frac{1}{2}N - M$

toresque description of the algebraic Bethe ansatz with open boundary conditions in terms of tensor networks.

II. MATRIX PRODUCT STATE FORM OF BETHE SOLUTIONS

Typically, Bethe-eigenstates are obtained as products of operators $B(\mu_j)$ applied on a certain vacuum state $|vac\rangle$, i.e.

$$|\Psi(\mu_1, \ldots, \mu_M)\rangle = B(\mu_1) \cdots B(\mu_M)|vac\rangle. \quad (1)$$

The parameters $\{\mu_j\}$ are thereby solutions of Bethe equations and the $B(\mu_j)$'s play the role of creation operators. In case of the antiferromagnetic Heisenberg model and the XXZ model with periodic boundary conditions, the vacuum corresponds to the state with all spins up and each operator $B(\mu_j)$ creates one down-spin. Thus, the product of M such operators applied to the vacuum creates a state with M down-spins, i.e. magnetization $S_z = N/2 - M$ (with N being the number of spins). $B(\lambda)$ is an operator acting on the whole Hilbert-space of dimension 2^N, but it has the well-structured form of a

Matrix Product Operator (MPO)[21] with virtual dimension 2. Indeed, as will be shown in Sec. III,

$$B(\lambda) = \sum_{\substack{k_1 \cdots k_N \\ l_1 \cdots l_N}} \langle 0 | \mathcal{L}_{l_1}^{k_1}(\lambda) \cdots \mathcal{L}_{l_N}^{k_N}(\lambda) | 1 \rangle o_{l_1}^{k_1} \otimes \cdots \otimes o_{l_N}^{k_N}$$

with $k, l \in \{0, 1\}$, $o_l^k = | k \rangle\langle l |$ ($0 \equiv \uparrow$, $1 \equiv \downarrow$) and $\mathcal{L}_l^k(\lambda)$ being 2×2 matrices dependent on the parameter λ. The product of operators $B(\mu_1) \cdots B(\mu_M)$ can be read as the contraction of the set of 4-index tensors $[\mathcal{L}_l^k(\mu_j)]_{r'}^r$ with respect to a rectangular grid, as shown in Fig. 1. Thereby, r, r', k and l label the left, right, up and down-indices, respectively. Explicitly, the matrices $\mathcal{L}_l^k(\lambda)$ read

$$\mathcal{L}_0^0(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & c(\lambda) \end{pmatrix}, \quad \mathcal{L}_1^0(\lambda) = \begin{pmatrix} 0 & 0 \\ b(\lambda) & 0 \end{pmatrix}$$

$$\mathcal{L}_0^1(\lambda) = \begin{pmatrix} 0 & b(\lambda) \\ 0 & 0 \end{pmatrix}, \quad \mathcal{L}_1^1(\lambda) = \begin{pmatrix} c(\lambda) & 0 \\ 0 & 1 \end{pmatrix}.$$

In case of the Heisenberg model $H_{XXX} = \sum_{j=1}^N h_{XXX}^{(j,j+1)}$ with

$$h_{XXX} = \frac{1}{2} \left[\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z - \mathbb{1} \right],$$

the functions $b(\lambda)$ and $c(\lambda)$ are

$$b(\lambda) = \frac{1}{1 + \lambda}, \quad c(\lambda) = \frac{\lambda}{1 + \lambda}.$$

For the XXZ model $H_{XXZ}(\Delta) = \sum_{j=1}^N h_{XXZ}^{(j,j+1)}(\Delta)$ with

$$h_{XXZ}(\Delta) = \frac{1}{2} \left[\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \Delta \left(\sigma_z \otimes \sigma_z - \mathbb{1} \right) \right],$$

the functions read

$$b(\lambda) = \frac{\sinh(2i\eta)}{\sinh(\lambda + 2i\eta)}, \quad c(\lambda) = \frac{\sinh(\lambda)}{\sinh(\lambda + 2i\eta)}.$$

The parameter η is related to the inhomogenity Δ in the XXZ model via $\Delta = \cos(2\eta)$.

Because of its "creation operator"-property, there is an inherent structure in the MPO $B(\mu)$: each summand in the MPO $B(\mu)$ must be non-zero only if $k_1 + \cdots + k_N = l_1 + \cdots + l_N + 1$. This global constraint can be reduced to the local constraint that the tensors $[\mathcal{L}_l^k(\mu)]_{r'}^r$ must be non-zero only if $r' = r + (k - l)$. This allows to interprete the virtual indices as "creation-annihilation" counters: the right index r' is equal to the left index r if the physical state is unchanged, it is increased if a down-spin is created and is decreased if a down-spin is annihilated. Thus, the virtual indices transfer the information on how many down-spins are created and anniliated from left to right. Since the left boundary-state is $\langle 0 |$ and the right boundary-state is $| 1 \rangle$, it is guaranteed that the whole MPO creates exactly one down-spin. With the restriction $r, r' \in \{0, 1\}$ there are 6 possible configurations that fulfill the local constraint. In other words, only 6 entries

m	D	
1	2	$2 = 1 \oplus 1$
2	4	$2 \otimes 2 = 1 \oplus 2 \oplus 1$
3	8	$2 \otimes 2 \otimes 2 = 1 \oplus 4 \oplus 2 \oplus 1$
4	16	$2 \otimes 2 \otimes 2 \otimes 2 = 1 \oplus 8 \oplus 4 \oplus 2 \oplus 1$
5	32	$2 \otimes 2 \otimes 2 \otimes 2 \otimes 2 = 1 \oplus 16 \oplus 8 \oplus 4 \oplus 2 \oplus 1$
6	64	$2 \otimes 2 \otimes 2 \otimes 2 \otimes 2 \otimes 2 = 1 \oplus 32 \oplus 16 \oplus 8 \oplus 4 \oplus 2 \oplus 1$

TABLE I. Disintegration of the matrices forming the MPS at step m, $| \Psi_m \rangle$, into blocks. The full size of the matrices is $D \times D$. The MPS has a conserved number of m down-spins.

of the tensor $[\mathcal{L}_l^k(\mu)]_{r'}^r$ are non-zero. These 6 non-zero entries are

$$\begin{array}{cc} [\mathcal{L}_0^0(\mu)]_0^0, & [\mathcal{L}_1^1(\mu)]_1^1 \\ [\mathcal{L}_1^0(\mu)]_0^1, & [\mathcal{L}_0^1(\mu)]_1^0 \\ [\mathcal{L}_0^0(\mu)]_1^1, & [\mathcal{L}_1^1(\mu)]_0^0, \end{array}$$

which is consistent with the matrices written above.

The multiplication of all MPOs with the product state $| vac \rangle$ evidently yields a Matrix Product State (MPS)[8,10] with bond-dimension 2^M. Since each MPO $B(\lambda)$ has the "creation operator"-property to create one down-spin, the MPS contains exactly M down-spins. Explicitly, the MPS reads

$$| \Psi \rangle = \sum_{k_1 \cdots k_N} \langle 0 | \langle 0 | \mathcal{A}^{k_1} \cdots \mathcal{A}^{k_N} | 0 \rangle | M \rangle | k_1, \ldots, k_N \rangle$$

with matrices $\mathcal{A}^k$ being block-diagonal in the sense that $\langle \alpha | \langle s | \mathcal{A}^k | \beta \rangle | s' \rangle \equiv [\mathcal{A}^k]_{\beta s'}^{\alpha s}$. α and β are the virtual indices that range from 0 to $D - 1$ (with D being the virtual dimension of the state). One the other hand, s and s' are the symmetry indices that transfer the information about the number of down-spins from left to right. The local constraint that guarantees this information transfer is $s' = s + k$. This constraint determines the blocks $[\mathcal{A}^k]_{-s'}^{-s}$ that are non-zero and allows a sparse storage of the state. The left boundary-state $\langle 0 |$ and the right boundary-state $| M \rangle$ fix the total number of down-spins of the MPS to M.

The MPS is constructed iteratively by applying the MPOs $B(\mu_1), \ldots, B(\mu_M)$ successively to the vacuum state $| vac \rangle$. The state after m multiplications is evidently a MPS with m down-spins which shall be denoted as

$$| \Psi_m \rangle = \sum_{k_1 \cdots k_N} \langle 0 | \langle 0 | \mathcal{A}_m^{k_1} \cdots \mathcal{A}_m^{k_N} | 0 \rangle | m \rangle | k_1, \ldots, k_N \rangle$$

with $\mathcal{A}_m^k$ being block-diagonal in the sense that $\langle \alpha | \langle s | \mathcal{A}_m^k | \beta \rangle | s' \rangle \equiv [\mathcal{A}_m^k]_{\beta s'}^{\alpha s}$ and fulfilling the constraint $s' = s + k$, as before. The application of the operator $B(\mu)$ to $| \Psi_m \rangle$ yields a state with $m + 1$ down-spins

$$| \Psi_{m+1} \rangle = \sum_{k_1 \cdots k_N} \langle 0 | \langle 0 | \mathcal{A}_{m+1}^{k_1} \cdots \mathcal{A}_{m+1}^{k_N} | 0 \rangle | m+1 \rangle | k_1, .., k_N \rangle.$$
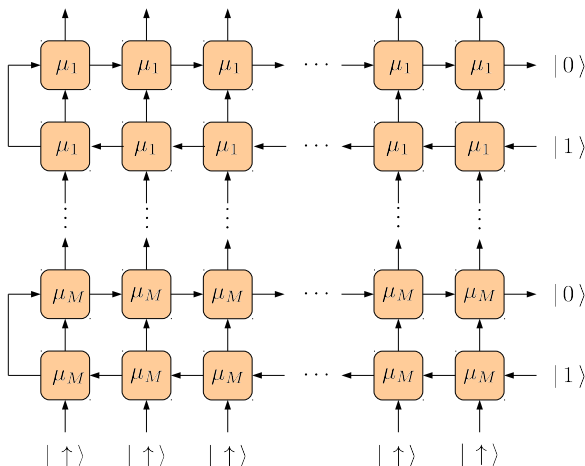
FIG. 2. Tensor network constituting the Bethe eigenstate of the Heisenberg model or XXZ model with open boundary conditions. $S_z = \frac{1}{2}N - M$

The matrices $\mathcal{A}^k_{m+1}$ emerge from tensor-products between $\mathcal{L}^k_l$ and $\mathcal{A}^l_m$, i.e. $\mathcal{A}^k_{m+1} = \sum_l \mathcal{L}^k_l \otimes \mathcal{A}^l_m$. In index notation,

$$\langle \alpha | \langle r | \langle s | \mathcal{A}^k_{m+1} | \beta \rangle | r' \rangle | s' \rangle = \sum_l \langle r | \mathcal{L}^k_l | r' \rangle \langle \alpha | \langle s | \mathcal{A}^l | \beta \rangle | s' \rangle.$$

Because of the constraints $s' = s + l$ and $r' = r + (k - l)$, $S = s + r$ and $S' = s' + r'$ suggest themselves as new symmetry indices. With this definition, $S' = S + k$, as desired. S and S' range from 0 to $m + 1$, since $s \in \{0, \ldots, m\}$ and $r \in \{0, 1\}$. For $S = 0$ and $S = m + 1$, there is the unique choice for $s = r = 0$ and $s = m, r = 1$, respectively. For $0 < S < m + 1$, either $s = S, r = 0$, or $s = S - 1, r = 1$. In this case, the index r must be kept to resolve this ambiguity. The index r can be incorporated into a new virtual index $\tilde{\alpha}$ as $\tilde{\alpha} = (\alpha, r)$. Thus, the dimension of the blocks doubles for $0 < S < m + 1$. The column indices S', s', r' and β can be treated in the same way: for $S' = 0$ and $S' = m + 1$, s' and r' are unambiguously defined; for $0 < S' < m + 1$ there is an ambiguity that can to be resolved by incorporating index r' into a new virtual index $\tilde{\beta} = (\beta, r')$. The matrices $\mathcal{A}^k_{m+1}$ in terms of the virtual indices $\tilde{\alpha}$ and $\tilde{\beta}$ and the symmetry indices S and S', i.e.

$$\langle \tilde{\alpha} | \langle S | \mathcal{A}^k_{m+1} | \tilde{\beta} \rangle | S' \rangle := \langle \alpha | \langle r | \langle s | \mathcal{A}^k_{m+1} | \beta \rangle | r' \rangle | s' \rangle,$$

have the desired block-form that fulfills the constraint $S' = S + k$. Please refer to Table I to see the dimensions of the block-representations that arise for different m's.

In the case of open boundary conditions, the Bethe Ansatz has the same form as in (1), merely the creation Operators are not single MPOs, but products of two MPOs[22,23]:

$$\mathcal{B}(\mu) = \sum_{s=0}^{1} \bar{B}_s(\mu) B_{1-s}(\mu)$$

$B_{1-s}(\mu)$ has the property to create $1 - s$ down-spins, whereas $\bar{B}_s(\mu)$ creates s down-spins ($s \in \{0, 1\}$), such that $\mathcal{B}(\mu)$ is a creation operator for exactly one down-spin, as before. In terms of the previously defined 2×2 matrices $\mathcal{L}^k_l(\mu)$, the MPOs read (see Sec. IV)

$$B_s(\mu) = \sum_{\substack{k_1 \cdots k_N \\ l_1 \cdots l_N}} \langle s | \mathcal{L}^{k_1}_{l_1}(\mu) \cdots \mathcal{L}^{k_N}_{l_N}(\mu) | 1 \rangle o^{k_1}_{l_1} \otimes \cdots \otimes o^{k_N}_{l_N}$$

and

$$\bar{B}_{1-s}(\mu) = \sum_{\substack{k_1 \cdots k_N \\ l_1 \cdots l_N}} \langle s | \mathcal{L}^{k_1}_{l_1}(\mu)^T \cdots \mathcal{L}^{k_N}_{l_N}(\mu)^T | 0 \rangle o^{k_1}_{l_1} \otimes \cdots \otimes o^{k_N}_{l_N}.$$

The virtual indices of $B_s(\mu)$ indicate the balance of created versus annihilated down-spins from left to right. This is due to the local constraint on $[\mathcal{L}^k_l(\mu)]^r_{r'}$ that $r' = r + (k - l)$, as mentioned before. Since the left boundary-vector is $\langle 0 |$ and the right boundary-vector is $| s \rangle$, the creation of s down-spins is guaranteed. In case of $\bar{B}_{1-s}(\mu)$, the MPO is built from the transposed matrices $\mathcal{L}^k_l(\mu)^T$, such that the local constraint on $[\mathcal{L}^k_l(\mu)^T]^r_{r'}$ is $r = r' + (k - 1)$ and the virtual indices count the creation-annihilation balance from right to left. With the right boundary vector $| s \rangle$ and the left boundary vector $\langle 1 |$, one down-spin is created for $s = 0$ and the number of down-spins is kept invariant for $s = 1$.

The tensor-network representation for the Bethe-state with open boundary conditions is shown in Fig. 2. It contains twice as many rows as the tensor-network for periodic boundary conditions, which makes the contraction more challenging, in principle. However, as we see numerically, after a multiplication with a MPO-pair $\mathcal{B}(\mu)$, the Schmidt-rank of the state only increases by a factor of 2 - not 4, as expected. This suggests that there should exist a representation with virtual dimension 2 also in the open boundary conditions-case.

III. THE ALGEBRAIC BETHE ANSATZ

Even though there exist numerous excellent reviews about the Algebraic Bethe Ansatz[2,24–27], we resketch here the Ansatz in the picturesque Tensor-Network language for sake of completeness. In this way, it is traceable, how the tensor networks shown in Figs. 1 and 2 form exact eigenstates of integrable systems.

A. The Yang-Baxter Algebra

In general, the starting point for the Algebraic Bethe Ansatz is the $R(\lambda, \mu)$-tensor

$$R^{\alpha\beta}_{\alpha'\beta'}(\lambda, \mu), \tag{2}$$

with $\alpha, \beta, \alpha', \beta'$ ranging from 1 to some "auxiliary" dimension d and λ, μ being some complex parameters. This
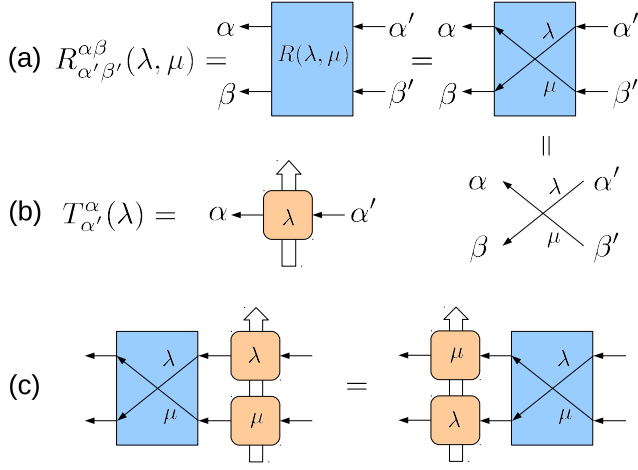
FIG. 3. (a) Visualization of the 4-index R-tensor $R(\lambda, \mu)$. In the abbreviated version, it is visualized as two crossing arrows with λ attached to the up-down arrow and μ attached to the down-up arrow. (b) Yang-Baxter algebra as 4-index tensor with two virtual indices (left-right) and two physical indices (up-down). (c) Defining equation for the Yang-Baxter algebra.
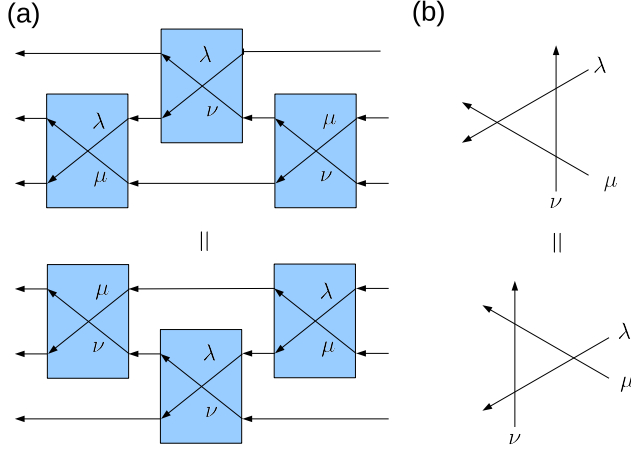

FIG. 4. (a) Yang-Baxter equation in tensor-network form. (b) Abbreviated version.

tensor defines the model under study, as will be shown later. Graphically, the tensor is represented by two crossing arrows, as shown in Fig. 3a, where λ and μ are associated to the up-down and down-up arrows, respectively. After joining indices $(\alpha\beta)$ and $(\alpha'\beta')$, the tensor (2) can also be interpreted as matrix $R(\lambda, \mu)$ acting on the vector space $V \otimes V$ (with $V = \mathbb{C}^d$).

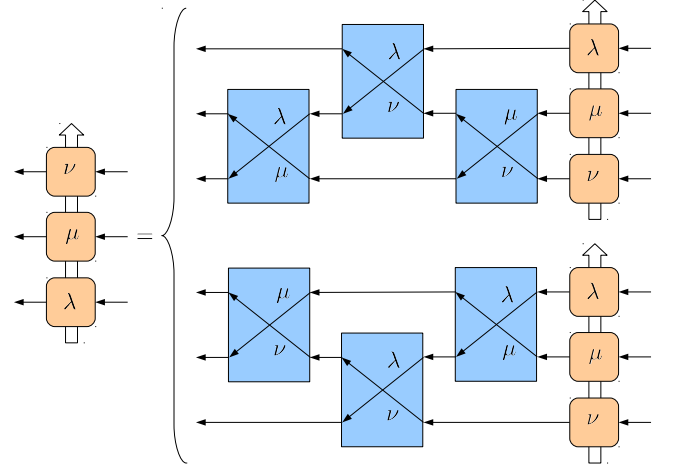The condition on the R-tensor (2) is that it fulfills

FIG. 5. Inversion of the ordering of 3 composed Yang-Baxter algebras using R-tensors. The inversion can be achieved in two ways, which makes necessary that the R-tensors fulfill the Yang-Baxter equation (Fig. 4).

Yang-Baxter equation (star-triangle relation). Writing

$$R^{(23)} = \mathbb{1} \otimes R$$
$$R^{(12)} = R \otimes \mathbb{1},$$

the Yang-Baxter equation reads

$$R^{(23)}(\lambda, \mu) R^{(12)}(\lambda, \nu) R^{(23)}(\mu, \nu)$$
$$= R^{(12)}(\mu, \nu) R^{(23)}(\lambda, \nu) R^{(12)}(\lambda, \mu).$$

The graphical representation of this equation is shown in Fig. 5. Another requirement is that solutions of the Yang-Baxter equation are regular, meaning that there exists a λ_0 and a ν_0, such that

$$R^{\alpha\beta}_{\alpha'\beta'}(\lambda_0, \nu_0) = \delta^\alpha_{\alpha'} \delta^\beta_{\beta'}. \tag{3}$$

The tensor $R(\lambda, \mu)$ defines the Yang-Baxter algebra $T^\alpha_{\alpha'}(\lambda)$ ($\alpha, \alpha' = 1, \ldots, d$) by the relation

$$R^{\alpha\beta}_{\alpha'\beta'}(\lambda, \mu) T^{\alpha'}_{\alpha''}(\lambda) T^{\beta'}_{\beta''}(\mu) = T^\alpha_{\alpha'}(\mu) T^\beta_{\beta'}(\lambda) R^{\alpha'\beta'}_{\alpha''\beta''}(\lambda, \mu)$$

As usual, common indices are summed over. Defining the Monodromy $T(\lambda)$ as the matrix of operators

$$T(\lambda) = \begin{pmatrix} T^1_1(\lambda) & \cdots & T^1_d(\lambda) \\ \vdots & \ddots & \vdots \\ T^d_1(\lambda) & \cdots & T^d_d(\lambda) \end{pmatrix},$$

the definition of the Yang-Baxter algebra can be written as

$$R(\lambda, \mu) \left[T(\lambda) \check{\otimes} T(\mu) \right] = \left[T(\mu) \check{\otimes} T(\lambda) \right] R(\lambda, \mu), \tag{4}$$

where the outer product "$\check{\otimes}$" acts in the space $V \otimes V$ in the sense that $[T(\mu)\check{\otimes}T(\lambda)]^{\alpha\beta}_{\alpha'\beta'} \equiv T^\alpha_{\alpha'}(\mu) T^\beta_{\beta'}(\lambda)$. $T(\lambda)$
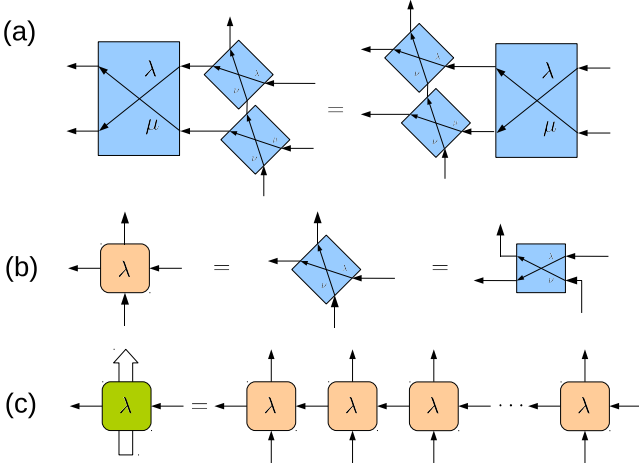
FIG. 6. (a) Deformation of the Yang-Baxter equation, such that it yields the fundamental representation of the Yang-Baxter algebra. (b) Connection between the fundamental representation and the R-tensor. (c) Formation of a more complex representation out of several fundamental representations.
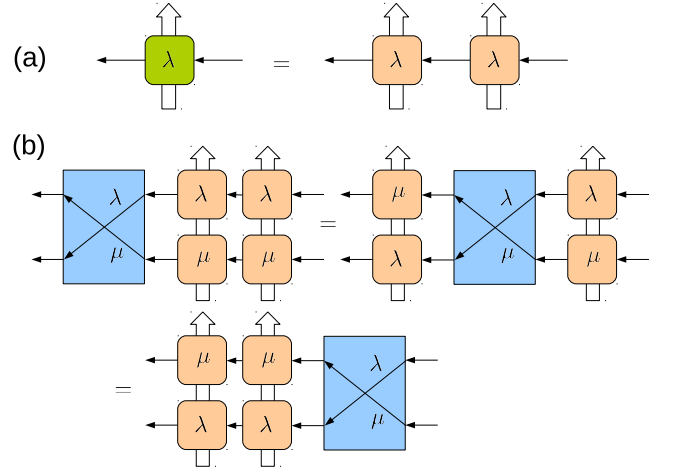

FIG. 7. (a) Co-multiplication property: formation of a new representation of the Yang-Baxter algebra out of two known representations. (b) Proof that the new representation still fulfills the defining equations for the Yang-Baxter algebra.

can be considered as a 4-index tensor: 2 "virtual" indices α, α' of dimension d select the operator $T^\alpha_{\alpha'}(\mu)$ within the matrix, and two "physical" indices operate as input- and output index of the operator. $T(\lambda)$ is represented graphically in Fig. 3b. The virtual indices are indicated as horizontal arrows; the physical input- and output indices are indicated as vertical in- and outgoing double-arrows. Using this graphical notation, the definition of the Yang-Baxter algebra assumes the simple form shown in Fig. 3c.

In this picture, $R(\lambda, \mu)$ has the property to permute the tensors $T(\lambda)$ and $T(\mu)$. There is, however, one ambiguity that arises: there are two ways to go from $T(\lambda)\check\otimes T(\mu)\check\otimes T(\nu)$ to $T(\nu)\check\otimes T(\mu)\check\otimes T(\lambda)$. This inversion of the ordering can be achieved either by exchanging firstly $\lambda \leftrightarrow \mu$, secondly $\lambda \leftrightarrow \nu$ and thirdly $\mu \leftrightarrow \nu$, or by exchanging firstly $\nu \leftrightarrow \mu$, secondly $\lambda \leftrightarrow \nu$ and thirdly $\lambda \leftrightarrow \mu$. This situation is depicted in Fig. 5. Thus, both

$$R^{(12)}(\mu,\nu)R^{(23)}(\lambda,\nu)R^{(12)}(\lambda,\mu)\,[T(\lambda)\check\otimes T(\mu)\check\otimes T(\nu)]$$
$$= \,[T(\lambda)\check\otimes T(\mu)\check\otimes T(\nu)]\,R^{(12)}(\mu,\nu)R^{(23)}(\lambda,\nu)R^{(12)}(\lambda,\mu)$$

and

$$R^{(23)}(\lambda,\mu)R^{(12)}(\lambda,\nu)R^{(23)}(\nu,\mu)\,[T(\lambda)\check\otimes T(\mu)\check\otimes T(\nu)]$$
$$= \,[T(\lambda)\check\otimes T(\mu)\check\otimes T(\nu)]\,R^{(23)}(\lambda,\mu)R^{(12)}(\lambda,\nu)R^{(23)}(\nu,\mu)$$

must be fulfilled. These two equations, however, are compatible, because $R(\lambda, \mu)$ was required to fulfill the Yang-Baxter equation. This makes the definition of the algebra $T^\alpha_{\alpha'}(\lambda)$ consistent.

One representation of the Yang-Baxter algebra is easy to obtain - which is the fundamental representation. This representation is formed by the operators $L^\alpha_{\alpha'}(\lambda, \nu)$ acting

on $\mathbb{C}^d$ defined as

$$[L^\alpha_{\alpha'}(\lambda,\nu)]^k_l = R^{k\alpha}_{\alpha'l}(\lambda,\nu). \qquad (5)$$

In the graphical picture, the operators correspond to a clockwise "rotation" of the R-tensor by 45 degrees, as shown in Fig. 6b. The two indices attached to the horizontal arrow then become the virtual indices of the operator, and the vertical arrow carries the physical indices. That these operators are a valid representation is due to the fact that the defining equation

$$R(\lambda,\mu)\,[L(\lambda,\nu)\check\otimes L(\mu,\nu)] = [L(\mu,\nu)\check\otimes L(\lambda,\nu)]\,R(\lambda,\mu)$$

is just a "distortion" of the Yang-Baxter equation, as shown in Fig. 6a. Up to now, the parameter ν in $L(\lambda, \nu)$ is arbitrary. Most conveniently it is to set $\nu = \nu_0$.

Once one representation $L(\lambda)$ is known, more complex representations are obtained by concatenating the $L(\lambda)$'s horizontally, as depicted in Fig. 6c. Here, operators $T^\alpha_{\alpha'}(\lambda)$ acting on $(\mathbb{C}^d)^{\otimes N}$ are constructed out of N simple operators $L^\alpha_{\alpha'}(\lambda)$ acting on $\mathbb{C}^d$ via

$$T^\alpha_{\alpha'}(\lambda) = \sum_{\alpha_2,\ldots,\alpha_N} L^\alpha_{\alpha_2}(\lambda) \otimes L^{\alpha_2}_{\alpha_3}(\lambda) \otimes \cdots \otimes L^{\alpha_N}_{\alpha'}(\lambda).$$

The outer product "$\otimes$" affects the physical indices. In index notation, the operators read

$$[T^\alpha_{\alpha'}(\lambda)]^{k_1\cdots k_N}_{l_1\cdots l_N} = \sum_{\alpha_2\cdots\alpha_N} [L^\alpha_{\alpha_2}(\lambda)]^{k_1}_{l_1}[L^{\alpha_2}_{\alpha_3}(\lambda)]^{k_2}_{l_2}\cdots [L^{\alpha_N}_{\alpha'}(\lambda)]^{k_N}_{l_N}.$$

The operators defined in such a way fulfill (4), because the R-tensor subsequently interchanges the operators $L^\alpha_{\alpha'}(\lambda)$ from left to right – as can be retraced from Fig. 7b for $N = 2$. Defining the matrices $\mathcal{L}^k_l(\lambda)$ as
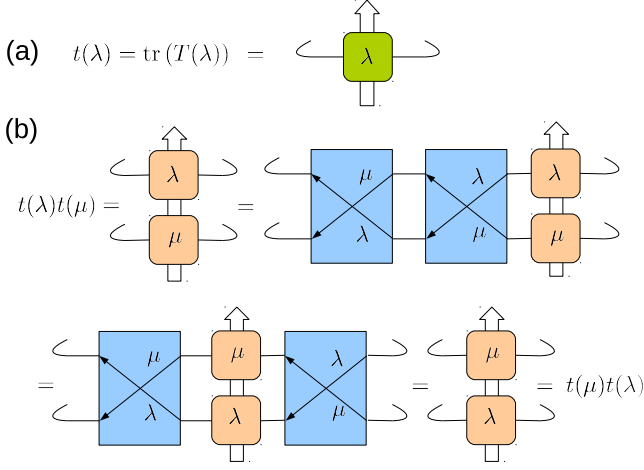
FIG. 8. (a) Definition of the transfer matrix. (b) Proof that all transfer matrices commute.

$\langle\,\alpha\,|\mathcal{L}_l^k(\lambda)|\,\alpha'\,\rangle := [L_{\alpha'}^{\alpha}(\lambda)]_l^k$, the operators $T_{\alpha'}^{\alpha}(\lambda)$ assume the form of MPOs,

$$T_{\alpha'}^{\alpha}(\lambda) = \sum_{\substack{k_1\cdots k_N \\ l_1\cdots l_N}} \langle\,\alpha\,|\mathcal{L}_{l_1}^{k_1}(\lambda)\cdots\mathcal{L}_{l_N}^{k_N}(\lambda)|\,\alpha'\,\rangle o_{l_1}^{k_1}\otimes\cdots\otimes o_{l_N}^{k_N},$$

with $o_l^k = |\,k\,\rangle\langle\,l\,|$.

The main building block of the Algebraic Bethe Ansatz is the transfer matrix $t(\lambda)$, obtained as the trace of the algebra $T_{\alpha'}^{\alpha}(\lambda)$,

$$t(\lambda) := \mathrm{tr}\,\{T(\lambda)\} \equiv \sum_{\alpha} T_{\alpha}^{\alpha}(\lambda).$$

The transfer matrix $t(\lambda)$ corresponds to $T_{\alpha'}^{\alpha}(\lambda)$ with contracted left and right indices α and α' (see Fig. 8a). In the MPO picture, $t(\lambda)$ is represented by an MPO with periodic boundary conditions.[28] Due to equation (4) that is fulfilled by the algebra, the transfer matrix has the property that $[t(\lambda), t(\mu)] = 0$ for all λ and μ. The way this property emerges from (4) can immediately be read off from Fig. 8b: starting out with the expression $t(\lambda)t(\mu)$, the identity in the form $\mathbb{1} = R(\lambda, \mu)^{-1}R(\lambda, \mu)$ can be inserted at the virtual bonds; secondly, $R(\lambda, \mu)$ can be used to exchange $T(\lambda)$ and $T(\mu)$; thirdly, the cyclic property of the trace can be used to eliminate $R(\lambda, \mu)$ and $R(\lambda, \mu)^{-1}$ in order to end up with $t(\mu)t(\lambda)$.

This property makes $t(\lambda)$ the generator of an infinite set of commuting observables: if $t(\lambda)$ is Taylor-expanded with respect to λ, $t(\lambda) = I_0 + \lambda I_1 + \lambda^2 I_2 + \ldots$, then $[I_j, I_k] = 0$ for all j and k. If one of the I_k's is equal to the Hamiltonian of a model, it is called *integrable*, since there exist infinitely many symmetries which commute mutually. In fact, any function of $t(\lambda)$ can be generating function for a set of commuting observables, like e.g. $\mathcal{F}(\lambda) = \log t(\lambda)$. The Taylor-expansion of this function reads

$$\mathcal{F}(\lambda) = \mathcal{F}(\lambda_0) + (\lambda - \lambda_0)\mathcal{F}'(\lambda_0) + O\left((\lambda - \lambda_0)^2\right).$$
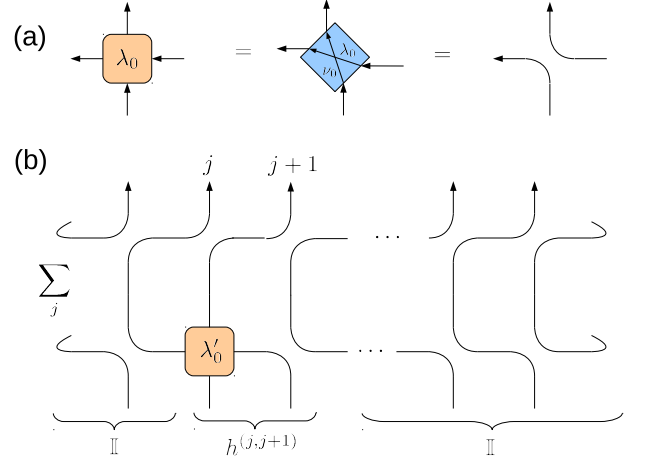

FIG. 9. (a) Precondition on the R-tensor: at some point $\lambda = \lambda_0$, the R-tensor decomposes into the outer product of two identities. (b) Logarithmic derivative of the transfer matrix $t(\lambda)$ at the point $\lambda = \lambda_0$. The first row represents $t(\lambda_0)^{-1}$, the second row $t'(\lambda_0)$.

It turns out that $\mathcal{F}'(\lambda_0)$ is local in the sense that

$$\mathcal{F}'(\lambda_0) \equiv \frac{d}{d\lambda}\log t(\lambda)\Big|_{\lambda=\lambda_0} = \sum_{i=1}^{N} h^{(i,i+1)}$$

with $h^{(i,i+1)}$ only acting on sites i and $i+1$. Thus, an integrable model is obtained described by a local Hamiltonian

$$H = \sum_{i=1}^{N} h^{(i,i+1)}. \tag{6}$$

Thereby,

$$h_{l_1 l_2}^{k_1 k_2} = \frac{d}{d\lambda}[L_{l_2}^{k_1}(\lambda)]_{l_1}^{k_2}\Big|_{\lambda=\lambda_0} \tag{7}$$

or

$$h = \frac{\partial}{\partial\lambda}R(\lambda, \nu_0)\Big|_{\lambda=\lambda_0},$$

respectively. To see this connection, it has to be realized that due to the regularity condition (3) $t(\lambda_0)$ is equal to the cyclic shift operator that shifts the whole lattice to the right by one site. The total momentum operator $\hat{P}$ is related to the cyclic shift operator according to

$$e^{i\hat{P}} = t(\lambda_0). \tag{8}$$

Graphically, $t(\lambda_0)$ is built from $L_{\alpha'}^{\alpha}(\lambda_0)$ shown in Fig. 9a. The way the local Hamiltonian H emerges by differentiating the non-local expression $\mathcal{F}(\lambda)$ is sketched in Fig. 9b. Since $\mathcal{F}'(\lambda_0) = t(\lambda_0)^{-1}t'(\lambda_0)$, the first row in the figure corresponds to the inverted cyclic shift operator $t(\lambda_0)^{-1}$

and the second row corresponds to the derivative $t'(\lambda_0)$. The derivative $t'(\lambda_0)$ disintegrates into a sum of N derivatives with respect to each of the tensors $L(\lambda)$ at sites $j = 1, \ldots, N$. As can be seen in the figure, term j has only support on two sites j and $j+1$ and thus corresponds to a two-site term that is related to the derivative $L'(\lambda_0)$ as formulated in (7).

Models that emerge in such a way from combinations of fundamental representations are fundamental models. Examples are the spin-1/2 Heisenberg model and XXZ model. In both cases, $d = 2$ and the R-matrix assumes the form

$$
R(\lambda, \mu) = \begin{pmatrix} 1 & & & \\ & b(\lambda, \mu) & c(\lambda, \mu) & \\ & c(\lambda, \mu) & b(\lambda, \mu) & \\ & & & 1 \end{pmatrix} \quad (9)
$$

Also, $b(\lambda, \mu)$ and $c(\lambda, \mu)$ are of difference form, i.e. $b(\lambda, \mu) = b(\lambda - \mu)$ and $c(\lambda, \mu) = c(\lambda - \mu)$. This yields a R-matrix of difference form, as well: $R(\lambda, \mu) = R(\lambda - \mu)$. Explicitly, the functions b and c read

$$
b(\lambda) = \frac{1}{1 + \lambda}
$$

$$
c(\lambda) = \frac{\lambda}{1 + \lambda}
$$

for the Heisenberg model and

$$
b(\lambda) = \frac{\sinh(2i\eta)}{\sinh(\lambda + 2i\eta)} \quad (10)
$$

$$
c(\lambda) = \frac{\sinh(\lambda)}{\sinh(\lambda + 2i\eta)} \quad (11)
$$

for the XXZ model. Evidently, in both cases, $R(0) = \mathbb{1}$, such that $\lambda_0 = 0$. In case of the Heisenberg model, $R'(0) = h_{XXX}$ with

$$
h_{XXX} = \frac{1}{2} \left[\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z - \mathbb{1} \right].
$$

In the XXZ-case,

$$
h_{XXZ}(\Delta) = \frac{1}{2} \left[\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \Delta \left(\sigma_z \otimes \sigma_z - \mathbb{1} \right) \right]
$$

is obtained via $R'(0) = 1/\sinh(2i\eta) h_{XXZ}(\Delta)$ with $\Delta = \cos(2\eta)$.

Models (fundamental and non-fundamental) with R-matrix (9) are $gl(2)$ generalized models. The Bethe ansatz for these models is especially simple and will be described in the following.

B.   Bethe Ansatz for $gl(2)$ generalized models

The Yang-Baxter Algebra with R-matrix (9) is generated by only 4 elements, such that the Monodromy assumes the form

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

with

$$
\begin{aligned}
A(\lambda) &= T_0^0(\lambda), & C(\lambda) &= T_0^1(\lambda) \\
B(\lambda) &= T_1^0(\lambda), & D(\lambda) &= T_1^1(\lambda)
\end{aligned} .
$$

The most important commutation relations of the algebra are

$$
B(\lambda)B(\mu) = B(\mu)B(\lambda)
$$

$$
A(\lambda)B(\mu) = \frac{1}{c(\mu, \lambda)} B(\mu)A(\lambda) - \frac{b(\mu, \lambda)}{c(\mu, \lambda)} B(\lambda)A(\mu)
$$

$$
D(\lambda)B(\mu) = \frac{1}{c(\lambda, \mu)} B(\mu)D(\lambda) - \frac{b(\lambda, \mu)}{c(\lambda, \mu)} B(\lambda)D(\mu).
$$

The precondition for the Ansatz is that a representation must exist, for which there is a pseudo-vacuum $|\,vac\,\rangle$ that is an eigenstate of $A(\lambda)$ and $D(\lambda)$ and that is annihiliated by $C(\lambda)$:

$$
A(\lambda)|\,vac\,\rangle = a(\lambda)|\,vac\,\rangle
$$

$$
D(\lambda)|\,vac\,\rangle = d(\lambda)|\,vac\,\rangle
$$

$$
C(\lambda)|\,vac\,\rangle = 0.
$$

The goal is to diagonalize the transfer matrix $t(\lambda) = A(\lambda) + D(\lambda)$. Since all transfer matrices commute, $[t(\lambda), t(\mu)] = 0$, they have a common system of eigenvectors. Thus, all eigenvectors are independent of λ. The eigenvalue problem reads

$$
t(\lambda)|\,\Psi\,\rangle = \tau(\lambda)|\,\Psi\,\rangle.
$$

The Bethe Ansatz

$$
|\,\Psi(\mu_1, \ldots, \mu_M)\,\rangle = B(\mu_1) \cdots B(\mu_M)|\,vac\,\rangle.
$$

fulfills the eigenvalue problem provided that the μ_k's fulfill the Bethe equations

$$
\frac{d(\mu_n)}{a(\mu_n)} = \prod_{\substack{j=1 \\ j \neq n}}^{M} \frac{c(\mu_n, \mu_j)}{c(\mu_j, \mu_n)} \quad (12)
$$

(n=1,…,M). The eigenvalue $\tau(\lambda)$ is then equal to

$$
\tau(\lambda) = a(\lambda) \prod_{j=1}^{M} \frac{1}{c(\mu_j, \lambda)} + d(\lambda) \prod_{j=1}^{M} \frac{1}{c(\lambda, \mu_j)}.
$$

The proof is obtained by utilizing algebraic relations only and can be gathered from appendix A. From $\tau(\lambda)$, the eigenvalue of the Hamiltonian (6) is obtained as

$$
E = \frac{\tau'(\lambda_0)}{\tau(\lambda_0)}. \quad (13)
$$

The total momentum is, according to (8), equal to

$$
p = -i \ln \tau(\lambda_0). \quad (14)
$$

C. Bethe Ansatz for the Heisenberg model and the XXZ model

In case of the Heisenberg model and XXZ model, this spezializes as follows: the matrices $\mathcal{L}_l^k(\lambda)$ that build up the MPOs $T_{\alpha'}^\alpha(\lambda)$ have block form. Written out, they read

$$\mathcal{L}_0^0(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & c(\lambda) \end{pmatrix}, \; \mathcal{L}_1^0(\lambda) = \begin{pmatrix} 0 & 0 \\ b(\lambda) & 0 \end{pmatrix}$$

$$\mathcal{L}_0^1(\lambda) = \begin{pmatrix} 0 & b(\lambda) \\ 0 & 0 \end{pmatrix}, \; \mathcal{L}_1^1(\lambda) = \begin{pmatrix} c(\lambda) & 0 \\ 0 & 1 \end{pmatrix}$$

These MPOs are symmetry conserving in the sense that $T_{\alpha'}^\alpha(\lambda)$ changes the number of down-spins by $\alpha' - \alpha$. This is due to the local constraint that $[\mathcal{L}_l^k(\lambda)]_{\alpha'}^\alpha$ are non-zero only if $\alpha' = \alpha + (k - l)$, as discussed in Sec. II.

Using these considerations, the vacuum state is obviously the state with no down-spins, namely

$$|\, vac\, \rangle = |\, 0\, \rangle \otimes \cdots \otimes |\, 0\, \rangle$$

$(0 \equiv\, \uparrow,\ 1 \equiv\, \downarrow)$. This state is annihilated by $C(\lambda)$ and is an eigenvector of $A(\lambda)$ and $D(\lambda)$ with eigenvalues

$$a(\lambda) = 1, \qquad d(\lambda) = c(\lambda)^N.$$

The Bethe-Ansatz state

$$|\, \Psi(\mu_1, \ldots, \mu_M)\, \rangle = B(\mu_1) \cdots B(\mu_M) |\, vac\, \rangle$$

is a state with M down-spins, i.e. with magnetization in z-direction equal to $S_z = \frac{1}{2}N - M$.

The Bethe equations obtained by the Algebraic Bethe Ansatz are equal to the equations obtained by the Coordinate Bethe Ansatz. In case of Heisenberg model, it is advantageous to introduce variables z_j that are related to μ_j in (12) via

$$\mu_j = \frac{z_j}{2i} - \frac{1}{2}$$

for a direct comparison with results of coordinate Bethe ansatz[1,29–32]: In terms of these variables, the Bethe equations read

$$\left(\frac{z_n - i}{z_n + i}\right)^N = \prod_{\substack{j=1 \\ j \neq n}}^M \frac{z_n - z_j - 2i}{z_n - z_j + 2i} \qquad (15)$$

with $n = 1, \ldots, M$. From the Bethe solutions $\{z_j\}$, the energy is obtained using (13) as

$$E = \frac{\tau'(0)}{\tau(0)} = -\sum_{j=1}^M \frac{4}{z_j^2 + 1}.$$

According to (14), the total momentum $\hat{P}$ has eigenvalue

$$p = -i \ln \tau(0) = \sum_{j=1}^M \left(-i \ln \frac{z_j + i}{z_j - i}\right).$$

The addends are usually referred to as magnon momenta that can be written as

$$p_j = \pi - 2 \arctan(z_j).$$

using the identity

$$\arctan(z) = \frac{1}{2i} \ln \frac{1 + iz}{1 - iz}.$$

In term of the magnon momenta p_j, the total momentum reads

$$p = \sum_{j=1}^M p_j \qquad (16)$$

and the energy is equal to

$$E = -2 \sum_{j=1}^M (1 - \cos(p_j)).$$

For solving the Bethe equations (15) it is advantageous to bring them to their their logarithmic form

$$N p_n = 2\pi I_n + \sum_{\substack{j=1 \\ j \neq n}}^M \Theta(p_n, p_j),$$

where

$$2 \cot \frac{\Theta(p, q)}{2} = \cot \frac{p}{2} - \cot \frac{q}{2}.$$

and I_j are integers $\in \{0, \ldots, N\}$. Solutions can then be found iteratively, as described in [31]. The ground state configuration for N even and $M = N/2$ is $(I_1, \ldots, I_M) = (1, 3, \ldots, N - 1)$.

In case of XXZ model, it is advantageous to introduce the variables z_j related to μ_j in (12) via

$$\mu_j = z_j - i\eta + i\frac{\pi}{2}$$

to compare with the Coordinate Bethe Ansatz[33,34]. The Bethe equations then read

$$\left(\frac{\cosh(z_n - i\eta)}{\cosh(z_n + i\eta)}\right)^N = \prod_{\substack{j=1 \\ j \neq n}}^M \frac{\sinh(z_n - z_j - 2i\eta)}{\sinh(z_n - z_j + 2i\eta)}.$$

From the Bethe solutions $\{z_j\}$, the energy is obtained as

$$E = \sinh(2i\eta) \frac{\tau'(0)}{\tau(0)} = 2 \sum_{j=1}^M \frac{\sin(2\eta)^2}{\cos(2\eta) + \cosh(2z_j)}.$$

The total momentum obtained from (14) is again of form (16) with
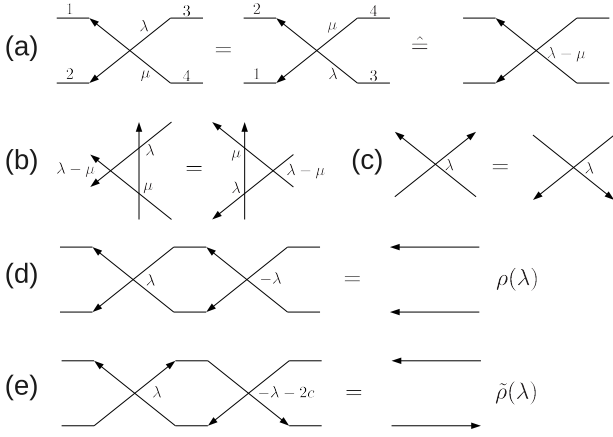
$$p_j = -2 \arctan\left(\tanh(z_j) \tan(\eta)\right).$$

FIG. 10. (a) Permutation symmetry. (b) Yang-Baxter equation. (b) Partial transpostion symmetry. (e) Crossing unitary condition. (d) Unitary condition.

In terms of the momenta p_j, the energy can be expressed as

$$E = -2 \sum_{j=1}^{M} (\Delta - \cos(p_j)).$$

The Bethe equations in their logarithmic form read

$$N p_n = 2\pi I_n + \sum_{\substack{j=1 \\ j \neq n}}^{M} \Theta(p_n, p_j),$$

with

$$\cot \frac{\Theta(p,q)}{2} = \frac{\Delta \sin \frac{p-q}{2}}{\cos \frac{p+q}{2} - \Delta \cos \frac{p-q}{2}} \qquad (17)$$

and $I_j \in \{0, \ldots, N\}$. The ground state configuration for N even and $M = N/2$ is again found with $(I_1, \ldots, I_M) = (1, 3, \ldots, N-1)$.

IV. ALGEBRAIC BETHE ANSATZ FOR OPEN BOUNDARY CONDITIONS

The method described for periodic boundary conditions is generalizeable to models with open boundary conditions and boundary fields[22,23,35]. We resketch here the Ansatz for open boundary conditions following closely Sklyanin[23] using a picturesque language.

For the following it is required that the R-tensor fulfills several conditions. To express these, it is convenient to define the permutation operator

$$P = \sum_{i,j} |j, i\rangle \langle i, j|$$

that permutes two indices. Using the matrix-notation $R(\lambda, \mu)$ from appendix III, i.e. considering the R-tensor as matrix acting on $V \otimes V$ with $V = \mathbb{C}^d$, a variant of the R-tensor with the first two indices permuted can be defined as

$$\mathcal{R}(\lambda, \mu) = P R(\lambda, \mu)$$

The basic assumption is that the R-tensor fullfills the symmetry condition

$$P R(\lambda, \mu) P = R(\lambda, \mu)$$

(see Fig. 10a). Then, the R-tensor can expressed just by two crossing arrows and it is not necessary to distinguish between them by marking them with the arguments. In fact, it is assumed in the following that R is of difference form, i.e. $R(\lambda, \mu) = R(\lambda - \mu)$. Thus, the tensor $R(\lambda - \mu)$ will be characterized by two crossing arrows together with the argument $\lambda - \mu$, as shown by the rightmost depiction in Fig. 10a. Using this notation, the Yang-Baxter equation assumes the form shown in Fig. 10b.

It is furthermore useful to define the partial transposition

$$[\mathcal{R}(\lambda)^{t_1}]^{\alpha\beta}_{\alpha'\beta'} = [\mathcal{R}(\lambda)]^{\alpha'\beta}_{\alpha\beta'},$$

which is equivalent to flipping the direction of "up-down" arrow. In analogy,

$$[\mathcal{R}(\lambda)^{t_2}]^{\alpha\beta}_{\alpha'\beta'} = [\mathcal{R}(\lambda)]^{\alpha\beta'}_{\alpha'\beta}$$

corresponds to flipping the direction of the "down-up" arrow. Accordingly, the partial transposition symmetry condition

$$\mathcal{R}(\lambda)^{t_1} = \mathcal{R}(\lambda)^{t_2}$$

is expressed by Fig. 10c.

Further conditions are the unitarity condition

$$\mathcal{R}(\lambda)\mathcal{R}(-\lambda) = \rho(\lambda) \qquad (18)$$

and the crossing unitarity condition

$$\mathcal{R}(\lambda)^{t_1} \mathcal{R}(-\lambda - 2c)^{t_1} = \tilde{\rho}(\lambda) \qquad (19)$$

with $\rho(\lambda)$ and $\tilde{\rho}(\lambda)$ being some scalar functions of λ and c denoting some constant characterizing the R-tensor. These conditions are represented by Figs. 10d and 10e.

A. Reflection Algebras

As the Bethe Ansatz for periodic boundary conditions is based on the Yang-Baxter algebra, the footing of the open boundary conditions Ansatz are the reflection algebras $\mathcal{K}^-(\lambda)$ and $\mathcal{K}^+(\lambda)$ spanned by $\{\mathcal{K}^-_{\alpha\beta}(\lambda)|\alpha, \beta = 1, \ldots, d\}$ and $\{\mathcal{K}^+(\lambda)_{\alpha\beta}|\alpha, \beta = 1, \ldots, d\}$. The graphical representation of these two algebras is shown in Fig. 11a:
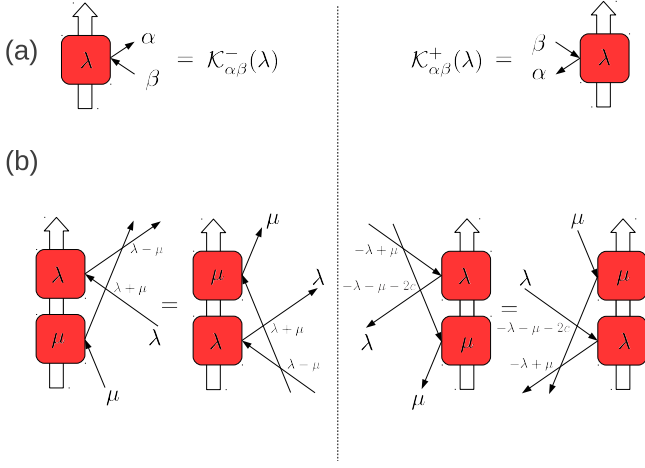
FIG. 11. (a) Graphical representation of the reflection algebras $\mathcal{K}^-(\lambda)$ and $\mathcal{K}^+(\lambda)$. The horizontal arrows indicate the virtual indices α and β; the vertical arrow indicates physical indices, i.e. the input- and output indices of the operators $\mathcal{K}^-_{\alpha\beta}(\lambda)$ and $\mathcal{K}^+_{\alpha\beta}(\lambda)$ respectively. (b) Defining equations for the reflection algebras (reflection equations). Each intersection of two lines represents an R-tensor. The argument of the R-tensor is written next to the intersection.

as in the case of the Yang-Baxter algebra, each of the two algebras is considered as a 4-index tensor with 2 "virtual" indices α and β of dimension d, represented by the horizontal arrows, and the 2 "physical" indices (corresponding to the input- and output indices of the operators $\mathcal{K}^-_{\alpha\beta}(\lambda)$ and $[\mathcal{K}^+_{\alpha\beta}(\lambda)]$ respectively), represented by the vertical arrows. The only difference to the Yang-Baxter case is that the virtual indices both are on the right-hand side of the tensor in case of $\mathcal{K}^-(\lambda)$ and on the left-hand side in case of $\mathcal{K}^+(\lambda)$. The correspondence to the Monodromy in the open boundary condition case is the matrix of operators

$$\mathcal{K}^{\pm}(\lambda) = \begin{pmatrix} \mathcal{K}^{\pm}_{11}(\lambda) & \cdots & \mathcal{K}^{\pm}_{1d}(\lambda) \\ \vdots & \ddots & \vdots \\ \mathcal{K}^{\pm}_{d1}(\lambda) & \cdots & \mathcal{K}^{\pm}_{dd}(\lambda) \end{pmatrix}.$$

The defining equations for the reflection algebras are the reflection equations, represented by the tensor network in Fig. 11b. In this figure, each intersection of two lines represents an R-tensor. The argument of the R-tensor is written next to the intersection. Algebraically, the reflection equations read

$$\mathcal{R}(\lambda - \mu)\overset{1}{\mathcal{K}}{}^-(\lambda)\mathcal{R}(\lambda + \mu)\overset{2}{\mathcal{K}}{}^-(\mu)$$
$$= \overset{2}{\mathcal{K}}{}^-(\mu)\mathcal{R}(\lambda + \mu)\overset{1}{\mathcal{K}}{}^-(\lambda)\mathcal{R}(\lambda - \mu)$$
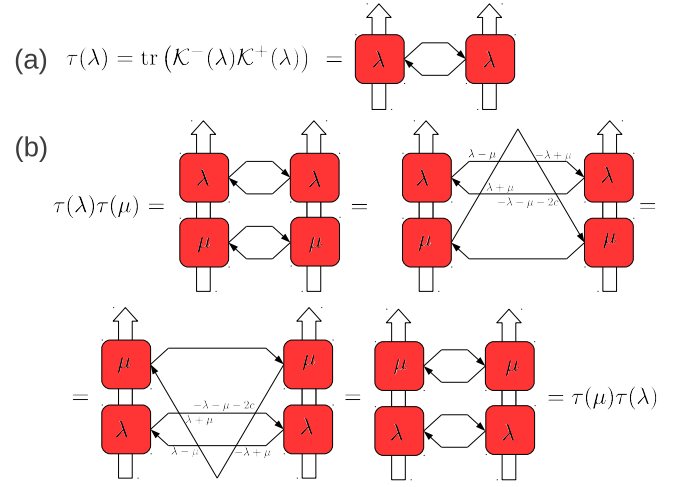

FIG. 12. (a) Definition of the transfer matrix. (b) Proof of the commuting property of the transfer matrix, $[\tau(\lambda), \tau(\mu)] = 0$.

and

$$\mathcal{R}(-\lambda + \mu)[\overset{1}{\mathcal{K}}{}^+(\lambda)]^{t_1}\mathcal{R}(-\lambda - \mu - 2c)[\overset{2}{\mathcal{K}}{}^+(\mu)]^{t_2}$$
$$= [\overset{2}{\mathcal{K}}{}^+(\mu)]^{t_2}\mathcal{R}(-\lambda - \mu - 2c)[\overset{1}{\mathcal{K}}{}^+(\lambda)]^{t_1}\mathcal{R}(-\lambda + \mu).$$

with

$$\overset{1}{\mathcal{K}}{}^{\pm}(\lambda) = \mathcal{K}^{\pm}(\lambda)\check{\otimes}\mathbb{1}$$
$$\overset{2}{\mathcal{K}}{}^{\pm}(\lambda) = \mathbb{1}\check{\otimes}\mathcal{K}^{\pm}(\lambda).$$

The outer product "$\check{\otimes}$" is thereby interpreted as in (4) and $\mathbb{1}$ is the $d \times d$ identity matrix.

Using these algebras, it is possible to define a commuting set of transfer matrices via

$$\tau(\lambda) = \mathrm{tr}\left(\mathcal{K}^-(\lambda)\mathcal{K}^+(\lambda)\right).$$

Graphically, $\tau(\lambda)$ corresponds to $\mathcal{K}^-(\lambda)$ and $\mathcal{K}^+(\lambda)$ being glued together, as shown in Fig. 12a. The commutativity of the transfer matrices, $[\tau(\lambda), \tau(\mu)] = 0$, can be proven using the unitary and crossing unitary conditions (18) and (19) and the reflection equations. The proof is sketched in Fig. 12b: starting out with $\tau(\lambda)\tau(\mu)$, the line connecting $\mathcal{K}^+(\mu)$ and $\mathcal{K}^-(\mu)$ can be pulled over the line lying above that connects $\mathcal{K}^-(\lambda)$ and $\mathcal{K}^+(\lambda)$ using (19) and over the topmost arrow connecting the two λ-algebras using (18). Next, the network is mirrored vertically by using the reflection equations. Finally, the drawn out line is pushed back using (18) and (19), which leads to $\tau(\mu)\tau(\lambda)$, as desired.

What is remaining is to find concrete representations of the reflection algebras. Examples of simple representations with physical dimension 1 have already been found.[22] More complex representations can be constructed by assembling a known representation with two
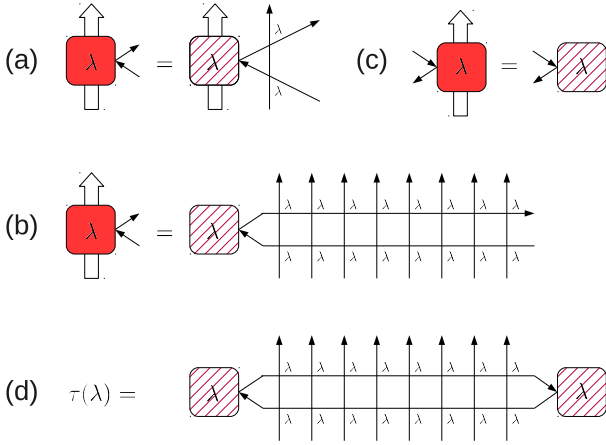
FIG. 13. (a) Composition of a new representation of $\mathcal{K}^-(\lambda)$ out of one pair of R-tensors and a known representation of $\mathcal{K}^-(\lambda)$ (that already fulfills the reflection equations). The known representation is indicated by the shaded surface. (b) Simple representation of $\mathcal{K}^+(\lambda)$ with physical dimension 1. (c) Composition of a complex representations of $\mathcal{K}^-(\lambda)$ of dimension d^N by attaching N pairs of R-tensors to a simple representation with physical dimension 1. (d) Transfer matrix built from the representations (b) and (c).
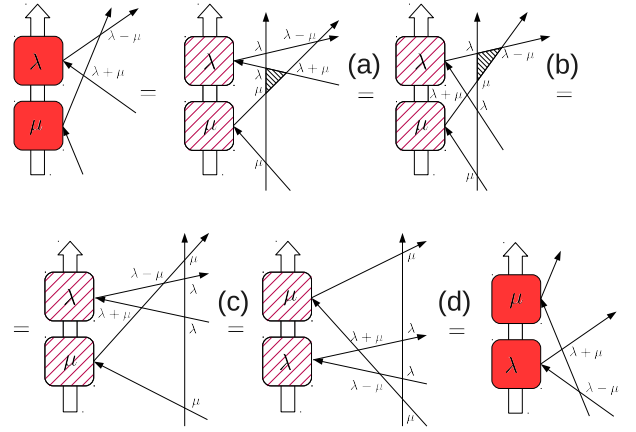
FIG. 14. Proof that the composed representation shown in Fig. 13a fulfills the reflection equations: the main idea of the first two steps (a) and (b) is to pull the vertical line rightmost by applying the Yang-Baxter equation twice. The three R-tensors to which the Yang-Baxter equation is applied are marked by the shaded triangles. The new situation now allows the application of the reflection equations, as shown in step (c). The last step (d) consists in pushing the vertical line back by applying the Yang-Baxter equation twice, such as in steps (a) and (b), but in reverse order.

R-tensors in the way shown in Fig. 13a. The physical dimension of the new representation is thereby increased by a factor d. That this assembly is indeed a valid representation can be proven using the Yang-Baxter equation and the reflection equations. The proof is sketched in Fig. 14.

Thus, starting out with a simple representation with physical dimension 1 for $\mathcal{K}^-(\lambda)$, a representation with physical dimension d^N is obtained after N iterations with the relation expressed in Fig. 13a. The structure of the representation after N iterations can be gathered from Fig. 13b. Assuming a simple representation with physical dimension 1 for $\mathcal{K}^+(\lambda)$ (depicted in Fig. 13c), the transfer matrix assumes the form shown in Fig. 13d.

For the sake of simplicity, we choose the simple representations with physical dimension 1 equal to the identity (which is a valid representation that fulfills the reflection equations). Using the notation for the fundamental representation of the Yang-Baxter algebra introduced in equation (5) and Fig. 6b, the representations of the algebras $\mathcal{K}^-(\lambda)$ and $\mathcal{K}^+(\lambda)$ look as shown in Fig. 15a. The transfer matrix assumes the form depicted in Fig. 15b. Algebraically, the representation of $\mathcal{K}^-(\lambda)$ is then the product of two MPOs,

$$\mathcal{K}^-_{\alpha\beta}(\lambda) = \sum_{s=1}^{d} \bar{K}^-_{s\alpha}(\mu) K^-_{s\beta}(\mu). \tag{20}$$

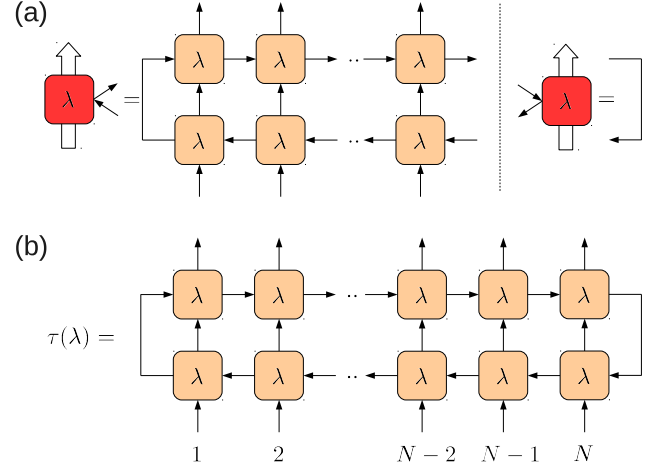In terms of the previously defined matrices $\mathcal{L}^k_l(\mu)$, the

FIG. 15. (a) Representation of $\mathcal{K}^-(\lambda)$ and $\mathcal{K}^+(\lambda)$ for open boundary conditions. (b) Transfer matrix built from these representations.

MPOs read

$$K^-_{s\beta}(\mu) = \sum_{\substack{k_1 \cdots k_N \\ l_1 \cdots l_N}} \langle s | \mathcal{L}^{k_1}_{l_1}(\mu) \cdots \mathcal{L}^{k_N}_{l_N}(\mu) | \beta \rangle o^{k_1}_{l_1} \otimes \cdots \otimes o^{k_N}_{l_N}$$

FIG. 16. Derivation of the open boundary condition Hamiltonian by derivative of the transfer matrix shown in Fig. 15b at the point λ_0. (part I).
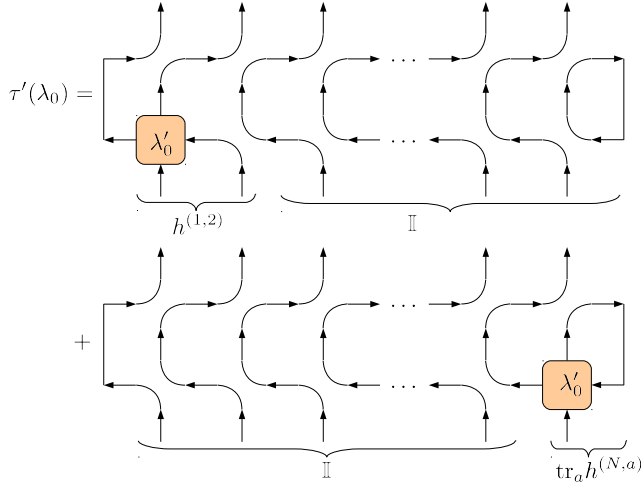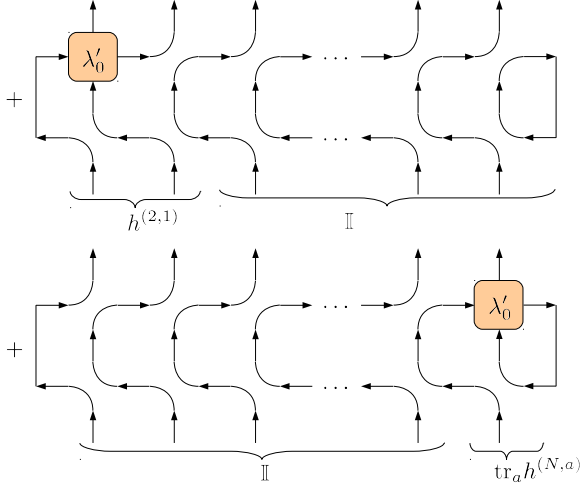

FIG. 17. Derivation of the open boundary condition Hamiltonian by derivative of the transfer matrix shown in Fig. 15b at the point λ_0. (part II).

and

$$\bar{K}^-_{s\alpha}(\mu) = \sum_{\substack{k_1\cdots k_N \\ l_1\cdots l_N}} \langle\, s\,|\mathcal{L}^{k_1}_{l_1}(\mu)^T\cdots\mathcal{L}^{k_N}_{l_N}(\mu)^T|\,\alpha\,\rangle o^{k_1}_{l_1}\otimes\cdots\otimes o^{k_N}_{l_N}$$

with $o^k_l = |\,k\,\rangle\langle\,l\,|$. The representation of $\mathcal{K}^+(\lambda)$ has physical dimension one and is equal to the identity with respect to the virtual indices, i.e.

$$\mathcal{K}^+_{\alpha\beta}(\lambda) = \delta_{\alpha\beta}.$$

The transfer matrix constructed in this way is indeed related to a local Hamiltonian with open boundary conditions. This Hamiltonian is obtained as the derivative of

the transfer matrix at the point λ_0 at which the R-tensor is equal to the identity (see (3)). Explicitly, the obtained Hamiltonian is of the form

$$H \equiv \sum_{i=1}^{N-1} h^{(i,i+1)} + \frac{1}{d}\mathrm{tr}_a h^{(N,a)} \tag{21}$$

and related to the transfer matrix via

$$H = \frac{1}{2d}\tau'(\lambda_0).$$

In (21), the symbol a refers to an auxiliary system that is traced out. Using the notation from appendix III, this relation is seen as follows: the derivative $\tau'(\lambda_0)$ disintegrates into a sum of $2N$ terms, each term containing one tensor differentiated at λ_0 and $2N-1$ tensors evaluated at λ_0. Due to the regularity condition (3) of the R-tensor, the tensors evaluated at λ_0 assume the simple form shown in Fig. 9a. As can be gathered from Figs. 16 and 17, each differentiated tensor at site i corresponds to a two-site term $h^{(i,i+1)}$ for $i = 1,\ldots,N-1$ (with h being defined in (7)). For $i = N$, two indices of the tensor are traced out, which leads to the one-site term $\mathrm{tr}_a h^{(N,a)}$.

B. Bethe Ansatz for the XXZ model with open boundary conditions

For the XXZ model, the virtual dimension d is equal to 2, such that the Monodromy can be written in the form

$$\mathcal{K}^-(\lambda) = \begin{pmatrix} \mathcal{A}(\lambda) & \mathcal{B}(\lambda) \\ \mathcal{C}(\lambda) & \mathcal{D}(\lambda) \end{pmatrix}$$

with

$$\begin{aligned} \mathcal{A}(\lambda) &= \mathcal{K}^-_{00}(\lambda), & \mathcal{C}(\lambda) &= \mathcal{K}^-_{10}(\lambda) \\ \mathcal{B}(\lambda) &= \mathcal{K}^-_{01}(\lambda), & \mathcal{D}(\lambda) &= \mathcal{K}^-_{11}(\lambda). \end{aligned}$$

The R-tensor has the form (9) with $b(\lambda)$ and $c(\lambda)$ being defined by (10) and (11). It fulfills the regularity condition (3) at the point $\lambda_0 = 0$, the unitarity condition (18) with $\rho(\lambda) = 1$ and the crossing unitarity condition (19) with $c = 2i\eta$ and $\tilde{\rho}(\lambda) = 1 - \sin(2\eta)^2/\sin(2\eta - i\lambda)^2$. Using representation (20) for $\mathcal{K}^-(\lambda)$, the R-tensor generates the Hamiltonian

$$H = \frac{1}{\sinh(2i\eta)}(H^{obc}_{XXZ}(\Delta) - \Delta)$$

with

$$H^{obc}_{XXZ}(\Delta) = \sum_{n=1}^{N-1} h_{XXZ}(\Delta).$$

The precondition for the Bethe Ansatz is that a representation must exist, for which there is a pseudo-vacuum $|\,vac\,\rangle$ that is an eigenstate of $\mathcal{A}(\lambda)$ and $\mathcal{D}(\lambda)$

and that is annihiliated by $\mathcal{C}(\lambda)$:

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

As argumented before, the operator $\mathcal{C}(\lambda)$ annihilates one down-spin, whereas $\mathcal{A}(\lambda)$ and $\mathcal{D}(\lambda)$ keep the number of down-spins constant, such that the state with all spins up is a valid pseudo-vacuum. The goal is now to diagonalize the transfer matrix $\tau(\lambda) = \mathcal{A}(\lambda) + \mathcal{D}(\lambda)$. Since all transfer matrices commute, $[\tau(\lambda), \tau(\mu)] = 0$, all eigenvectors are independent of λ. The eigenvalue problem reads

$$\tau(\lambda)|\,\Psi\,\rangle = \tau(\lambda)|\,\Psi\,\rangle.$$

The Bethe Ansatz

$$|\,\Psi(\mu_1, \ldots, \mu_M)\,\rangle = \mathcal{B}(\mu_1) \cdots \mathcal{B}(\mu_M)|\,vac\,\rangle.$$

fulfills the eigenvalue problem provided that the μ_j's fulfill the Bethe equations. The proof is based upon the algebraic relations between $\mathcal{A}(\lambda)$, $\mathcal{B}(\lambda)$, $\mathcal{C}(\lambda)$ and $\mathcal{D}(\lambda)$ and is described in detail in [23].

Defining the momenta p_j via

$$p_j = i \ln \frac{\mu_j}{\mu_j + \eta},$$

the Bethe equations in their logarithmic form read[36,37]

$$(N+1)p_n = \pi I_n + \Theta(p_n, -p_n) + \sum_{\substack{j=1 \\ j \neq n}}^{M} \frac{\Theta(p_n, -p_j) + \Theta(p_n, p_j)}{2}$$

with $\Theta(p, q)$ being defined in (17). The ground state for N even and $M = N/2$ corresponds to $(I_1, \ldots, I_M) = (1, 3, \ldots, N-1)$. The energy eigenvalue of $H_{XXZ}^{obc}(\Delta)$ for a configuration $(p_1, \ldots, p_M)$ is obtained as

$$E_{XXZ}^{obc}(\Delta) = -2\sum_{j=1}^{M}(\Delta - \cos(p_j)).$$

V. CONCLUSIONS

Summing up, we have sketched the Algebraic Bethe Ansatz using the pictoresque language of tensor networks. *In a future paper, the method will be extended to [three-dimensional] space lattices and its physical implications for cohesion, ferromagnetism and electrical conductivity will be derived.*[1]

Appendix A: Algebraic Derivation of the Bethe Equations

For completeness, we sketch here the derivation of the Bethe Equations using algebraic relations. We thereby follow Korepin[2].

The goal is to find eigenvectors of $t(\lambda) = A(\lambda) + D(\lambda)$ using algebraic relations between $A(\lambda), B(\lambda), C(\lambda)$ and $D(\lambda)$. The commutation relations that are required are

$$B(\lambda)B(\mu) = B(\mu)B(\lambda) \tag{A1}$$
$$A(\lambda)B(\mu) = f(\lambda, \mu)B(\mu)A(\lambda) + g(\lambda, \mu)B(\lambda)A(\mu) \tag{A2}$$
$$D(\lambda)B(\mu) = f(\mu, \lambda)B(\mu)D(\lambda) + g(\mu, \lambda)B(\lambda)D(\mu) \tag{A3}$$

Here, the abbreviations $f(\lambda, \mu) = 1/c(\mu, \lambda)$ and $g(\lambda, \mu) = -b(\mu, \lambda)/c(\mu, \lambda)$ are used.

The Bethe Ansatz reads

$$|\,\Psi(\mu_1, \ldots, \mu_M)\,\rangle = B(\mu_1) \cdots B(\mu_M)|\,vac\,\rangle, \tag{A4}$$

where $|\,vac\,\rangle$ is a state that is an eigenvector of $A(\lambda)$ and $D(\lambda)$ with eigenvalues $a(\lambda)$ and $d(\lambda)$, and that is annihilated by $C(\lambda)$. $A(\lambda)$ applied to $|\,\Psi(\mu_1, \ldots, \mu_M)\,\rangle$ using relation (A2) yields in principle 2^M terms, because each commutation of $A(\lambda)$ with a $B(\mu_k)$ yields 2 terms and it takes M commutations to move $A(\lambda)$ from left to right. However, these two terms are not arbitrary. Both terms only perform exchange operators: the f-term in (A2) exchanges the operators A and B, *but not* their arguments; the g-term, on the other hand, exchanges the operators A and B *and* their arguments. Due to this, after M commutations the following conditions must hold:

- Every term must contain M B's and one A.

- The $M + 1$ coefficients $(\lambda, \mu_1, \ldots, \mu_M)$ are distributed among the M B's and the one A.

Since all B's commute, there are only 2 cases: either λ is argument of A – then the term looks like

$$B(\mu_1) \cdots B(\mu_M)A(\lambda)|\,vac\,\rangle. \tag{A5}$$

Or, λ is argument of one of the B's. Then the term is of the form

$$B(\lambda) \prod_{j \neq n} B(\mu_j)A(\mu_n)|\,vac\,\rangle \tag{A6}$$

with $n \in \{1, \ldots, M\}$. Thus, the 2^M terms can be collected into $M + 1$ linearly independent terms:

$$A(\lambda)|\,\Psi(\mu_1, \ldots, \mu_M)\,\rangle = \Lambda\, B(\mu_1) \cdots B(\mu_M)A(\lambda)|\,vac\,\rangle$$
$$+ \sum_{n=1}^{M} \Lambda_n B(\lambda) \prod_{j \neq n} B(\mu_j)A(\mu_n)|\,vac\,\rangle$$

What remains to be done is the calculation of the coefficients Λ and Λ_n.

The expression (A5) is obviously obtained after M commutations using the f-term in (A2). The g-term must not be applied, because it introduces a $B(\lambda)$. Thus

$$\Lambda = \prod_{j=1}^{M} f(\lambda, \mu_j).$$

To obtain (A6), it is convenient to rewrite the Bethe Ansatz (A4) as

$$|\,\Psi(\mu_1,\ldots,\mu_M)\,\rangle = B(\mu_n) \prod_{j\neq n} B(\mu_j)|\,vac\,\rangle.$$

This is possible for all n, since all B's commute. Since expression (A6) must not contain $B(\mu_n)$, the first commutation with $A(\lambda)$ must be performed using the g-term in (A2). The expression then reads

$$g(\lambda, \mu_n) B(\lambda) A(\mu_n) \prod_{j\neq n} B(\mu_j)|\,vac\,\rangle.$$

All further commutations must use the f-term, because another use of the g-term would introduce $B(\mu_n)$ in the expression again. Thus, the coefficients must be

$$\Lambda_n = g(\lambda, \mu_n) \prod_{j\neq n} f(\mu_n, \mu_j)$$

The application of $D(\lambda)$ to $|\,\Psi(\mu_1,\ldots,\mu_M)\,\rangle$ can be treated in a similar way using relations (A1) and (A3). Again, the application yields $M+1$ terms

$$D(\lambda)|\,\Psi(\mu_1,\ldots,\mu_M)\,\rangle = \tilde{\Lambda}\, B(\mu_1)\cdots B(\mu_M)D(\lambda)|\,vac\,\rangle$$
$$+ \sum_{n=1}^{M} \tilde{\Lambda}_n B(\lambda) \prod_{j\neq n} B(\mu_j)D(\mu_n)|\,vac\,\rangle$$

The coefficients are

$$\tilde{\Lambda} = \prod_{j=1}^{M} f(\mu_j, \lambda).$$

and

$$\tilde{\Lambda}_n = g(\mu_n, \lambda) \prod_{j\neq n} f(\mu_j, \mu_n).$$

Thus, $|\,\Psi(\mu_1,\ldots,\mu_M)\,\rangle$ is an eigenvector of $t(\lambda) = A(\lambda) + D(\lambda)$ if

$$a(\mu_n)\Lambda_n + d(\mu_n)\tilde{\Lambda}_n = 0$$

for $n = 1,\ldots, M$. These relations are the Bethe Ansatz equations, which can be written in the form

$$\frac{d(\mu_n)}{a(\mu_n)} = \prod_{\substack{j=1 \\ j\neq n}}^{M} \frac{c(\mu_n, \mu_j)}{c(\mu_j, \mu_n)}$$

under the assumption that $g(\lambda, \mu)$ is an odd function in the sense that $g(\lambda, \mu) = -g(\mu, \lambda)$ (as it is the case for the Heisenberg model and the XXZ model).

The eigenvalue $\tau(\lambda)$ is obtained as

$$\tau(\lambda) = a(\lambda)\Lambda + d(\lambda)\tilde{\Lambda},$$

which can be expressed as

$$\tau(\lambda) = a(\lambda) \prod_{j=1}^{M} \frac{1}{c(\mu_j, \lambda)} + d(\lambda) \prod_{j=1}^{M} \frac{1}{c(\lambda, \mu_j)}.$$

[1] H. Bethe, Z. Phys. **71**, 205 (1931).
[2] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge University Press, 1993).
[3] N. A. Slavnov, Russ. Math. Surv. **62**, 727 (2007).
[4] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Phys. Rev. Lett. **59**, 799 (1087).
[5] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Commun. Math. Phys. **115**, 477 (1988).
[6] S. R. White, Phys. Rev. Lett **69**, 2863 (1992).
[7] S. R. White, Phys. Rev. B **48**, 10345 (1993).
[8] F. Verstraete and J. Cirac, Phys. Rev. B **73**, 094423 (2006), cond-mat/0505140.
[9] D. Perez-Garcia, F. Verstraete, M. Wolf, and J. Cirac, Quantum Inf. Comput. 7, 401 (2007) **7**, 401 (2007), arXiv:quant-ph/0608197v1.
[10] F. Verstraete, J. I. Cirac, and V. Murg, Adv. Phys. **57 (2)**, 143 (2008).
[11] S. Singh, R. N. C. Pfeifer, and G. Vidal, Phys. Rev. A **82**, 050301 (2010).
[12] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066v1 (2004).
[13] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. A **75**, 033605 (2007), cond-mat/0611522.
[14] V. Murg, F. Verstraete, and J. I. Cirac, Phys. Rev. B **79**, 195119 (2009).
[15] G. Vidal, arXiv:cond-mat/0512165v2 (2005).
[16] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).
[17] H. Katsura and I. Maruyama, J. Phys. A: Math. Theor. **43**, 175003 (2010).
[18] F. C. Alcaraz and M. J. Lazo, J. Phys. A: Math. Gen. **37**, 4149 (2004).
[19] F. C. Alcaraz and M. J. Lazo, J. Phys. A: Math. Gen. **37**, L1 (2004).
[20] F. C. Alcaraz and M. J. Lazo, J. Phys. A: Math. Gen. **39**, 11335 (2006).
[21] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett **93**, 207204 (2004), cond-mat/0406426.
[22] I. V. Cherednik, Theor. Math. Phys. **61**, 977 (1984).
[23] E. K. Sklyanin, J. Phys. A: Math. Gen. **21**, 2375 (1988).
[24] C. Gómez, M. Ruiz-Altaba, and G. Sierra, *Quantum Groups in Two-Dimensional Physics* (Cambridge University Press, 1996).
[25] F. H. L. Essler, H. Frahm, F. Göhmann, A. Klümper, and

V. E. Korepin, *The one-dimensional Hubbard model* (Cambridge University Press, 2005).

[26] A. A. Zvyagin, *Finite Size Effects in Correlated Electron Models: Exact Results* (Imperial College Press, 2005).

[27] J. Plunkett, *An introduction to the algebraic Bethe ansatz* (VDM Verlag, 2009).

[28] F. Verstraete, D. Porras, and J. I. Cirac, Phys. Rev. Lett **93**, 227205 (2004), cond-mat/0404706.

[29] L. Hulthén, Arkiv. Mat. Astron. Fysik. **26A**, No. 11 (1938).

[30] M. Karbach and G. Müller, Computers in Physics **11**, 36 (1997), cond-mat/9809162.

[31] M. Karbach, K. Hu, and G. Müller, arXiv:cond-mat/9809163v1 (1998).

[32] M. Karbach, K. Hu, and G. Müller, arXiv:cond-mat/0008018v1 (2000).

[33] C. N. Yang and C. P. Yang, Phys. Rev. **150**, 321 (1966).

[34] R. Orbach, Phys. Rev. **112**, 309 (1958).

[35] N. Kitanine, K. K. Kozlowski, J. M. Maillet, G. Niccoli, N. A. Slavnov, and V. Terras, J. Stat. Mech. , P10009 (2007).

[36] M. Gaudin, Phys. Rev. A **4**, 386 (1971).

[37] F. C. Alcaraz, M. N. Barber, M. T. Batchelor, R. J. Baxter, and G. R. W. Quispel, J. Phys. A: Math. Gen. **20**, 6397 (1987).