Scalar products of Bethe vectors in models with $gl(2|1)$ symmetry 1. Super-analog of Reshetikhin formula

A Hutsalyuk$^1$, A Liashyk$^{2,3}$, S Z Pakuliak$^{1,4}$, E Ragoucy$^{5,7}$ and N A Slavnov$^6$

1 Moscow Institute of Physics and Technology, Dolgoprudny, Moscow reg., Russia
2 Bogoliubov Institute for Theoretical Physics, NAS of Ukraine, Kiev, Ukraine
3 National Research University Higher School of Economics, Russia
4 Laboratory of Theoretical Physics, JINR, Dubna, Moscow reg., Russia
5 Laboratoire de Physique Théorique LAPTH, CNRS and Université de Savoie, BP 110, F-74941 Annecy-le-Vieux Cedex, France
6 Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, Russia

E-mail: hutsalyuk@gmail.com, a.liashyk@gmail.com, stanislav.pakuliak@jinr.ru, eric.ragoucy@lapth.cnrs.fr and nslavnov@mi.ras.ru

Received 24 June 2016
Accepted for publication 5 September 2016
Published 19 October 2016

Abstract
We study the scalar products of Bethe vectors in integrable models solvable by the nested algebraic Bethe ansatz and possessing $gl(2|1)$ symmetry. Using explicit formulas of the monodromy matrix entries’ multiple actions onto Bethe vectors we obtain a representation for the scalar product in the most general case. This explicit representation appears to be a sum over partitions of the Bethe parameters. It can be used for the analysis of scalar products involving on-shell Bethe vectors. As a by-product, we obtain a determinant representation for the scalar products of generic Bethe vectors in integrable models with $gl(1|1)$ symmetry.

Keywords: algebraic Bethe ansatz, integrable systems, scalar product, Bethe vectors

1. Introduction

The study of quantum integrable systems in the context of the quantum inverse scattering method (QISM) has been one of the great successes of the Leningrad school [1–3]. As one of the early members of this school, Petr Kulish has made several important contributions to the...
field. Among them, it is worth noticing the study of quantum integrable models solvable by the nested algebraic Bethe ansatz [4–6], the search for solutions to the Yang–Baxter equation [7, 8], and the development of the QISM for models with open boundaries [9–11]. He also contributed to generalizations to the supersymmetric versions of these models [7, 12].

In this article, we would like to make a contribution to the field he was fond of, that is, supersymmetric quantum integrable models solvable by the nested algebraic Bethe ansatz. This paper is the continuation of a series of articles devoted to quantum integrable models possessing \( \mathfrak{gl}(2|1) \) symmetry. In [13] we built explicit representations for the Bethe vectors in these models. In [14] we found multiple actions of the monodromy matrix entries onto these vectors. Now we consider the problem of the scalar products of Bethe vectors.

The scalar products of Bethe vectors play a very important role in the algebraic Bethe ansatz. They are a necessary tool for calculating form factors and correlation functions within this framework. The first results in this field concern \( \mathfrak{gl}(2) \)-based models and their \( q \)-deformations. They were obtained in [15–17], where in particular the Izergin–Korepin formula (for scalar products) was given. Concerning scalar products in models with a \( \mathfrak{gl}(3) \)-invariant \( R \)-matrix, the first result was obtained in [18]. There, an analog of the Izergin–Korepin formula for the scalar product of generic Bethe vectors was produced (the Reshetikhin formula) and a determinant representation for the norm of the eigenvectors of the transfer matrix was found. Similar results for models based on a \( q \)-deformed \( \mathfrak{gl}(3) \) algebra were obtained in [19, 20].

In this paper we consider scalar products in models with \( \mathfrak{gl}(2|1) \) symmetry. At this stage of study our goal is to obtain an analog of the Reshetikhin formula for these models. This means that we are going to find an explicit representation for the scalar product of generic Bethe vectors, in which the result is given as a sum over partitions of the Bethe parameters (sum formula). Certainly, this type of formula is not convenient for direct applications. However, it provides a key for studying particular cases of scalar products, in which the sum over partitions can be reduced to a single determinant. Such determinant representations for particular cases of scalar products were already found for models with \( \mathfrak{gl}(2) \) and \( \mathfrak{gl}(3) \) symmetries [20–23]. In all these cases the starting point was a sum formula.

The article is organized as follows. In section 2 we introduce the model under consideration and specify our conventions and notation. Section 3 contains explicit formulas for the Bethe vectors and multiple actions of the monodromy matrix entries onto them. In section 4 we present the main result of the paper. There we define scalar products and give a sum formula for them. The sum formula is proved in the remaining part of the paper. In section 5 we study successive actions of the monodromy matrix entries onto Bethe vectors. In section 6 we obtain an explicit determinant representation for the highest coefficient of the scalar product. Section 7 deals with a particular case of scalar products in \( \mathfrak{gl}(1|1) \)-based models. Section 8 contains several alternative representations for the highest coefficient. We have collected auxiliary formulas in four appendices. In particular, appendix A is devoted to the properties of the partition function of a six-vertex model with domain wall boundary conditions (DWPF). In appendix B we describe a relationship of sums over partitions and multiple contour integrals. Appendix C contains several identities allowing us to calculate certain multiple sums of rational functions. Finally, in appendix D we present the reduction properties of a determinant introduced in section 6.
2. Description of the model

2.1. \(\text{gl}(2|1)\)-based models

The \(R\)-matrix of \(\text{gl}(2|1)\)-based models acts in the tensor product \(\mathbb{C}^{2|1} \otimes \mathbb{C}^{2|1}\), where \(\mathbb{C}^{2|1}\) is the \(\mathbb{Z}_2\)-graded vector space with the grading \([1] = [2] = 0, [3] = 1\). The \(R\)-matrix has the form

\[
R(u, v) = \mathbb{1} + g(u, v)P, \quad g(u, v) = \frac{c}{u - v},
\]

(2.1)

where \(\mathbb{1}\) is the identity matrix, \(P\) is the graded permutation operator \([7]\) and \(c\) is a constant. The monodromy matrix \(T(u)\) is also graded according to the rule \([T_{ij}(u)] = [i] + [j]\). It satisfies the RTT-relation

\[
R(u, v)(T(u) \otimes \mathbb{1})(\mathbb{1} \otimes T(v)) = (\mathbb{1} \otimes T(v))(T(u) \otimes \mathbb{1})R(u, v).
\]

(2.2)

Equation (2.2) holds in the tensor product \(\mathbb{C}^{2|1} \otimes \mathbb{C}^{2|1} \otimes \mathcal{H}\), where \(\mathcal{H}\) is the Hilbert space of a Hamiltonian under consideration. Tensor products of \(\mathbb{C}^{2|1}\) spaces are graded as follows:

\[
(\mathbb{1} \otimes e_{ij}) \cdot (e_{kl} \otimes \mathbb{1}) = (-1)^{|i||j|+|k||l|} e_{il} \otimes e_{kj},
\]

(2.3)

where \(e_{ij}\) are the elementary units: \((e_{ij})_{ab} = \delta_{ia}\delta_{bj}\).

Algebra (2.2) possesses an antimorphism \([13]\)

\[
\psi(T_{ij}(u)) = (-1)^{|i||j|} T_{ji}(u), \quad \psi(AB) = (-1)^{|A||B|} \psi(B) \psi(A),
\]

(2.4)

where \(A\) and \(B\) are arbitrary operators of fixed gradings. It follows from (2.4) that

\[
\psi(A_1 \ldots A_n) = (-1)^{\sum k} \psi(A_n) \ldots \psi(A_1), \quad \psi_a = \sum_{1 \leq i < j \leq n} [A_i] : [A_j].
\]

(2.5)

The RTT-relation (2.2) implies a set of scalar commutation relations for the monodromy matrix elements. For our purpose, we will need only the following ones:

\[
[T_{ik}(v_1), T_{kj}(v_2)] = 0, \quad \text{if} \quad [T_{ik}] = 0,
\]

\[
h(v_1, v_2)T_{ij}(v_1)T_{ji}(v_2) = h(v_2, v_1)T_{ij}(v_2)T_{ji}(v_1), \quad j = 1, 2.
\]

(2.6)

Finally, the graded transfer matrix is defined as the supertrace of the monodromy matrix

\[
T(u) = \text{str} \, T(u) = \sum_{j=1}^{3} (-1)^{|j|} T_{jj}(u).
\]

(2.7)

It is a generating function of the integrals of motion, due to the relation \([T(u), T(v)] = 0\).

2.2. Notation and known results

In this paper we use the same notation and conventions as in \([14]\). Besides the function \(g(u, v)\) we shall also use two functions

\[
f(u, v) = 1 + g(u, v) = \frac{u - v + c}{u - v} \quad \text{and} \quad h(u, v) = \frac{f(u, v)}{g(u, v)} = \frac{u - v + c}{c}.
\]

(2.8)
These functions have the following obvious properties:

\[ g(u, v) = -g(v, u), \quad h(u, v + c) = \frac{1}{g(u, v)}, \quad f(u, v + c) = \frac{1}{f(v, u)} \quad (2.9) \]

We denote sets of variables by a bar: \( \bar{x}, \bar{u}, \bar{v} \) etc. Individual elements of the sets are denoted by Latin subscripts: \( u_j, v_i \) etc. As a rule, the number of elements in the sets is not shown explicitly in the equations; however we give these cardinalities in special comments to the formulas. The notation \( \bar{u} = \{u_1, \ldots, u_n\} = \{u_1 \pm c, \ldots, u_n \pm c\} \). A union of sets is denoted by braces: \( \{u, v\} \equiv \bar{u} \cup \bar{v} \).

As a rule, subsets of variables are labeled by roman subscripts: \( u_I, v_i \) etc. The only exception will be the notation \( \bar{u}_j \) for a subset that is complementary to the element \( u_j \), that is, \( \bar{u}_j = \bar{u} \setminus \{u_j\} \). The notation \( \bar{u} : \{\bar{u}_1, \bar{u}_2\} \) means that the set \( \bar{u} \) is divided into two subsets \( \bar{u}_1 \) and \( \bar{u}_2 \) such that \( \{\bar{u}_1, \bar{u}_2\} = \bar{u} \) and \( \bar{u}_1 \cap \bar{u}_2 = \emptyset \). We assume that the elements in every subset of variables are ordered in such a way that the sequence of their subscripts is strictly increasing. We call this ordering natural order.

In order to avoid too cumbersome formulas we use a shorthand notation for the products of functions depending on one or two variables. That is, if the functions \( g, f, h \) depend on a set of variables, one should take the product over the corresponding set. For example,

\[ h(\bar{u}, \bar{v}) = \prod_{u_j \in \bar{u}} h(u_j, v_j); \quad f(\bar{u}_j, \bar{v}_j) = \prod_{v_j \in \bar{v}_j} f(u_j, v_j); \quad g(\bar{v}_1, \bar{v}_2) = \prod_{v_j \in \bar{v}_1} \prod_{v_k \in \bar{v}_2} g(v_j, v_k). \quad (2.10) \]

We use the same prescription for the products of even commuting operators \( T_{ij} \) and for vacuum eigenvalues \( \lambda_k \) and \( r_k \) (see (3.1), (3.2)), for instance,

\[ \lambda_2(\bar{v}) = \prod_{v_j \in \bar{v}} \lambda_2(v_j); \quad r_3(\bar{u}) = \prod_{u_j \in \bar{u}} r_3(u_j); \quad T_{21}(\bar{v}_1) = \prod_{v_j \in \bar{v}_1} T_{21}(v_j). \quad (2.11) \]

For the odd operators \( T_{i\beta} \) and \( T_{j\gamma} \), which exchange with a multiplication factor, we introduce for an arbitrary set \( \bar{v} = \{v_1, \ldots, v_n\} \)

\[ T_{j\beta}(\bar{v}) = \frac{T_{j\beta}(v_1) \ldots T_{j\beta}(v_n)}{\prod_{m \geq l \geq 1} h(v_l, v_m)}, \quad T_{i\gamma}(\bar{v}) = \frac{T_{i\gamma}(v_1) \ldots T_{i\gamma}(v_n)}{\prod_{m \geq l \geq 1} h(v_m, v_l)}, \quad j = 1, 2. \quad (2.12) \]

Due to the commutation relations (2.6) the operator products (2.12) are symmetric over the parameters \( \bar{v} \) and play the same role as the bosonic products \( T_{ij}(\bar{u}) \).

We would like to stress that the convention on the shorthand notation for the products concerns functions (operators) depending on one or two variables only. It should not be applied to functions which by definition might depend on many variables.

An example of a function depending on two sets of variables is the DWPF [15, 17]. We denote it by \( K_n(\bar{u}|\bar{v}) \). It depends on two sets of variables \( \bar{u} \) and \( \bar{v} \), and the subscript shows that \( \#\bar{u} = \#\bar{v} = n \). The function \( K_n \) has the following determinant representation
\[
K_n(\bar{u}|\bar{v}) = \Delta'_n(\bar{u})\Delta_n(\bar{v})h(\bar{u}, \bar{v})\det_n \left( \frac{g(u_j, v_k)}{h(u_j, v_k)} \right),
\]  
(2.13)

where \(\Delta'_n(\bar{u})\) and \(\Delta_n(\bar{v})\) are
\[
\Delta'_n(\bar{u}) = \prod_{j<k} g(u_j, u_k), \quad \Delta_n(\bar{v}) = \prod_{j>k} g(v_j, v_k).
\]  
(2.14)

It is easy to see that \(K_n\) is symmetric over \(\bar{u}\) and symmetric over \(\bar{v}\); however \(K_n(\bar{u}|\bar{v}) = K_n(\bar{v}|\bar{u})\). Some other properties of DWPF are given in appendix A.

3. Necessary tools

In this section we recall some explicit formulas for the Bethe vectors and describe the multiple actions of the operators \(T_{ij}\) onto them.

3.1. Bethe vectors

Generic Bethe vectors and their dual vectors are denoted respectively by \(B_{uv;ab}(\bar{u}|\bar{v})\) and \(C_{uv;ab}(\bar{u}|\bar{v})\). They are parameterized by two sets of complex parameters (Bethe parameters) \(\bar{u} = u_1, \ldots, u_a\) and \(\bar{v} = v_1, \ldots, v_b\) with \(a, b = 0, 1, \ldots\). The reader can find several explicit representations for the Bethe vectors in terms of the monodromy matrix entries in [13]. Here we use some of them (see below). Vectors \(|0\rangle = B_{0,0}(\bar{u}; \bar{v})\) and \(|0\rangle = C_{0,0}(\bar{u}; \bar{v})\) respectively are called a pseudovacuum vector and a dual pseudovacuum vector. They are eigenvectors of the diagonal entries of the monodromy matrix
\[
T_i(u)|0\rangle = \lambda_i(u)|0\rangle, \quad \langle 0|T_i(u) = \lambda_i(u)\langle 0|, \quad i = 1, 2, 3.
\]  
(3.1)

where \(\lambda_i(u)\) represents some scalar functions. In the framework of the generalized model [15] considered in this paper, they remain free functional parameters. Below it will be convenient to deal with the ratios of these functions
\[
r_1(u) = \frac{\lambda_1(u)}{\lambda_2(u)}, \quad r_3(u) = \frac{\lambda_3(u)}{\lambda_2(u)}.
\]  
(3.2)

Now let us give one of the explicit representations for the Bethe vectors obtained in [13, 24]
\[
B_{uv;ab}(\bar{u}; \bar{v}) = \sum_{\bar{u}_1} K_n(\bar{v}_1|\bar{u}_1) f(\bar{u}_1, \bar{u}_b)g(\bar{v}_b, \bar{v}_1)_{\lambda_1(\bar{v})\lambda_2(\bar{u})f(\bar{v}, \bar{u}) T_{12}(\bar{u}\bar{u}) T_{13}(\bar{v}\bar{v}) T_{23}(\bar{u}\bar{v}) |0\rangle.
\]  
(3.3)

Here \(K_n(\bar{v}_1|\bar{u}_1)\) is the DWPF (2.13). The sum is taken over partitions \(\bar{u} \mapsto \{\bar{u}_1, \bar{u}_b\}\) and \(\bar{v} \mapsto \{\bar{v}_1, \bar{v}_b\}\), where \#\(\bar{u}_1 = \#\bar{u} = n\), and \(n = 0, 1, \ldots, \min(a, b)\). Recall that the notation \(T_{ij}(\bar{u}\bar{u})\), \(g(\bar{v}_b, \bar{v}_1)\), and so on means the products of the operators (functions) over the corresponding subset (see (2.10)–(2.12)).

Explicit representations for dual Bethe vectors will play a more important role in our calculations. We use two of them [13, 24]:
\[
C_{uv;ab}(\bar{u}; \bar{v}) = (-1)^{\epsilon_{ab}} \sum_{\bar{u}_1} K_n(\bar{v}_1|\bar{u}_1) f(\bar{u}_1, \bar{u}_b)g(\bar{v}_b, \bar{v}_1)_{\lambda_2(\bar{v})\lambda_2(\bar{u})f(\bar{v}, \bar{u}) T_{21}(\bar{u}\bar{u}) T_{12}(\bar{v}\bar{v}) T_{31}(\bar{v}_1),
\]  
(3.4)
and

\[
\mathcal{C}_{a,b}(\tilde{u}; \tilde{v}) = (-1)^{\frac{a+b}{2}} \sum g(\tilde{v}_1, \tilde{u}_1) \frac{f(\tilde{u}_\Phi, \tilde{u}_1) f(\tilde{v}_1, \tilde{u}_\Phi) g(\tilde{u}_\Phi, \tilde{v}_1) h(\tilde{u}_1, \tilde{u}_1)}{\lambda_2(\tilde{u}) \lambda_2(\tilde{v}) f(\tilde{v}, \tilde{u})} \times \langle 0 | T_{22}(\tilde{v}_\Phi) T_{31}(\tilde{u}_1) T_{21}(\tilde{u}_\Phi) \rangle.
\]

(3.5)

Here the sum is taken over the same partitions of the sets \(\tilde{u}\) and \(\tilde{v}\) as in (3.3).

3.2. Multiple actions of \(T\)

Explicit formulas for the multiple actions of the operators \(T_{ij}\) onto Bethe vectors were obtained in [14]. For our goal we need the actions of \(T_{ij}\) with \(i > j\). Below we give the corresponding formulas. Everywhere in this section \(\eta = \{\tilde{u}, \xi\}\), \(\xi = \{\tilde{v}, \zeta\}\), \#\(\tilde{u}\) = \(a\), \#\(\tilde{v}\) = \(b\), and \#\(\xi\) = \(n\).

3.2.1. Multiple action of \(T_{21}\). The multiple action of the operators \(T_{21}\) onto Bethe vectors reads

\[
T_{21}(\xi) B_{a,b}(\tilde{u}; \tilde{v}) = \lambda_2(\xi) h(\tilde{z}, \tilde{z}) \sum r_1(\eta_1) \frac{f(\eta_\Phi, \eta_1) f(\eta_\Phi, \eta_\Phi) g(\tilde{z}_\Phi, \xi_1) g(\tilde{z}_\Phi, \xi_1)}{h(\eta_1, \tilde{z}) h(\xi_1, \tilde{z})} \times K_n(\xi_\Phi + c) B_{a-b}(\eta_\Phi; \xi_\Phi).
\]

(3.6)

Here the function \(K_n\) represents the DWPF (2.13). The sum is taken over partitions \(\xi = (\xi_\Phi, \xi_\Phi)\) and \(\eta = (\eta_\Phi, \eta_\Phi, \eta_\Phi)\) with \#\(\xi_\Phi\) = \#\(\eta_1\) = \#\(\eta_\Phi\) = \(n\). If \(n > a\), then the product \(T_{21}(\xi)\) annihilates \(B_{a,b}(\tilde{u}; \tilde{v})\).

3.2.2. Multiple action of \(T_{31}\). The multiple action of \(T_{31}\) has the following form:

\[
T_{31}(\xi) B_{a,b}(\tilde{u}; \tilde{v}) = (-1)^{\frac{a+b}{2}} \lambda_2(\xi) h(\tilde{z}, \tilde{z}) \sum r_3(\xi_1) r_1(\eta_1) \frac{g(\tilde{z}_\Phi, \xi_1) g(\tilde{z}_\Phi, \xi_1) h(\xi_1, \tilde{z})}{h(\eta_1, \tilde{z}) h(\xi_1, \tild{z})} \times K_n(\xi_\Phi + c) B_{a-b}(\eta_\Phi; \xi_\Phi).
\]

(3.7)

Here the sum is taken over partitions \(\xi = (\xi_\Phi, \xi_\Phi, \xi_\Phi)\) and \(\eta = (\eta_\Phi, \eta_\Phi, \eta_\Phi, \eta_\Phi)\) with \#\(\xi_\Phi\) = \#\(\xi_\Phi\) = \#\(\eta_1\) = \#\(\eta_\Phi\) = \(n\). If \(n > \min(a, b)\), then the product \(T_{31}(\xi)\) annihilates \(B_{a,b}(\tilde{u}; \til{v})\).

Actually, for this action we will need only the particular case \(n = a\). Then

\[
T_{31}(\xi) B_{a,b}(\tilde{u}; \til{v}) = (-1)^{\frac{a+b}{2}} \lambda_2(\xi) h(\til{z}, \til{z}) \sum r_3(\xi_1) r_1(\eta_1) \frac{g(\til{z}_\Phi, \xi_1) g(\til{z}_\Phi, \xi_1) h(\xi_1, \til{z})}{h(\eta_1, \til{z}) h(\xi_1, \til{z})} \times K_n(\eta_\Phi + c) B_{0,b-a}(\xi_\Phi; \xi_\Phi).
\]

(3.8)

Here the sum is taken over partitions \(\xi = (\xi_\Phi, \xi_\Phi, \xi_\Phi)\) and \(\eta = (\eta_\Phi, \eta_\Phi)\) with \#\(\xi_\Phi\) = \#\(\xi_\Phi\) = \#\(\eta_1\) = \#\(\eta_\Phi\) = \(n\).
3.2.3. Multiple action of $T_{32}$

The multiple action of $T_{32}(z)$ reads

$$
T_{32}(z)B_{a,b}(\bar{a}; \bar{v}) = (-1)^{\#\xi} \lambda_2(z) h(\zeta, z) \sum r_3(\xi) \frac{f(\eta_1, \eta_2)g(\xi_{\#} \zeta_{\#})g(\xi_{\#} \zeta_{\#})g(\xi_{\#} \zeta_{\#})}{h(\eta_1, z) h(\xi, z) h(\zeta, z) f(\xi, \eta_2)}
\times h(\eta_1, \eta_2) B_{a,b-n}(\eta_1; \zeta_{\#}).
$$

(3.9)

Here the sum is taken over partitions $\bar{\xi} \Rightarrow (\xi, \bar{\zeta}, \bar{\eta})$ and $\eta \Rightarrow (\eta_1, \eta_2)$ with $\#\zeta = \#\bar{\xi} = \#\eta = n$. If $n > b$, then the result of this action vanishes.

If $a = 0$ and $n = b$, then

$$
T_{32}(z)B_{0,b}(\bar{a}; \bar{v}) = (-1)^{\#\xi} \lambda_2(z) \sum r_3(\xi) g(\xi_{\#} \zeta_{\#}) |0\rangle.
$$

(3.10)

The sum is taken over partitions $\bar{\xi} \Rightarrow (\xi, \bar{\zeta}, \bar{\eta})$ with $\#\bar{\xi} = \#\bar{\eta} = n = n$.

4. General form of the scalar product

The scalar product of Bethe vectors is defined as

$$
S_{a,b} \equiv S_{a,b}(\bar{a}^C; \bar{v}^{|B^B}; \bar{v}^B) = C_{a,b}(\bar{a}^C; \bar{v}^{|B^B})B_{a,b}(\bar{a}^B; \bar{v}^B),
$$

(4.1)

where all the Bethe parameters are generic complex numbers. We have added the superscripts $C$ and $B$ to the sets $a$, $b$ in order to stress that the vectors $S_{a,b}$ and $B_{a,b}$ may depend on different sets of parameters.

Being a scalar function, the scalar product is invariant under the action of the antimorphism $\psi$ (2.4)

$$
\psi(S_{a,b}(\bar{a}^C; \bar{v}^{|B^B}; \bar{v}^B)) = S_{a,b}(\bar{a}^C; \bar{v}^{|B^B})|\bar{a}^B; \bar{v}^B\rangle.
$$

(4.2)

On the other hand, acting with $\psi$ on the rhs of (4.1) and using the explicit representations (3.3) and (3.4) for the Bethe vectors we find

$$
\psi(C_{a,b}(\bar{a}^C; \bar{v}^{|B^B})B_{a,b}(\bar{a}^B; \bar{v}^B)) = C_{a,b}(\bar{a}^B; \bar{v}^B)B_{a,b}(\bar{a}^B; \bar{v}^B) = S_{a,b}(\bar{a}^B; \bar{v}^{|B^B}) \bar{v}^{|B^B}.
$$

(4.3)

Here we have used $\psi(T_{3j}) = T_{3j}$ and $\psi(T_{3j}) = -T_{3j}$ for $j = 1, 2$ (see (2.4)). Then using (2.5), and the fact that the total number of odd operators $T_{3j}$ and $T_{3j}$ with $j = 1, 2$ in the scalar product is equal $2b$, we arrive at (4.3). Thus, we conclude that the scalar product is invariant under the permutation of the sets $\{\bar{a}^C; \bar{v}^{|B^B}\} \leftrightarrow \{\bar{a}^B; \bar{v}^{|B^B}\}$:

$$
S_{a,b}(\bar{a}^C; \bar{v}^{|B^B}; \bar{v}^B) = S_{a,b}(\bar{a}^B; \bar{v}^{|B^B}; \bar{v}^B).
$$

(4.4)

In order to calculate the scalar product one can take an explicit formula for the dual Bethe vector (3.4) or (3.5) and then use the formulas of the multiple actions (3.6)–(3.10). Basing on these formulas we can present the scalar product of Bethe vectors in the following schematic form:

$$
S_{a,b}(\bar{a}^C; \bar{v}^{|B^B}; \bar{v}^B) = \sum r_1(\bar{w}_1) r_3(\bar{w}_1) W_{\text{par}}(\bar{w}_1; \bar{w}_1; \bar{w}_1).
$$

(4.5)

Here the set $\bar{w}$ is the union of all the Bethe parameters: $\bar{w} = \{\bar{a}^C, \bar{a}^B, \bar{v}^C, \bar{v}^{|B^B}\}$. The sum is taken over partitions of this set into three subsets $\bar{w} \Rightarrow (\bar{w}_1, \bar{w}_1, \bar{w}_1)$. The function $W_{\text{par}}$ represents some rational coefficients. Their explicit forms are not important for now. We stress in (4.5) that a part of the Bethe parameters $\bar{w}_1$ becomes the arguments of the function $r_1$, while the parameters $\bar{w}_1$ become the arguments of the function $r_3$. The remaining parameters $\bar{w}_1$ enter the rational functions $W_{\text{par}}$ only.
Let us call the set \{\tilde{u}^C, \tilde{u}^B\} parameters of the \(u\)-type. Correspondingly, we call the set \{\bar{v}^C, \bar{v}^B\} parameters of the \(v\)-type.

**Proposition 4.1.** The set \(\bar{w}_i\) in (4.5) consists of parameters of the \(u\)-type only, while the set \(\tilde{w}_i\) consists of parameters of the \(v\)-type, that is, \(\bar{w}_i \subset \{\tilde{u}^C, \tilde{u}^B\}\) and \(\tilde{w}_i \subset \{\bar{v}^C, \bar{v}^B\}\). Moreover, \(#\bar{w}_i = a\) and \(#\tilde{w}_i = b\).

**Proof.** Let us prove that \(\bar{w}_i \subset \{\tilde{u}^C, \tilde{u}^B\}\). For this we take the dual Bethe vector in the form (3.5). Let us fix a partition \(\tilde{u}^C \Rightarrow \{\tilde{u}_i^C, \tilde{u}_i^B\}\) in (3.5), such that \(#\tilde{u}_i^C = n, n = 0, 1, \ldots, \min(a, b)\). Calculating the scalar product we first act with the operators \(T_{21}(\tilde{u}_i^C)\) onto the Bethe vector. Then due to (3.6) we obtain a sum over partitions of the set \(\{\bar{u}_i^C, \bar{u}_i^B\}\). The terms of this sum are proportional to the products of the function \(r_i(\eta_i)\), where \(\eta_i \subset \{\bar{u}_i^C, \bar{u}_i^B\}\) and \(#\eta_i = a - n\). Hence, the parameters \(\eta_i\) are of the \(u\)-type.

Next, we act with the operators \(T_{31}(\tilde{v}_i^C)\) onto the obtained Bethe vectors via (3.8). We get new partitions of the set \(\{\bar{u}_i^C, \bar{u}_i^B\} \setminus \eta_i\) and new products of function \(r_i\), say, \(r_i(\eta_i)\). Obviously, \(\eta_i \subset \{\tilde{v}_i^C, \tilde{v}_i^B\}\) and \(#\eta_i = n\). Thus, the total number of the function \(r_i\) is equal to \(a\), and all its arguments are of the \(u\)-type.

Finally, we should act with the product of the operators \(T_{32}(\tilde{v}_i^C)\). But due to (3.9) this action does not produce new functions \(r_i\). Thus, we have proved that \(\bar{w}_i \subset \{\bar{v}^C, \bar{v}^B\}\) and \(#\tilde{w}_i = b\).

Similarly, one can prove that \(\tilde{w}_i \subset \{\bar{v}^C, \bar{v}^B\}\) and \(#\bar{w}_i = a\). However, for this one should take representation (3.4) for the dual Bethe vector. Then all the functions \(r_i\) will be produced under the successive actions of the operators \(T_{32}(\bar{v}_i^C)\) and \(T_{31}(\bar{v}_i^C)\). Repeating the considerations above we prove that all the parameters \(\tilde{w}_i\) are of the \(v\)-type and their total number is equal to \(b\). \(\square\)

Due to proposition 4.1 one can recast (4.5) in the form

\[
S_{a,b}(\bar{u}^C; \bar{v}^C|\bar{u}^B; \bar{v}^B) = \sum r_1(\bar{\eta}_1) r_2(\bar{\xi}_1) W_{\text{part}}(\bar{\eta}_1; \bar{\eta}_2; \bar{\xi}_1; \bar{\xi}_2).
\]

(4.6)

Here \(\bar{\eta} = \{\tilde{u}^C, \tilde{u}^B\}\) and \(\bar{\xi} = \{\bar{v}^C, \bar{v}^B\}\). The sum is taken over partitions \(\eta = \{\eta_1, \eta_2\}\) and \(\xi = \{\xi_1, \xi_2\}\), such that \(#\eta_1 = a\) and \(#\eta_2 = b\). Setting in (4.6)

\[
\eta_1 = \{\tilde{u}_1^C, \tilde{u}_1^B\}, \quad \eta_2 = \{\tilde{u}_2^C, \tilde{u}_2^B\}, \quad \#\tilde{u}_i^C = \#\tilde{u}_i^B = k, \quad k = 1, \ldots, a;
\]

\[
\xi_1 = \{\bar{v}_1^C, \bar{v}_1^B\}, \quad \xi_2 = \{\bar{v}_2^C, \bar{v}_2^B\}, \quad \#\bar{v}_i^B = \#\bar{v}_i^C = n, \quad n = 1, \ldots, b,
\]

(4.7)

we arrive at a representation

\[
S_{a,b}(\bar{u}^C; \bar{v}^C|\bar{u}^B; \bar{v}^B) = \sum r_1(\bar{u}_i^C) r_1(\bar{a}_i^B) r_2(\bar{v}_i^C) r_2(\bar{v}_i^B) W_{\text{part}} \left( \frac{\bar{a}_1^B, \bar{a}_2^B, \bar{a}_1^C, \bar{a}_2^C}{\bar{v}_1^B, \bar{v}_1^C, \bar{v}_2^B, \bar{v}_2^C} \right)
\]

(4.8)

Here the sum runs over all the partitions \(\bar{a}^C \Rightarrow \{\bar{a}_1^C, \bar{a}_2^C\}\), \(\bar{a}^B \Rightarrow \{\bar{a}_1^B, \bar{a}_2^B\}\), \(\bar{v}^C \Rightarrow \{\bar{v}_1^C, \bar{v}_2^C\}\) and \(\bar{v}^B \Rightarrow \{\bar{v}_1^B, \bar{v}_2^B\}\) with \(#\bar{a}_i^C = \#\bar{a}_i^B\) and \(#\bar{v}_i^C = \#\bar{v}_i^B\). The function \(W_{\text{part}}\) represents rational coefficients, which depend on the partitions but do not depend on the functions \(r_1\) and \(r_2\). We also have extracted explicitly the product \(f(\bar{v}_i^C, \bar{a}_i^C) f(\bar{v}_i^B, \bar{a}_i^B)\), which plays the role of a normalization factor.

**Definition 4.1.** We call the highest coefficient \(Z_{a,b}(\bar{u}^C; \bar{v}^C|\bar{v}^B)\) the function \(W_{\text{part}}\) that corresponds to the extreme partitions \(\bar{a}_i^C = \bar{a}_i^B = \emptyset\) and \(\bar{v}_i^C = \bar{v}_i^B = \emptyset\):
\[
W_{\text{part}}\left(\tilde{u}^C, \tilde{u}^B, \varnothing, \varnothing; \tilde{v}^C, \tilde{v}^B, \varnothing, \varnothing\right) = Z_{\alpha, b}(\tilde{u}^C; \tilde{u}^B|\tilde{v}^C; \tilde{v}^B). \quad (4.9)
\]

In other words, \(Z_{\alpha, b}(\tilde{u}^C; \tilde{u}^B|\tilde{v}^C; \tilde{v}^B)\) is the coefficient of the product \(r_1(\tilde{u}^C)r_3(\tilde{v}^B)\).

One can also define the conjugated highest coefficient corresponding to the extreme partition \(\tilde{u}^C = \tilde{u}^B = \varnothing\) and \(\tilde{v}^C = \tilde{v}^B = \varnothing\), that is, the coefficient of the product \(r_1(\tilde{u}^B)r_3(\tilde{v}^C)\). However, due to (4.4) it is clear that
\[
W_{\text{part}}\left(\varnothing, \varnothing, \tilde{u}^C, \tilde{u}^B; \tilde{v}^C, \tilde{v}^B, \varnothing, \varnothing\right) = Z_{\alpha, b}(\tilde{u}^B; \tilde{u}^C|\tilde{v}^B; \tilde{v}^C). \quad (4.10)
\]

We will show that all other coefficients \(W_{\text{part}}\) are equal to the bilinear combinations of the highest coefficient and its conjugate.

**Proposition 4.2.** For a fixed partition with \(\# \tilde{u}^C = \# \tilde{u}^B = k\) and \(\# \tilde{v}^C = \# \tilde{v}^B = n\), (where \(k = 0, \ldots, a\) and \(n = 0, \ldots, b\)), the coefficient \(W_{\text{part}}\) has the form
\[
W_{\text{part}}\left(\tilde{u}^C, \tilde{u}^B, \tilde{v}^C, \tilde{v}^B\right) = f(\tilde{u}^B, \tilde{v}^C)f(\tilde{u}^C, \tilde{v}^B)g(\tilde{v}^B, \tilde{v}^C)g(\tilde{v}^C, \tilde{v}^B)Z_{\alpha - k, n}(\tilde{u}^C; \tilde{u}^B|\tilde{v}^C; \tilde{v}^B)Z_{\alpha, b - n}(\tilde{a}^B; \tilde{a}^C|\tilde{v}^B; \tilde{v}^C). \quad (4.11)
\]

The main goal of this paper is to find an explicit formula for the highest coefficient \(Z_{\alpha, b}\) and to prove representation (4.11) for the coefficient \(W_{\text{part}}\).

Comparing (4.11) with the analogous formula for \(\text{gl}(3)\)-based models [18] one can see that they are very similar. It is enough to replace the product \(g(\tilde{v}^B, \tilde{v}^C)g(\tilde{v}^C, \tilde{v}^B)\) in (4.11) with the product \(f(\tilde{v}^B, \tilde{v}^C)f(\tilde{v}^C, \tilde{v}^B)\) and we obtain the formula of the paper [18]. One should remember, however, that the highest coefficients also have different representations in models described by \(\text{gl}(2|1)\) and \(\text{gl}(3)\) algebras. In particular, we will see that in the case under consideration the highest coefficient \(Z_{\alpha, b}\) admits a single determinant representation, while in the \(\text{gl}(3)\) case such a determinant formula is not known.

5. Successive actions

In the previous section we have described how the scalar product depends on the function \(r_k\). Our goal now is to explicitly find the rational coefficients \(W_{\text{part}}\). For this we calculate the successive action of the operators \(T_{ij}\) with \(i > j\) onto a generic Bethe vector. This calculation is quite similar to the one of the work in [14], where we computed the multiple actions of the monodromy matrix entries.

5.1. Successive action of \(T_{31}(x) T_{21}(y)\)

We start with the successive action of the products \(T_{31}(\tilde{x})T_{21}(\tilde{y})\). Let \(# \tilde{x} = n\) and \(# \tilde{y} = a - n\) where \(n = 0, 1, \ldots, \min(a, b)\). Define
\[
G_{\alpha, \beta}(\tilde{x}, \tilde{y}) = \frac{T_{31}(\tilde{x})T_{21}(\tilde{y})}{\lambda_2(\tilde{x})\lambda_2(\tilde{y})}B_{\alpha, \beta}(\tilde{u}; \tilde{v}). \quad (5.1)
\]
Using (3.6) and (3.8) successively, we obtain

\[
G_{n,a}(\bar{x}, \bar{y}) = (-1)^{\text{sign}(\bar{y})} h(\bar{x}, \bar{y}) h(\bar{y}, \bar{x}) \sum_{\tau} \frac{f(\eta_B \eta_i) f(\eta_B \eta_i)}{h(\xi, \eta)} g(\xi, \eta) 
\]
\[
\times K_{a,n}(\bar{y} | \eta_B + c) K_{a,n}(\eta | \xi_i + c) h(\xi, \eta) r_f(\eta_i) 
\]
\[
\times g(\xi, \eta) g(\xi, \eta) g(\xi, \bar{\xi} i) g(\xi, \bar{\xi}) f(\bar{\xi}, \eta) 
\]
\[
\times \frac{f(\eta_i, \eta_i)}{f(\xi, \eta_i)} h(\eta_i, \eta_i) K_n(\eta_i) \xi_i + c) B_{0, b, n}(\emptyset; \eta_i). 
\]

The sum is organized as follows. First the sets \( \{ \bar{y}, \bar{u} \} \) and \( \{ \bar{y}, \bar{v} \} \) are divided respectively into subsets \( \{ \eta, \eta_B, \eta_i \} \) and \( \{ \xi, \xi_B, \xi_i \} \) with the restriction \#\( \xi_i = \#\eta_i = n \). Then the

\[
G_{n,a}(\bar{x}, \bar{y}) = (-1)^{\text{sign}(\bar{y})} h(\bar{x}, \bar{y}) h(\bar{y}, \bar{x}) \sum_{\tau} \frac{f(\eta_B \eta_i) f(\eta_B \eta_i)}{h(\xi, \eta)} g(\xi, \eta) 
\]
\[
\times K_{a,n}(\bar{y} | \eta_B + c) K_{a,n}(\eta | \xi_i + c) h(\xi, \eta) r_f(\eta_i) 
\]
\[
\times g(\xi, \eta) g(\xi, \eta) g(\xi, \bar{\xi} i) g(\xi, \bar{\xi}) f(\bar{\xi}, \eta) 
\]
\[
\times \frac{f(\eta_i, \eta_i)}{f(\xi, \eta_i)} h(\eta_i, \eta_i) K_n(\eta_i) \xi_i + c) B_{0, b, n}(\emptyset; \eta_i). 
\]

(5.2)

Observe that the restrictions \( \bar{x} \cap \bar{\eta} = \emptyset \) and \( \bar{x} \cap \bar{\xi} = \emptyset \) hold automatically due to the presence of the product \( f(\eta_B, \bar{x}) \) in the denominator of (5.3) and the product \( g(\xi, \bar{\xi}) \) in the denominator of (5.4). Indeed, \( 1/f(\eta_B, \bar{x}) = 0 \) as soon as \( \bar{x} \cap \bar{\eta} = \emptyset \) and \( 1/g(\xi, \bar{\xi}) = 0 \) as soon as \( \bar{x} \cap \bar{\xi} = \emptyset \). Actually, one can easily see that the condition \( \bar{x} \cap \bar{\eta} = \emptyset \) holds, although the product \( f(\bar{x}, \eta_B) \) in the denominator of (5.3) is compensated by the same product in the numerator of (5.4). Indeed, we have seen that \( \bar{x} \cap \bar{\xi} = \emptyset \), that is to say, \( \bar{x} \subset \{ \xi, \xi_B, \xi_i \} \). But in this case \( \bar{x} \cap \bar{\eta} = \emptyset \) due to the products of the \( f \)-functions in the denominator of (5.4).

Thus, we can recast (5.2) as follows:

\[
G_{n,a}(\bar{x}, \bar{y}) = (-1)^{\text{sign}(\bar{y})} h(\bar{x}, \bar{y}) h(\bar{y}, \bar{x}) \sum_{\tau} \frac{f(\eta_B \eta_i) f(\eta_B \eta_i)}{h(\xi, \eta)} g(\xi, \eta) 
\]
\[
\times K_{a,n}(\bar{y} | \eta_B + c) K_{a,n}(\eta | \xi_i + c) h(\xi, \eta) r_f(\eta_i) 
\]
\[
\times g(\xi, \eta) g(\xi, \eta) g(\xi, \bar{\xi} i) g(\xi, \bar{\xi}) f(\bar{\xi}, \eta) 
\]
\[
\times \frac{f(\eta_i, \eta_i)}{f(\xi, \eta_i)} h(\eta_i, \eta_i) K_n(\eta_i) \xi_i + c) B_{0, b, n}(\emptyset; \eta_i). 
\]

(5.5)
Here we have also used
\[ h(\tilde{x}, \bar{x})h(\tilde{x}, \bar{x}) = \frac{h(\tilde{x}, \bar{x})}{h(\tilde{x}, \bar{x})}. \] (5.6)

In (5.5) the sum is taken over partitions \( \tilde{n} \Rightarrow \{ \tilde{n}_1, \tilde{n}_2, \tilde{n}_3, \tilde{n}_4 \} \) and \( \tilde{\xi} \Rightarrow \{ \tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3, \tilde{\xi}_4 \} \). The restrictions are imposed on the cardinalities of the subsets only.

One can reduce the number of subsets in (5.5). Let \( \tilde{n}_1 = \{ \tilde{n}_1, \tilde{n}_4 \} \). Then (5.5) takes the form
\[
\begin{align*}
G_{n,a}(\tilde{x}, \bar{y}) &= \left(-1\right)^{\nu(n+1)} h(\tilde{x}, \bar{x})h(\tilde{x}, \bar{x}) \sum_{\tilde{n}_1=\{\tilde{n}_1, \tilde{n}_4\}} r_1(\tilde{n}_1) r_3(\tilde{\xi}) \frac{f(\tilde{n}_1, \tilde{n}_4)f(f(\tilde{n}_1, \tilde{n}_4))}{f(\tilde{\xi}_1, \tilde{n}_4)f(\tilde{\xi}_1, \tilde{n}_4)} K_{a-n}(\bar{y}|\tilde{n}_4 + c) \\
&\quad \times f(\tilde{n}_1, \tilde{n}_4)f(f(\tilde{n}_1, \tilde{n}_4))g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)
\end{align*}
\]
\( \times \mathbb{B}_{0,b-a}(\tilde{n}_1; \tilde{\xi}_1) K_{a-n}(\tilde{\xi}_1|\tilde{n}_1 + c) K_{n}(\tilde{\xi}_1|\tilde{n}_1)f(\tilde{n}_1, \tilde{n}_4) = \frac{(1)^n K_n(\tilde{\xi}_1, \tilde{n}_1)|\tilde{n}_1 + c)}{f(\tilde{\xi}_1, \tilde{n}_4)f(\tilde{\xi}_1, \tilde{n}_4)}.
\] (5.7)

We see that the sum over partitions \( \tilde{n}_1 = \{ \tilde{n}_1, \tilde{n}_4 \} \) involves the terms in the last line only. This sum can be computed via lemma C.1. Using (C.1) we find
\[
\sum_{\tilde{n}_1=\{\tilde{n}_1, \tilde{n}_4\}} K_{a-n}(\tilde{n}_1|\tilde{\xi}_1 + c) K_{n}(\tilde{\xi}_1|\tilde{n}_1 + c)f(\tilde{n}_1, \tilde{n}_4)
= \frac{(1)^n K_n(\tilde{\xi}_1, \tilde{n}_1)|\tilde{n}_1 + c)}{f(\tilde{\xi}_1, \tilde{n}_4)f(\tilde{\xi}_1, \tilde{n}_4)}.
\] (5.8)

Thus, (5.7) takes the form
\[
\begin{align*}
G_{n,a}(\tilde{x}, \bar{y}) &= \left(-1\right)^{\nu(n+1)} h(\tilde{x}, \bar{x})h(\tilde{x}, \bar{x}) \sum_{\tilde{n}_1=\{\tilde{n}_1, \tilde{n}_4\}} r_1(\tilde{n}_1) r_3(\tilde{\xi}) K_{a-n}(\bar{y}|\tilde{n}_4 + c) \\
&\quad \times f(\tilde{n}_1, \tilde{n}_4)f(f(\tilde{n}_1, \tilde{n}_4))g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)
\end{align*}
\]
\( \times K_{a-n}(\tilde{\xi}_1, \tilde{n}_1) \mathbb{B}_{0,b-a}(\tilde{n}_1; \tilde{\xi}_1). \) (5.9)

Now we define \( \tilde{\xi}_0 = \{ \tilde{\xi}_1, \tilde{\xi}_4 \} \). Then (5.9) can be written as
\[
\begin{align*}
G_{n,a}(\tilde{x}, \bar{y}) &= \left(-1\right)^{\nu(n+1)} h(\tilde{x}, \bar{x})h(\tilde{x}, \bar{x}) \sum_{\tilde{n}_1=\{\tilde{n}_1, \tilde{n}_4\}} r_1(\tilde{n}_1) r_3(\tilde{\xi}) K_{a-n}(\bar{y}|\tilde{n}_4 + c) K_{n}(\tilde{\xi}_0|\tilde{n}_1 + c) \\
&\quad \times f(\tilde{n}_1, \tilde{n}_4)f(f(\tilde{n}_1, \tilde{n}_4))g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)g(\tilde{\xi}_1, \tilde{n}_4)
\end{align*}
\]
\( \times \mathbb{B}_{0,b-a}(\tilde{n}_1; \tilde{\xi}_1) K_{a-n}(\tilde{\xi}_1|\tilde{n}_1 + c) K_{n}(\tilde{\xi}_0|\tilde{n}_1 + c) G_{n,a}(\tilde{x}, \bar{y})h(\tilde{x}, \bar{x}) = \frac{(1)^n K_n(\tilde{\xi}_0, \tilde{n}_1)|\tilde{n}_1 + c)}{f(\tilde{\xi}_1, \tilde{n}_4)f(\tilde{\xi}_1, \tilde{n}_4)\bar{y} g(\tilde{\xi}_0, \tilde{x})}. \) (5.10)

The sum over partitions \( \tilde{\xi}_0 = \{ \tilde{\xi}_1, \tilde{\xi}_4 \} \) involves the terms in the last line only. It can be computed via (C.13):
\[
\sum g(\hat{\xi}_{\text{II}}, \hat{\zeta}) g(\hat{\xi}_{\text{I}}, \hat{\eta}) h(\hat{\xi}_{\text{O}}, \hat{\eta}) \frac{(-1)^{n(a-n)}}{g(\hat{\xi}_{\text{O}}, \hat{\eta})} \sum g(\hat{\xi}_{\text{II}}, \hat{\zeta}) g(\hat{\xi}_{\text{I}}, \hat{\eta}) \frac{(-1)^{n(a-n)}}{h(\hat{\xi}_{\text{O}}, \hat{\eta})} g(\hat{\xi}_{\text{I}}, \hat{\eta}) g(\hat{\xi}_{\text{O}}, \hat{\eta}) y - c) = \frac{h(\hat{\xi}_{\text{I}}, \hat{\eta})}{h(\hat{\xi}_{\text{O}}, \hat{\eta})}.
\]

Thus, we arrive at
\[
G_{a,a}(\hat{\xi}_{\text{O}}, \hat{\eta}) = \frac{(-1)^{n(a-1)}}{2} h(\hat{\xi}_{\text{I}}, \hat{\eta}) h(\hat{\xi}_{\text{O}}, \hat{\eta}) \sum \frac{r_1(\hat{\eta}_0)r_3(\hat{\xi}_i)}{h(\hat{\xi}_{\text{O}}, \hat{\eta}) h(\hat{\xi}_{\text{I}}, \hat{\eta})} K_{a-a}(\hat{\xi}_{\text{O}} | \hat{\eta}_{\text{I}} + c) K_{a}(\hat{\xi}_{\text{I}} | \hat{\eta}_0) \times \frac{f(\hat{\eta}_{\text{I}}, \hat{\eta}_0)f(\hat{\eta}_{\text{I}}, \hat{\eta}_0)h(\hat{\eta}_{\text{I}}, \hat{\eta}_0)g(\hat{\xi}_{\text{II}}, \hat{\zeta}) g(\hat{\xi}_{\text{I}}, \hat{\eta}) g(\hat{\xi}_{\text{O}}, \hat{\eta}) h(\hat{\xi}_{\text{I}}, \hat{\eta}) h(\hat{\xi}_{\text{O}}, \hat{\eta})}{f(\hat{\eta}_{\text{I}}, \hat{\eta}_0)f(\hat{\xi}_{\text{II}}, \hat{\eta}_0)f(\hat{\xi}_{\text{I}}, \hat{\eta}_0)g(\hat{\xi}_{\text{II}}, \hat{\zeta}) g(\hat{\xi}_{\text{I}}, \hat{\eta}) g(\hat{\xi}_{\text{O}}, \hat{\eta})} \mathbb{P}_{0,b-a}(\varnothing; \hat{\xi}_{\text{II}}). \quad (5.12)
\]

Finally, after a relabeling of the subsets \( \hat{\eta}_0 \to \hat{\xi}_1, \hat{\eta}_1 \to \hat{\eta}_0, \hat{\eta}_0 \to \hat{\eta}_1, \hat{\xi}_2 \to \hat{\xi}_0, \hat{\xi}_0 \to \hat{\xi}_2, \hat{\xi}_{\text{II}} \to \hat{\xi}_{\text{II}}, \) we recast \( (5.12) \) as follows:
\[
G_{a,a}(\hat{\xi}_{\text{O}}, \hat{\eta}) = \frac{(-1)^{n(a-1)}}{2} h(\hat{\xi}_{\text{I}}, \hat{\eta}) h(\hat{\xi}_{\text{O}}, \hat{\eta}) \sum \frac{r_1(\hat{\eta}_0)r_3(\hat{\xi}_i)}{h(\hat{\xi}_{\text{O}}, \hat{\eta}) h(\hat{\xi}_{\text{I}}, \hat{\eta})} K_{a-a}(\hat{\xi}_{\text{O}} | \hat{\eta}_{\text{I}} + c) K_{a}(\hat{\xi}_{\text{I}} | \hat{\eta}_0) \times \frac{f(\hat{\eta}_{\text{I}}, \hat{\eta}_0)f(\hat{\eta}_{\text{I}}, \hat{\eta}_0)h(\hat{\eta}_{\text{I}}, \hat{\eta}_0)g(\hat{\xi}_{\text{II}}, \hat{\zeta}) g(\hat{\xi}_{\text{I}}, \hat{\eta}) g(\hat{\xi}_{\text{O}}, \hat{\eta}) h(\hat{\xi}_{\text{I}}, \hat{\eta}) h(\hat{\xi}_{\text{O}}, \hat{\eta})}{f(\hat{\eta}_{\text{I}}, \hat{\eta}_0)f(\hat{\xi}_{\text{II}}, \hat{\eta}_0)f(\hat{\xi}_{\text{I}}, \hat{\eta}_0)g(\hat{\xi}_{\text{II}}, \hat{\zeta}) g(\hat{\xi}_{\text{I}}, \hat{\eta}) g(\hat{\xi}_{\text{O}}, \hat{\eta})} \mathbb{P}_{0,b-a}(\varnothing; \hat{\xi}_{\text{II}}). \quad (5.13)
\]

We recall that the cardinalities of the subsets are
\[
\#\hat{\eta}_1 = a, \quad \#\hat{\eta}_2 = n, \quad \#\hat{\eta}_{\text{II}} = a - n, \quad \#\hat{\xi}_1 = n, \quad \#\hat{\xi}_2 = a, \quad \#\hat{\xi}_{\text{II}} = b - n. \quad (5.14)
\]

**Remark 5.1.** Strictly speaking, the sets \( \hat{\eta} \) and \( \hat{\xi} \) in equation \( (5.13) \) should be understood as
\[
\hat{\eta} = \{ \hat{x} + \epsilon_1, \hat{y} + \epsilon_1, \hat{u} + \epsilon_1 \}, \quad \hat{\xi} = \{ \hat{x} + \epsilon_2, \hat{y} + \epsilon_2, \hat{u} + \epsilon_2 \}, \quad \text{at } \epsilon_k \to 0, \quad k = 1, 2. \quad (5.15)
\]

The point is that the individual factors in \( (5.13) \) may have singularities if we set \( \epsilon_k = 0 \). For instance, if \( \hat{\xi} \cap \hat{\eta} \neq \varnothing \), then the DWPF \( K_{a}(\hat{\xi} | \hat{\eta}) \) is singular. However, these poles are compensated by the product \( f(\hat{\xi}, \hat{\eta})^{-1} \). Therefore, for the appropriate evaluation of the limit we should have \( \epsilon_k \neq 0 \). In order to lighten the formulas we do not write these auxiliary parameters \( \epsilon_k \) explicitly, but one has to keep them in mind when doing the calculations.

**5.2. Successive action of** \( T_{32}(\hat{x}) T_{31}(\hat{x}) T_{21}(\hat{y}) \)**

Now let \( \#\hat{\xi} = b - n \). Then we define
\[
\mathbb{P}_{32}(\hat{x}) T_{31}(\hat{x}) T_{21}(\hat{y}) = \mathbb{P}_{32}(\hat{x}) G_{a,a}(\hat{x}, \hat{y}) = H_{a,a,b}(\hat{x}, \hat{y}, \hat{z}, \hat{\xi}) | 0 \). \quad (5.16)
\]
In order to act with $T_{\alpha\beta}(\vec{z})$ onto $G_{n,a}(\vec{x}, \vec{y})$ we should use (3.10). Let us denote the union \{\vec{x}, \vec{y}\} as $\vec{a}^C$ (as it will be in the case of the scalar product). Then we obtain

$$H_{n,a,b}(\vec{x}, \vec{y}, \vec{z}) = (-1)^{a+b} \frac{\prod_{i=1}^{n} (\vec{a}_i^{\alpha} + 1) - \prod_{i=1}^{a} (\vec{a}_i^{\alpha} + 1)}{\prod_{i=1}^{n} (\vec{a}_i^{\alpha} + 1) - \prod_{i=1}^{a} (\vec{a}_i^{\alpha} + 1)} h(\vec{v}, \vec{a})h(\vec{a}, \vec{a}^{c}) \sum r(\vec{a}_1) \cdots r(\vec{a}_n) \frac{h(\vec{a}_1^{\alpha}, \vec{a}^{c})}{h(\vec{a}_1^{\alpha}, \vec{a}^{c})}$$

$$\times K_{a-n}(\vec{y}|\vec{u} + c)K_{a}(\vec{z}|\vec{y}) \frac{f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})}{f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})} h(\vec{z}, \vec{y})$$

$$\times \frac{g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})}{h(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})}.$$

(5.17)

Here the partitions of the set $\vec{\eta}$ remain the same as in (5.13). The partitions of the remaining variables are organized as follows. We first have the partitions of the set \{\vec{a}^C, \vec{v}\} = $\vec{\xi}$ \Rightarrow $\{\vec{\xi}, \vec{u}, \vec{u}\}$. Then we combine $\{\vec{z}, \vec{y}\}$ and obtain additional partitions $\{\vec{z}, \vec{y}\}$ \Rightarrow $\{\vec{z}, \vec{y}\}$ with the restriction $\#\vec{z} = \#\vec{y} = b - n$.

We should substitute $\vec{\xi}_{\vec{u}} = \{\vec{\xi}, \vec{u}, \vec{u}\}$ into (5.17). Then, using

$$\frac{g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})}{f(\vec{z}, \vec{y})f(\vec{z}, \vec{y})f(\vec{z}, \vec{y})f(\vec{z}, \vec{y})} = g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})$$

(5.18)

we arrive at

$$H_{n,a,b}(\vec{x}, \vec{y}, \vec{z}) = (-1)^{a+b} \frac{\prod_{i=1}^{n} (\vec{a}_i^{\alpha} + 1) - \prod_{i=1}^{a} (\vec{a}_i^{\alpha} + 1)}{\prod_{i=1}^{n} (\vec{a}_i^{\alpha} + 1) - \prod_{i=1}^{a} (\vec{a}_i^{\alpha} + 1)} h(\vec{v}, \vec{a})h(\vec{a}, \vec{a}^{c}) \sum r(\vec{a}_1) \cdots r(\vec{a}_n) \frac{h(\vec{a}_1^{\alpha}, \vec{a}^{c})}{h(\vec{a}_1^{\alpha}, \vec{a}^{c})}$$

$$\times K_{a-n}(\vec{y}|\vec{u} + c)K_{a}(\vec{z}|\vec{y}) \frac{f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})}{f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})} h(\vec{z}, \vec{y})$$

$$\times \frac{g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})}{h(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})}.$$

(5.19)

Here we have denoted $\vec{\xi}$ the union $\{\vec{z}, \vec{a}^C, \vec{v}\}$. This set is divided into four subsets $\vec{\xi} \Rightarrow \{\vec{\xi}, \vec{\xi}_{\vec{u}}, \vec{\xi}_{\vec{u}}, \vec{\xi}_{\vec{u}}\}$ with the cardinalities $\#\vec{\xi} = \#\vec{\xi}_{\vec{u}} = b - n$, $\#\vec{\xi}_{\vec{u}} = n$, and $\#\vec{\xi}_{\vec{u}} = a$.

Let $\vec{\xi}_0 = \{\vec{\xi}, \vec{\xi}_{\vec{u}}\}$. Then

$$H_{n,a,b}(\vec{x}, \vec{y}, \vec{z}) = (-1)^{a+b+n(b-1)+2a+n-1} h(\vec{v}, \vec{a})h(\vec{a}, \vec{a}^{c}) \sum r(\vec{a}_1) \cdots r(\vec{a}_n) \frac{h(\vec{a}_1^{\alpha}, \vec{a}^{c})}{h(\vec{a}_1^{\alpha}, \vec{a}^{c})} K_{a-n}(\vec{y}|\vec{u} + c)$$

$$\times K_{a}(\vec{z}|\vec{y}) \frac{f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})}{f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})f(\vec{y}^{\alpha}, \vec{y})} h(\vec{z}, \vec{y})$$

$$\times \frac{g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})}{h(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})}.$$

(5.20)

The sum over partitions $\vec{\xi}_0 \Rightarrow \{\vec{\xi}, \vec{\xi}_{\vec{u}}\}$ involves the terms in the last line only. It can be computed via (C.13):

$$\sum \frac{g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})}{h(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})g(\vec{z}, \vec{y})} = \frac{(-1)^{b-n}}{h(\vec{z}, \vec{y})} \sum g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z})g(\vec{z}, \vec{z}) = \frac{h(\vec{z}, \vec{y})}{h(\vec{z}, \vec{y})}.$$

(5.21)
Substituting this into (5.20) we find

\[
H_{n,a,b}(\xi, \eta, \bar{\xi}, \bar{\eta}) = (-1)^{a+n(b+1)+\frac{b(b+1)}{2}} h(\bar{\eta}, a^C) h(\bar{u}^C, a^C) \sum r_1(\bar{\eta}_1) r_3(\bar{\xi}_0) \frac{H(\bar{\xi}_0, a^C)}{H(\bar{\xi}_0, a^C)}
\times f(\bar{\eta}_0, \bar{\eta}_1) h(\bar{\eta}_0, \bar{\eta}_1) h(\bar{\xi}_0, \bar{\xi}_1) h(\bar{\xi}_1, \bar{\xi}_0) g(\xi_0, \xi_1) g(\xi_1, \xi_0) h(\bar{\xi}_0, \bar{\eta}_1) g(\bar{\xi}_0, \bar{\eta}_1) h(\bar{\xi}_0, \bar{\eta}_1) g(\bar{\xi}_0, \bar{\eta}_1)
\times K_{a-n}(\bar{\eta}_0) K_{a-\bar{\eta}_0}(\bar{\eta}_0 + c) K_{a}(\bar{\xi}_0 | \bar{\xi}_0).
\]

(5.22)

Finally, relabeling \(\bar{\xi}_0 \rightarrow \bar{\xi}_1\) and \(\bar{\xi}_1 \rightarrow \bar{\xi}_0\) we arrive at

\[
H_{n,a,b}(\xi, \eta, \bar{\xi}, \bar{\eta}) = (-1)^{a+n(b+1)+\frac{b(b+1)}{2}} h(\bar{\eta}, a^C) h(\bar{u}^C, a^C) \sum r_1(\bar{\eta}_1) r_3(\bar{\xi}_0) \frac{H(\bar{\xi}_0, a^C)}{H(\bar{\xi}_0, a^C)}
\times f(\bar{\xi}_0, \bar{\xi}_1) h(\bar{\xi}_0, \bar{\xi}_1) h(\bar{\eta}_0, \bar{\eta}_1) h(\bar{\eta}_0, \bar{\eta}_1) g(\xi_0, \xi_1) g(\xi_1, \xi_0) h(\bar{\xi}_0, \bar{\eta}_1) g(\bar{\xi}_0, \bar{\eta}_1) h(\bar{\xi}_0, \bar{\eta}_1) g(\bar{\xi}_0, \bar{\eta}_1)
\times h(\bar{\xi}_0, \bar{\eta}_1) f(\bar{\xi}_0, \bar{\eta}_1) h(\bar{\xi}_0, \bar{\eta}_1) h(\bar{\xi}_0, \bar{\eta}_1) g(\xi_0, \xi_1) g(\xi_1, \xi_0) h(\bar{\xi}_0, \bar{\eta}_1) g(\bar{\xi}_0, \bar{\eta}_1) h(\bar{\xi}_0, \bar{\eta}_1) g(\bar{\xi}_0, \bar{\eta}_1)
\]

(5.23)

Recall that in this formula \(\bar{\eta} = \{\bar{\xi}, \bar{\eta}, \bar{u}\}\), \(\bar{\xi} = \{\bar{\eta}, \bar{\xi}, \bar{u}^C\}\), and we denote \(\bar{u}^C = \{\bar{\xi}, \bar{\eta}\}\). The sum is taken over partitions \(\bar{\eta} \Rightarrow \{\bar{\eta}_1, \bar{\eta}_2, \bar{\eta}_3\}\) and \(\bar{\xi} \Rightarrow \{\bar{\xi}_1, \bar{\xi}_2, \bar{\xi}_3\}\). We also set \(\bar{\eta}_0 = \{\bar{\eta}_1, \bar{\eta}_2\}\). The cardinalities of the subsets are

\[
\#\bar{\eta}_1 = a, \quad \#\bar{\eta}_2 = n, \quad \#\bar{\eta}_3 = a - n,
\]

\[
\#\bar{\xi}_1 = b, \quad \#\bar{\xi}_2 = a, \quad \#\bar{\xi}_3 = b - n.
\]

(5.24)

6. Highest coefficients

Equation (5.23) allows us to obtain an explicit representation for the scalar product of Bethe vectors. Using (3.5) we find

\[
S_{a,b} = \frac{(-1)^{\frac{a+b}{2}}}{f(\bar{u}^C, a^C) \sum g(\bar{v}_1^C, \bar{u}_1^C) f(\bar{u}_1^C, \bar{u}_1^C) f(\bar{v}_1^C, \bar{v}_1^C)}
\times g(\bar{v}_1^C, \bar{v}_1^C) h(\bar{u}_1^C, \bar{u}_1^C) H_{n,a,b}(\bar{u}_1^C, \bar{u}_1^C, \bar{v}_1^C).
\]

(6.1)

The sum is taken over partitions \(\bar{u}^C \Rightarrow \{\bar{u}_1^C, \bar{u}_1^C\}\) and \(\bar{v}^C \Rightarrow \{\bar{v}_1^C, \bar{v}_1^C\}\), where \(\#\bar{v}_1^C = \#\bar{u}_1^C = n\), and \(n = 0, 1, \ldots, \min(a, b)\). The function \(H_{n,a,b}(\bar{u}_1^C, \bar{u}_1^C, \bar{v}_1^C)\) itself is given as a sum over partitions described in (5.23). That is, the union \(\{\bar{u}^C, \bar{v}^C\}\) is divided into three subsets and the union \(\{\bar{v}_1^C, \bar{u}_1^C, \bar{v}_1^C\}\) is also divided into three subsets. Although the resulting formula is explicit, it is inconvenient for later use. Therefore, we will try to simplify it. To do this, we introduce a new function.

Definition 6.1. Let \(\bar{x}, \bar{y}, \bar{i}, \bar{s}, \bar{t}\) be five sets of generic complex numbers with cardinalities \(\#\bar{x} = n, \#\bar{y} = m\), and \(\#\bar{t} = n + m\). The cardinalities of the sets \(\bar{i}\) and \(\bar{s}\) are not fixed. Define a function

\[
J_{n,m}(\bar{x}, \bar{y}; \bar{i}, \bar{s}; \bar{t}) = \Delta_{n+m}(\bar{t}) \Delta_n(\bar{t}) \Delta_m(\bar{t}) \det_{\bar{t}}.
\]

(6.2)
where
\[
J_{jk} = \frac{g(\beta_j, x_k)}{h(\beta_j, x_k)}, \quad k = 1, \ldots, n;
\]
\[
J_{j,k+n} = \frac{g(\beta_j, \tilde{y})}{h(\beta_j, \tilde{s})}, \quad k = 1, \ldots, m; \tag{6.3}
\]

Developing the determinant in (6.2) with respect to the first \(n\) columns (see appendix C for more details) we obtain a presentation of \(J_{n,m}\) as a sum over partitions of the set \(\bar{\beta}\):
\[
J_{n,m}(\bar{x}; \bar{y}, \bar{s}; \bar{\beta}) = \sum K_{n}(\bar{x}, \bar{y}) \frac{g(\beta_{\bar{n}}, y) g(\beta_{\bar{s}}, \tilde{y})}{h(\beta_{\bar{n}}, \bar{x}) h(\beta_{\bar{s}}, \bar{s})}.
\]  
Here the sum is taken over partitions \(\bar{\beta} \Rightarrow \{\beta_1, \beta_2\}\), such that \(#\beta_1 = n \text{ and } #\beta_2 = m\).

6.1. First representation for the highest coefficient

Let us find the highest coefficient \(Z_{uu vv}^{ab BC BC}{\bar{v}^B, \bar{v}^B}\). We recall that up to the normalization factor \((f(v^C, \bar{a}^C) f(v^B, \bar{a}^B))^{-1}\) it is the rational coefficient of the product \(r_1(\bar{u}^B)r_2(\bar{v}^C)\) (see (4.10), (4.11)).

Obviously, for this we should set \(\bar{\eta}_1 = \bar{a}^B\) and \(\bar{\xi}_1 = \bar{v}^C\) in (5.23). However, \(\bar{\xi}_1 \subset \{\bar{v}^C, \bar{a}^C, \bar{v}^B\}\). Hence, one can have \(\bar{\xi}_1 = \bar{v}^C\) if and only if \(\bar{v}^B = \bar{v}^C\), and thus, \(\bar{v}^C = \emptyset\). But \(#\bar{v}^C = n + m\) in (6.1); therefore \(\bar{a}^C = \emptyset\) and \(n = 0\). Thus, (6.1) takes the form
\[
\frac{r_1(\bar{u}^B)r_2(\bar{v}^C)}{f(v^C, \bar{a}^C)f(\bar{v}^B, \bar{a}^B)} = \frac{(-1)^{a^B + b}}{f(v^C, \bar{a}^C)} H_{a, n, b} (\emptyset, \bar{a}^C, \bar{v}^C) \bigg|_{\bar{\eta}_1 = \bar{u}^B; \bar{\xi}_1 = \bar{v}^C}.
\]  
Substituting the conditions \(\bar{\eta}_1 = \bar{a}^B\) and \(\bar{\xi}_1 = \bar{v}^C\) into (5.23) we should take into account that \(#\bar{v}^C = n = 0\) (see (5.24)). Hence, \(\bar{\eta}_1 = \emptyset\), which implies \(\eta_{\bar{n}} = \bar{a}^C\). Thus, substituting these subsets into (5.23) we find
\[
r_1(\bar{u}^B)r_2(\bar{v}^C) Z_{a, b}(\bar{a}^B; \bar{v}^B, \bar{v}^B, \bar{v}^C) = (-1)^a h(\bar{v}^B, \bar{a}^C) h(\bar{u}^B, \bar{a}^C) r_1(\bar{u}^B)r_2(\bar{v}^C) \times K_{a}(\bar{a}^C, \bar{a}^C + c) \sum K_{\bar{a}}(\bar{a}^B) g(\xi_{\bar{n}}, \bar{a}^C) g(\xi_{\bar{a}}, \bar{v}^C) h(\xi_{\bar{n}}, \bar{a}^B)
\]  
where the sum is taken over partitions \(\{\bar{a}^C, \bar{v}^B\} = \bar{\xi} \Rightarrow \{\xi_{\bar{n}}, \xi_{\bar{a}}\}\) with \(#\xi_{\bar{n}} = a\) and \(#\xi_{\bar{a}} = b\). Due to (A.1) we conclude that \(K_{a}(\bar{a}^C, \bar{a}^C + c) = (-1)^a\), and we arrive at
\[
Z_{a, b}(\bar{a}^B; \bar{v}^B, \bar{v}^C) = h(\bar{v}^B, \bar{a}^B) h(\bar{a}^C, \bar{v}^C) \sum K_{a}(\xi_{\bar{n}}|\bar{a}^B) g(\xi_{\bar{n}}, \bar{v}^C) g(\xi_{\bar{a}}, \bar{v}^B) h(\xi_{\bar{n}}, \bar{a}^B)
\]  
Finally, using \(\{\bar{a}^C, \bar{v}^B\} = \bar{\xi}\) we recast (6.7) as follows:
\[
Z_{a, b}(\bar{a}^B; \bar{v}^B, \bar{v}^C) = \sum K_{a}(\xi_{\bar{a}}|\bar{a}^B) g(\xi_{\bar{a}}, \bar{v}^C) g(\xi_{\bar{n}}, \bar{v}^B) h(\xi_{\bar{a}}, \bar{a}^C)
\]  
\[
= h(\bar{v}^B, \bar{a}^B) h(\bar{a}^C, \bar{v}^B) \sum K_{a}(\xi_{\bar{n}}|\bar{a}^B) g(\xi_{\bar{n}}, \bar{v}^C) g(\xi_{\bar{a}}, \bar{v}^B) h(\xi_{\bar{n}}, \bar{a}^B)
\]  
\[
= h(\bar{v}^B, \bar{a}^B) h(\bar{a}^C, \bar{v}^B) \sum K_{a}(\xi_{\bar{n}}|\bar{a}^B) g(\xi_{\bar{n}}, \bar{v}^C) g(\xi_{\bar{a}}, \bar{v}^B) h(\xi_{\bar{n}}, \bar{a}^B)
\]  
\[
= h(\bar{v}^B, \bar{a}^B) h(\bar{a}^C, \bar{v}^B) \sum K_{a}(\xi_{\bar{n}}|\bar{a}^B) g(\xi_{\bar{n}}, \bar{v}^C) g(\xi_{\bar{a}}, \bar{v}^B) h(\xi_{\bar{n}}, \bar{a}^B)
\]  
\[
= h(\bar{v}^B, \bar{a}^B) h(\bar{a}^C, \bar{v}^B) \sum K_{a}(\xi_{\bar{n}}|\bar{a}^B) g(\xi_{\bar{n}}, \bar{v}^C) g(\xi_{\bar{a}}, \bar{v}^B) h(\xi_{\bar{n}}, \bar{a}^B)
\]
Comparing (6.8) and (6.4) we conclude that
\[ Z_{a,b}(\bar{u}^B; \bar{u}^C|\bar{v}^B; \bar{v}^C) = h(\bar{v}^B, \bar{u}^B) h(\bar{u}^C, \bar{u}^B) J_{a,b}(\bar{u}^B; \bar{v}^C|\bar{u}^B, \bar{u}^C). \] (6.9)
Thus, we have obtained an explicit representation for the highest coefficient
\[ Z_{a,b}(\bar{u}^B; \bar{u}^C|\bar{v}^B, \bar{v}^C) \] in terms of the determinant of the \((a + b) \times (a + b)\) matrix \(J_{a,b}\) (6.3).

### 6.2. Second highest coefficient

In order to obtain the second highest coefficient \(Z_{a,b}(\bar{u}^C; \bar{u}^B|\bar{v}^C; \bar{v}^B)\) it is enough to make the replacements \(\bar{a}^C \rightarrow \bar{a}^B\) and \(\bar{v}^C \rightarrow \bar{v}^B\) in (6.9). On the other hand, this coefficient should arise if we set \(\bar{h}_1 = \bar{u}^C\) and \(\bar{v}_1 = \bar{v}^B\) in (5.23). However, if we do so, then we do not obtain (6.9) with the replacements mentioned above. Instead, we obtain a much more sophisticated formula involving many sums over partitions. This 'break of symmetry' occurs because we use a specific representation (3.5) for the dual Bethe vector. If we used equation (3.4) for \(C_{a,b}^{\bar{u}^B; \bar{v}^B}\), then we would have an analog of (6.9) for \(Z_{a,b}(\bar{u}^C; \bar{u}^B|\bar{v}^C; \bar{v}^B)\); however, we would have a more complex formula for \(Z_{a,b}(\bar{u}^C; \bar{u}^B|\bar{v}^B; \bar{v}^C)\).

A 'complex' formula for the highest coefficient provides us with a very non-trivial identity for \(Z_{a,b}(\bar{u}^C; \bar{u}^B|\bar{v}^C, \bar{v}^B)\), which will be used later. In order to obtain this identity we first make several additional summations in (6.1). Let us rewrite this equation explicitly

\[ S_{a,b} = \frac{1}{f(\bar{v}^C, \bar{u}^C)} \sum (-1)^{a+n(b+1)} g(\bar{v}^C, \bar{u}^C) h(\bar{u}_1^C, \bar{u}_1^C) h(\bar{v}_1^C, \bar{u}_1^C) \]

\[ \times g(\bar{v}_1^C, \bar{v}_1^C) h(\bar{u}_1^C, \bar{u}_1^C) r_1(\bar{h}_1) r_1(\bar{h}_1) J_a(\bar{v}_1^C|\bar{h}_1^a) J_{a=n}(\bar{u}_1^C|\bar{h}_1^a) + c \]

\[ \times \frac{1}{g(\bar{v}_1^C, \bar{v}_1^C)} h(\bar{h}_1^a) h(\bar{h}_1^a) h(\bar{h}_1^a) h(\bar{h}_1^a) h(\bar{h}_1^a) J_{a=n}(\bar{h}_1^a) \]

\[ \times g(\bar{v}_1^C, \bar{v}_1^C) h(\bar{v}_1^C, \bar{v}_1^C) g(\bar{v}_1^C, \bar{v}_1^C) g(\bar{v}_1^C, \bar{v}_1^C) \]

\[ \times \frac{1}{g(\bar{v}_1^C, \bar{v}_1^C)} g(\bar{v}_1^C, \bar{v}_1^C) J_{a=n}(\bar{h}_1^a). \] (6.10)

The sum over partitions into subsets \(\bar{h}_1^a\) and \(\bar{h}_1^a\), as well as the sum over partitions \(\bar{a}^C \rightarrow \{a_1^C, a_1^C\}\), can be computed in terms of the function \(J(6.2)\). Let \(\xi_0 = \{\xi_0, \xi_0\}\). Then

\[ \sum_{\xi_0 \rightarrow \{\xi_0, \xi_0\}} K_a(\xi_0|\eta_0) g(\xi_0, \xi_0) g(\xi_0, \xi_0) g(\xi_0, \xi_0) \]

\[ = (-1)^{a+b+c} g(\bar{v}_1^C, \bar{v}_1^C) h(\bar{v}_1^C, \bar{v}_1^C) \]

\[ \times J_{a=n}(\bar{h}_1^a, \bar{h}_1^a). \] (6.11)

Similarly, one can verify that

\[ \sum_{\bar{a}^C \rightarrow \{\bar{a}^C, \bar{a}^C\}} h(\bar{v}_1^C, \bar{v}_1^C) h(\bar{v}_1^C, \bar{v}_1^C) h(\bar{u}_1^C, \bar{u}_1^C) \]

\[ = (-1)^{a+b+c} h(\bar{v}_1^C, \bar{v}_1^C) J_{a-n}(\bar{h}_1^a, \bar{h}_1^a, \bar{h}_1^a, \bar{h}_1^a). \] (6.12)
Substituting these results into (10.6) we find
\[ S_{a,b} = \frac{h(\tilde{v}^B, \tilde{u}^C)h(\tilde{a}^C, \tilde{u}^C)}{f(\tilde{v}^C, \tilde{a}^C)} \sum (-1)^{b+n(a+b+1)} \eta_i \eta_j \sum_{J_{a,b-n}} (\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \]

\[ \times J_{a,n-\eta_i}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \eta_i \eta_j \hat{r}_3(\tilde{\xi}_I) J_{n-b-n}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \]

\[ \times h(\tilde{v}_I^C, \tilde{u}_I^C; \tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) h(\tilde{\xi}_I, \tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C). \]  

(6.13)

The sum is taken over partitions:

1. \( \tilde{v}^C \) \( \Rightarrow \{ \tilde{v}_I^C \} \) with \# \( \tilde{v}_I^C = n \) and \# \( \tilde{u}_I^C = b - n \);
2. \( \{ \tilde{a}^C, \tilde{u}^B \} = \tilde{\xi} \) \( \Rightarrow \{ \tilde{\xi}_I \} \) with \# \( \tilde{\xi}_I = \) \( a \) and \# \( \tilde{\eta}_I = \) \( b \) - \( n \);
3. \( \{ \tilde{a}^C, \tilde{b}^B \} = \tilde{\eta} \) \( \Rightarrow \{ \tilde{\eta}_I \} \) and \( \tilde{\eta}_I \) \( \Rightarrow \{ \tilde{\eta}_I \} \) with \# \( \tilde{\eta}_I = \) \( a \), \# \( \tilde{\eta}_I = \) \( n \), and \# \( \tilde{\eta}_I = \) \( a \) - \( n \).

In all these partitions \( n = 0, 1, \ldots, \min(a,b) \).

Due to proposition 4.1 the function \( r_3 \) depends on variables of the \( \nu \)-type only. Hence, \( \tilde{\xi}_I \cap \tilde{a}^C = \emptyset \), that is, \( \tilde{a}^C \subset \tilde{\xi}_I \). Therefore, we can set \( \tilde{\xi}_I = \{ \tilde{a}^C, \tilde{\xi}_I \} \). Substituting this into (6.13) we obtain

\[ S_{a,b} = \frac{h(\tilde{v}^B, \tilde{u}^C)}{f(\tilde{v}^C, \tilde{a}^C)} \sum (-1)^{b+n(a+b+1)} \eta_i \eta_j \sum_{J_{a,b-n}} (\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \]

\[ \times J_{a,n-\eta_i}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \eta_i \eta_j \hat{r}_3(\tilde{\xi}_I) J_{n-b-n}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \]

\[ \times h(\tilde{v}_I^C, \tilde{u}_I^C; \tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) h(\tilde{\xi}_I, \tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C). \]  

(6.14)

In this formula \( \xi = \{ \tilde{v}_I^C, \tilde{b}^B \} \), \# \( \xi = b \) and \# \( \tilde{\xi}_I = b - n \). All other subsets are the same as in (6.13).

Now everything is ready for formulating the second representation for the highest coefficient \( Z_{a,b}(\tilde{a}^C, \tilde{b}^B; \tilde{v}^C, \tilde{b}^B) \). For this we set \( \tilde{\eta}_I = \tilde{a}^C \) and \( \tilde{\xi}_I = \tilde{b}^B \). Then automatically \( \tilde{\eta}_I = \tilde{a}^C, \tilde{\xi}_I = \tilde{b}^B \) and we also can set \( \tilde{\eta}_I = \tilde{b}^B \), \( \tilde{\eta}_I = \tilde{b}^B \). Substituting this into (6.14) and keeping in mind remark 5.1 we obtain

\[ Z_{a,b}(\tilde{a}^C; \tilde{b}^B) = \frac{h(\tilde{v}^B, \tilde{u}^B)}{f(\tilde{v}^C, \tilde{a}^C)} \sum (-1)^{n(a+b+1)} \eta_i \eta_j \sum_{J_{a,b-n}} (\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \]

\[ \times J_{a,n-\eta_i}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \eta_i \eta_j \hat{r}_3(\tilde{\xi}_I) J_{n-b-n}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C) \]

\[ \times h(\tilde{v}_I^C, \tilde{u}_I^C; \tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) h(\tilde{\xi}_I, \tilde{\eta}_I; \tilde{v}_I^C, \tilde{u}_I^C, \tilde{a}_I^C). \]  

(6.15)

Using (D.3) and (D.4) we find

\[ \lim_{\tilde{\eta}_I \rightarrow \tilde{a}^C} J_{a,b-n}(\tilde{\eta}_I; \tilde{v}_I^C, \tilde{a}_I^C) \]

\[ \times \lim_{\tilde{\xi}_I \rightarrow \tilde{b}^B} g(\tilde{v}_I^C, \tilde{u}_I^C) g(\tilde{v}_I^C, \tilde{u}_I^C) = (-1)^{b+n} g(\tilde{v}_I^C, \tilde{a}_I^C), \]  

(6.16)
and we thus arrive at
\[ Z_{a,b}(\tilde{u}^C; \tilde{v}^B, \tilde{v}^C) = f(\tilde{u}^B, \tilde{v}^C)f(\tilde{v}^B, \tilde{u}^B) \sum (-1)^{n(a+b)} \]
\[ \times J_{3n-n}(\tilde{u}^B_0, \tilde{v}^C_0) \tilde{u}^C + c; \tilde{u}^B + c(\tilde{u}^C - c)h(\tilde{u}^B_1, \tilde{u}^B_1) \]
\[ \times f(\tilde{v}^C_1, \tilde{u}^C_1)h(\tilde{v}^C_1, \tilde{v}^B_1)g(\tilde{v}^C_1, \tilde{v}^C_1)g(\tilde{v}^B_1, \tilde{v}^B_1) \]
\[ h(\tilde{v}^B_1, \tilde{v}^B_1)g(\tilde{u}^B_1, \tilde{u}^B_1). \]  
(6.17)

Here the sum is taken over partitions \( \tilde{u}^B = \{ \tilde{u}^B_1, \tilde{u}^B_0 \} \) and \( \tilde{v}^C = \{ \tilde{v}^C_1, \tilde{v}^C_0 \} \) with \( \#\tilde{v}^C_1 = n \) and \( \#\tilde{u}^B_0 = a - n \).

This is the second representation for the highest coefficient discussed above. It will play the key role below; therefore we formulate it as a proposition.

**Proposition 6.1.** For arbitrary sets of complex numbers \( \bar{t}, \bar{x}, \bar{s}, \bar{y} \) with cardinalities \( \#\bar{t} = \#\bar{x} = a \) and \( \#\bar{s} = \#\bar{y} = b \) the following identity holds:

\[ Z_{a,b}(\bar{t}, \bar{x}|\bar{s}, \bar{y}) = f(\bar{x}, \bar{t})f(\bar{y}, \bar{x}) \sum (-1)^{n(a+b)} J_{3n-n}(\bar{s}_1; \bar{t}_1 + c; \bar{x} + c|\bar{t} - c) \]
\[ \times h(\bar{s}_1, \bar{x}_1)f(\bar{x}_1, \bar{x}_1) \frac{f(\bar{t}_1, \bar{t}_1)h(\bar{s}_1, \bar{t}_1)g(\bar{s}_1, \bar{x}_1)g(\bar{t}_1, \bar{x}_1)}{h(\bar{y}, \bar{x}_1)g(\bar{s}_1, \bar{t}_1)}. \]  
(6.18)

Here \( \ell_2 = \#s_1, m_1 = \#s_1 \). The sum is taken over partitions
\[ \bar{x} \Rightarrow \{ \bar{x}_1, \bar{x}_2 \}, \quad \bar{s} \Rightarrow \{ \bar{s}_1, \bar{s}_2 \}, \]
with a restriction \( \#\bar{s}_1 = \#\bar{s}_1 \) (which is equivalent to \( \ell_2 + n_1 = a \)).

**Proof.** Setting in (6.17) \( \bar{u}^C = \bar{t}, \bar{u}^B = \bar{x}, \bar{v}^C = \bar{s}, \) and \( \bar{v}^B = \bar{y} \) we obtain (6.18).

### 6.3. General formula for the scalar product

Now we turn back to equation (6.14). To proceed further we should specify all the subsets.

Let \( \bar{\psi}^C = \{ \bar{\psi}_1^C, \bar{\psi}_0^C, \bar{\psi}_1^C \} \) and \( \bar{\psi}^B = \{ \bar{\psi}_1^B, \bar{\psi}_0^B \} \). We set
\[ \bar{\psi}_1^C = \bar{\psi}_1^C, \quad \bar{\psi}_0^C = \{ \bar{\psi}_1^C, \bar{\psi}_0^C \}, \]
\[ \bar{\psi}_1^B = \{ \bar{\psi}_1^B, \bar{\psi}_0^B \}, \quad \bar{\psi}_0^B = \{ \bar{\psi}_1^B, \bar{\psi}_0^B \}. \]

It is easy to see that the following conditions for the cardinalities hold:
\[ n_{ii} = n, \quad n_i + n_{ii} = b - n, \quad m_i + m_{ii} = b, \quad m_0 = n. \]

Let also
\[ \bar{u}^C = \{ \bar{u}_1^C, \bar{u}_0^C \}, \quad \bar{u}_0^C = \{ \bar{u}_1^C, \bar{u}_0^C \}, \]
\[ \bar{u}_1^B = \{ \bar{u}_1^B, \bar{u}_0^B \}, \quad \bar{u}_0^B = \{ \bar{u}_1^B, \bar{u}_0^B \}, \]
\[ \bar{v}_1^C = \{ \bar{v}_1^C, \bar{v}_0^C \}, \quad \bar{v}_0^C = \{ \bar{v}_1^C, \bar{v}_0^C \}. \]

We have the following conditions for the cardinalities:
\[ k_1 + k_0 = \ell_1 + \ell_0 = a, \quad k_i = \ell_i, \quad k_0 = \ell_0. \]  
(6.23)
Observe that we do not fix a distribution of the parameters \( \tilde{a}^C \) and \( \tilde{a}^B \) among the subsets \( \tilde{\eta}_i \) and \( \tilde{\eta}_II \). It is important, however, that \( \# \tilde{\eta}_I = n = n_{ii} \) and \( \# \tilde{\eta}_II = a - n \).

Using (D.3) and (D.4) we obtain

\[
J_{a,b;\nu}(\tilde{a}^C; \tilde{\eta}_I \cup \{ \tilde{a}^C \}; \tilde{\eta}_II \cup \{ \tilde{a}^B \}) = (-1)^{b - a} g(\psi^\nu, \tilde{a}^C) g(\psi^\nu, \tilde{a}^B) h(\psi^\nu, \tilde{a}^C) g(\psi^\nu, \tilde{a}^B) h(\psi^\nu, \tilde{a}^C) \times J_{I,n}(a^B; \psi^C|[a^B]; a^B|\{a^C, \psi^B\}).
\]

Due to (6.9) this function reduces to the highest coefficient

\[
J_{I,n}(a^B; \psi^C|[a^B]; a^B|\{a^C, \psi^B\}) = \frac{Z_{I,n}(a^B, a^C|\{a^B, a^C\}; \psi^C)}{h(a^B, a^B) h(\psi^C, \psi^B)}.
\]

Now we substitute (6.24) and (6.25) into (6.14). We also write explicitly the products \( g(\psi^C, \psi^C), g(\tilde{\eta}_I, \tilde{\eta}_I) \), and combine \( \{\tilde{\psi}_i^C, \tilde{\psi}_i^C\} = \psi^C \). Then we have

\[
S_{a,b} = \frac{1}{f(\psi^C, \psi^C)} \sum \left[r(\tilde{a}_0^B) r(\tilde{a}_0^C) r(\tilde{a}_0^B) r(\tilde{a}_0^C) \times J_{I,n}(a^B; \psi^C|[a^C, \psi^C]; a^B|\{a^C, \psi^B\})(\tilde{a}_0^B, \tilde{a}_0^B|\{a^C, \psi^C\}) \times (-1)^{n(a + b - n)} h(\tilde{\eta}_I, \tilde{\eta}_II) f(\tilde{\eta}_I, \tilde{\eta}_II) f(\tilde{\eta}_I, \tilde{\eta}_II)\right].
\]

Here the sum is organized as follows. First, we have partitions

\[
\tilde{a}^C \Rightarrow \{\tilde{a}_0^C, \tilde{a}_0^C\}, \quad \tilde{a}^B \Rightarrow \{\tilde{a}_0^B, \tilde{a}_0^B\}, \quad \# \tilde{a}_0^C = \# \tilde{a}_0^B = c, \quad \# \tilde{\eta}_I = n = n_{ii}, \quad \# \tilde{\eta}_II = a - n.
\]

After this we have two additional partitions: the set \( \psi^C \) is divided into subsets \( \psi_i^C \) and \( \psi_i^C \); the union of the subsets \( \tilde{\eta}_I = \{\tilde{\psi}_i^C, \tilde{\psi}_i^C\} \) is divided into subsets \( \tilde{\eta}_II \) and \( \tilde{\eta}_II \) (see the terms in braces in (6.26)). At the same time, we have one restriction for the cardinalities \( \# \psi_i^C = \# \psi_i^C = n_{ii} \).

Let us separate this additional sum over partitions in the braces of (6.26)

\[
\mathcal{F}(\tilde{a}^C; \tilde{\eta}_I; \psi_i^C, \psi_i^C) = \sum_{\tilde{\eta}_II} J_{a,b;\nu}(\tilde{a}^C; \tilde{\eta}_II, \psi_i^C; \tilde{\eta}_II, \psi_i^C) \times (-1)^{n(a + b - n)} h(\tilde{\eta}_II, \tilde{\eta}_II) f(\tilde{\eta}_II, \tilde{\eta}_II) f(\tilde{\eta}_II, \tilde{\eta}_II).
\]

Comparing (6.28) with (6.18) one can see that they coincide after appropriate identification of the subsets and their cardinalities. That is, (6.28) turns into (6.18) under the replacements \( b \rightarrow b \), \( \tilde{a}^C \rightarrow \tilde{a}^C \), \( \tilde{\eta}_I \rightarrow \tilde{\eta}_I \), \( \psi_i^C \rightarrow \psi_i^C \), and \( \psi_i^C \rightarrow \psi_i^C \). Thus, due to proposition 6.1 we obtain

\[
\mathcal{F}(\tilde{a}^C; \tilde{\eta}_I; \psi_i^C, \psi_i^C) = \frac{Z_{a,b;\nu}(\tilde{a}^C; \tilde{\eta}_II, \psi_i^C, \psi_i^C)}{f(\tilde{\eta}_II, \tilde{a}^C) f(\psi_i^C, \tilde{\eta}_II)}.
\]
Thus, substituting this into (6.26) we arrive at
\[ S_{a,b} = \frac{1}{f(v^c, v^\alpha)} \sum_{r_1} \langle \hat{u}^C, r_1 \rangle \langle \hat{a}^B, r_3 \rangle \langle \hat{v}^C, r_2 \rangle Z_{a,b,n}(\hat{u}^B, \hat{a}^C, v^C) \frac{f(\hat{u}^B, \hat{a}^C)}{f(v^B, v^\alpha)} \times f(\hat{u}^C, \hat{a}^C) g(\hat{v}^C, \hat{v}^C) g(\hat{v}^B, \hat{v}^B) \]
\[ \times \frac{Z_{a,b-n}(\hat{u}^C; \eta_0^C, \hat{v}^B, \hat{v}^B)}{f(\eta_0, \hat{a}^C) f(\hat{v}^B, \hat{v}^B)}. \]  
(6.30)

The next step is to simplify the ratio \( Z_{a,b-n}(\hat{u}^C; \eta_0^C, \hat{v}^B, \hat{v}^B) \). It can be done via (6.9), (6.3):
\[ \frac{Z_{a,b-n}(\hat{u}^C; \eta_0^C, \hat{v}^B, \hat{v}^B)}{f(\eta_0, \hat{a}^C) f(\hat{v}^B, \hat{v}^B)} = \frac{f(v^C, \hat{a}^C)}{f(\hat{v}^B, \hat{a}^C)} Z_{a-b-n}(\hat{u}^B, \hat{v}^B, \hat{v}^B). \]  
(6.31)

Substituting this into (6.30) we obtain
\[ S_{a,b} = \sum_{r_1} \langle \hat{u}^C, r_1 \rangle \langle \hat{a}^B, r_3 \rangle \langle \hat{v}^C, r_2 \rangle \frac{f(\hat{a}^B, \hat{v}^C)}{f(\hat{v}^B, \hat{v}^C)} g(\hat{v}^C, \hat{v}^C) g(\hat{v}^B, \hat{v}^B) \]
\[ \times \frac{Z_{a,b-n}(\hat{u}^B, \hat{v}^B, \hat{v}^B)}{f(\hat{v}^B, \hat{v}^C)} Z_{a-b,n}(\hat{u}^B, \hat{v}^B, \hat{v}^B). \]  
(6.32)

It is easy to see that after appropriately relabeling the subsets we arrive at
\[ S_{a,b} = \sum_{r_1} \langle \hat{u}^C, r_1 \rangle \langle \hat{a}^B, r_3 \rangle \langle \hat{v}^C, r_2 \rangle \frac{f(\hat{a}^B, \hat{v}^C)}{f(\hat{v}^B, \hat{v}^C)} g(\hat{v}^C, \hat{v}^C) g(\hat{v}^B, \hat{v}^B) \]
\[ \times \frac{Z_{a-k,n}(\hat{u}^B, \hat{v}^B, \hat{v}^B)}{f(\hat{v}^B, \hat{v}^C)} Z_{a-k,n}(\hat{u}^B, \hat{v}^B, \hat{v}^B). \]  
(6.33)

where \( k = \#\hat{u}^B = \#\hat{a}^B \) and \( n = \#\hat{v}^C = \#\hat{v}^B \). Comparing this result with (4.8) and (4.11) we see that proposition 4.2 is proved.

### 7. Scalar product in the \( gl(1|1) \) sector

Consider a particular case of the subalgebra \( gl(1|1) \), generated by the operators \( T_{33}(u), T_{22}(u), T_{32}(u) \) and \( T_{32}(u) \). In this case one should set \( \hat{a}^C = \hat{a}^B = \varnothing \) in the formulas for the scalar product. Then the highest coefficient is simplified as
\[ Z_{0,b}(\varnothing; \varnothing; \varnothing) = \Delta_b(\varnothing) \Delta_b(\varnothing) \det g(s_j, v_k) = g(s, v), \]
(7.1)
where we use an explicit representation for the Cauchy determinant
\[ \det g(u_j, v_k) = g(\hat{u}, \hat{v}) \]  
\[ \Delta_m(\hat{u}) \Delta_m(\hat{v}). \]  
(7.2)

The scalar product (6.33) takes the form
\[ S_{0,b} = \sum_{r_1} \langle \hat{v}^C, r_1 \rangle \langle \hat{v}^B, r_3 \rangle g(\hat{v}^C, \hat{v}^C) g(\hat{v}^B, \hat{v}^B) g(\hat{v}^B, \hat{v}^B) \]
\[ \times \frac{\Delta_b(\hat{v}^C) \Delta_b(\hat{v}^C) \det [g(\hat{v}^C, \hat{v}^B) (r_3(\hat{v}^B) - r_3(\hat{v}^C))]}{f(\hat{v}^C, \hat{v}^B)}. \]  
(7.3)

Indeed, developing the determinant in (7.4) via the Laplace formula and using (7.2), (C.5), we obtain the sum (7.3).
Thus, the scalar product of Bethe vectors in $gl(1|1)$ integrable models admits a determinant representation without any restriction on the Bethe parameters. This is not surprising, as these models are equivalent to free fermions [26, 27].

8. Different representations for the highest coefficient

If $\varphi^c = \varphi^b = \emptyset$, then formula (6.33) describes the scalar product in the $gl(2)$-based models. In this case the scalar product admits a determinant representation if one of the Bethe vectors is an eigenvector of the transfer matrix. One expects that in the general $gl(2|1)$ case the sum over partitions in (6.33) can also be reduced to a single determinant for some particular cases of Bethe vectors. To make this reduction one should have different representations for the highest coefficient $Z_{a,b}(\bar{t}; \bar{x}|\bar{s}; \bar{y})$. In this section we give several formulas for $Z_{a,b}$ in terms of sums over partitions and multiple contour integrals.

We have already obtained an expression for $Z_{a,b}$ as the determinant of an $(a + b) \times (a + b)$ matrix

$$Z_{a,b}(\bar{t}; \bar{x}|\bar{s}; \bar{y}) = h(\bar{w}, \bar{t}) \Delta_{a+b}(\bar{w}) \Delta_{a}(\bar{t}) \Delta_{b}(\bar{y}) \det_{a+b} \mathcal{F}_{jk},$$

(8.1)

where $\bar{w} = \{x, s\}$ and the matrix $\mathcal{F}_{jk}$ is defined in (6.3):

$$\mathcal{F}_{jk} = \frac{g(w_j, t_k)}{h(w_j, t_k)}, \quad k = 1, \ldots, a;$$

$$\mathcal{F}_{j,k+a} = \frac{h(w_j, s_x)}{h(w_j, s_{\bar{t}})}, \quad k = 1, \ldots, b;$$

(8.2)

Developing the determinant with respect to the first columns we obtain

$$Z_{a,b}(\bar{t}; \bar{x}|\bar{s}; \bar{y}) = \sum \mathcal{K}_{a}(\bar{w}_1|\bar{t}) h(\bar{w}_1, \bar{x}) g(\bar{w}_1, \bar{y}) g(\bar{w}_1, \bar{t}_1).$$

(8.3)

The sum is taken over partitions $\{\bar{t}, \bar{x}\} = \bar{w}_1 \Rightarrow \#\bar{w}_1 = a$ and $\#\bar{t}_1 = b$.

Let us give several alternative representations for the highest coefficient.

- As a sum over partitions of $\bar{t}$ and $\bar{y}$:

$$Z_{a,b}(\bar{t}; \bar{x}|\bar{s}; \bar{y}) = \text{f}(\bar{s}, \bar{t}) f(\bar{y}, \bar{x}) \sum \mathcal{K}_{b}(\bar{y}_1|\bar{t}) \frac{h(\bar{t}, \bar{y}_1)}{h(\bar{s}, \bar{y}_1)} \mathcal{K}_{a}(\bar{x}|\bar{t}_1).$$

(8.4)

Here the sum is taken over partitions $\{\bar{t}, \bar{y} + c\} = \bar{\eta} \Rightarrow \#\bar{\eta}_1 = a$ and $\#\bar{\eta}_b = b$.

- As a sum over partitions of $\bar{t}$ and $\bar{x}$:

$$Z_{a,b}(\bar{t}; \bar{x}|\bar{s}; \bar{y}) = (-1)^a h(\bar{s}, \bar{x}) h(\bar{x}, \bar{y}) g(\bar{x}, \bar{y}) g(\bar{s}, \bar{y}) \sum \mathcal{K}_{a}(\bar{t} - c(\bar{\xi}_1)|\bar{t}) \frac{h(\bar{t}, \bar{\xi}_1)}{g(\bar{\xi}_1, \bar{y})} g(\bar{\xi}_1, \bar{\xi}_2) \mathcal{K}_{a}(\bar{t} - c|\bar{y}).$$

(8.5)

Here the sum is taken over partitions $\{\bar{t}, \bar{x} + c\} = \bar{\xi} \Rightarrow \#\bar{\xi}_1 = a$.

- As a sum over partitions of $\bar{s}$ and $\bar{y}$:

$$Z_{a,b}(\bar{t}; \bar{x}|\bar{s}; \bar{y}) = (-1)^{a+b} f(\bar{t}, \bar{s}) f(\bar{t}, \bar{t}) \sum \mathcal{K}_{a+b}(\bar{t}_1|\bar{t}) \mathcal{K}_{a+b}(\bar{x}|\bar{t}_1) \mathcal{K}_{a+b}(\bar{t}_1, \bar{t}_2|\bar{s}, \bar{y}).$$

(8.6)
Here the sum is taken over partitions \( \{ \bar{x} - c, \bar{y} \} = \bar{\nu} \Rightarrow \{ \bar{\nu}_1, \bar{\nu}_2 \} \) such that \( \# \bar{\nu}_1 = \# \bar{\nu}_2 = b \).

All the sum formulas listed above follow from (8.3) and can be proved via reduction of the sums over partitions to multiple contour integrals of the Cauchy type. Let us show how this method works.

Consider a \( b \)-fold integral
\[
I = \frac{(-1)^b}{(2\pi i)^b b!} \oint_{\bar{\nu}} K_{a+b}(\bar{w}|\{I, \bar{z} + c\}) \frac{g(\bar{z}, \bar{y})}{\Delta_b(\bar{z}) \Delta_b'(\bar{z})} \, d\bar{z},
\]
where \( \bar{w} = \{ \bar{x}, \bar{z} \} \) and \( d\bar{z} = dz_1, ..., dz_b \). We have used a subscript \( \bar{w} \) on the integral symbol in order to stress that the integration contour for every \( z_j \) surrounds the set \( \bar{w} \) in the anticlockwise direction. We also assume that the integration contours do not contain any other singularities of the integrand. Similar prescription will be kept for all other integral representations considered below.

The only poles of the integrand within the integration contours are the points \( z_j = \bar{w}_j \). Evaluating the integral by the residues in these poles we obtain (see appendix B for details)
\[
I = (-1)^b \sum K_a(\bar{w}|\{I, \bar{w} + c\}) h(\bar{w}, \bar{y}) g(\bar{w}, \bar{y}) g(\bar{w}, \bar{y}), \bar{x}),
\]
where the sum is taken over partitions of \( \bar{w} \) into subsets \( \bar{w}_1 \) and \( \bar{w}_2 \) with \( \# \bar{w}_1 = a \) and \( \# \bar{w}_2 = b \). Due to (A.1) we have
\[
K_{a+b}(\bar{w}|\{I, \bar{w} + c\}) = (-1)^b K_a(\bar{w}_1|\bar{f}),
\]
and comparing the obtained sum with (8.3) we immediately obtain \( I = Z_{a,b}(\bar{f}; \bar{x}|\bar{x}; \bar{y}) \).

Similarly, one can check that the sum over partitions in (8.3) can be presented as an \( a \)-fold contour integral
\[
Z_{a,b}(\bar{f}; \bar{x}|\bar{x}; \bar{y}) = (-1)^b h(\bar{w}, \bar{z}) g(\bar{y}, \bar{w}) \frac{K_a(\bar{w}|\bar{f}) g(\bar{w}, \bar{w})}{(2\pi i)^a a!} \oint_{\bar{w}} h(\bar{z}, \bar{y}) g(\bar{z}, \bar{y}) \frac{\Delta_a(\bar{z}) \Delta_a'(\bar{z})}{\Delta_a(\bar{z}) \Delta_a'(\bar{z})} \, d\bar{z},
\]
where now \( d\bar{z} = dz_1, ..., dz_a \). Indeed, taking the residues in the points \( \bar{z} = \bar{w}_1 \) we obtain
\[
Z_{a,b}(\bar{f}; \bar{x}|\bar{x}; \bar{y}) = (-1)^b h(\bar{w}, \bar{z}) g(\bar{y}, \bar{w}) \sum K_a(\bar{w}|\bar{f}) g(\bar{w}_1, \bar{w}_1) \frac{h(\bar{w}_1, \bar{x}) g(\bar{w}_1, \bar{y})}{h(\bar{w}_1, \bar{x}) g(\bar{w}_1, \bar{y})}.
\]
Multiplying the terms of the sum with the prefactor \( h(\bar{w}, \bar{z}) g(\bar{y}, \bar{w}) \) we arrive at (8.3).

Let us turn back to integral (8.7). Obviously, it can be calculated by taking the residues in the poles outside the original integration contour. It is easy to see that for arbitrary \( z_j \) the integrand behaves as \( 1/z_j^3 \) at \( z_j \rightarrow \infty \). Hence, the residue at infinity vanishes. The poles outside the original integration contours are in \( z_j = \bar{y}_k \) and \( z_j = \bar{x}_k - c \) (the poles at \( z_j = \bar{x}_k - c \) are compensated by the zeros of the product \( h(\bar{z}, \bar{x}) \)). Thus, we can move the original contour surrounding \( \bar{w} \) to the points \( D = \{ \bar{y}, \bar{x} - c \} \)
\[
Z_{a,b}(\bar{f}; \bar{x}|\bar{x}; \bar{y}) = \frac{1}{(2\pi i)^b b!} \oint_{\bar{w}} K_{a+b}(\bar{w}|\{I, \bar{z} + c\}) h(\bar{z}, \bar{x}) \frac{h(\bar{z}, \bar{y}) g(\bar{z}, \bar{w})}{\Delta_b(\bar{z}) \Delta_b'(\bar{z})} \, d\bar{z}.
\]
It is convenient to transform the integrand, applying (A.2) to \( K_{a+b}(\bar{w}|\{I, \bar{z} + c\}) \). Then substituting \( \bar{w} = \{ \bar{x}, \bar{z} \} \) and using the elementary properties of \( f(\bar{z}, \bar{w}) \) we obtain
and after simplification we arrive at

$$Z_{a,b} (\vec{t}; \vec{x} \mid \vec{y} ) = \frac{\mathcal{K}_{a+b} (\vec{t} - c, \vec{\bar{z}} ) \mathcal{g} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{h} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{g} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{g} (\vec{\bar{z}}, \vec{\bar{x}} ) \mathcal{g} (\vec{\bar{z}}, \vec{\bar{s}} ) }{\mathcal{f} (\vec{\bar{z}}, \vec{\bar{x}} ) \mathcal{f} (\vec{\bar{z}}, \vec{\bar{s}} ) \mathcal{\Delta}_b (\vec{\bar{z}} ) \mathcal{\Delta}_b (\vec{\bar{z}} ) } \times \mathcal{d} \vec{\bar{z}} .$$  

(8.13)

Now all the poles are explicitly combined in the product $\mathcal{p} \mathcal{g}$. Hence, the result of the integration gives the sum over partitions of $\mathcal{p} = \{ \bar{v}_1, \bar{v}_2 \}$ with $\# \bar{v}_1 = \# \bar{v}_2 = b$, which coincides with (8.6).

Applying (A.1) to the DWPF $\mathcal{K}_{a+b} (\{ \bar{v}_1, \bar{v}_2 \} \mid \{ \bar{x}, \bar{y} \} )$ in (8.6), we have

$$\mathcal{K}_{a+b} (\{ \bar{v}_1, \bar{v}_2 \} \mid \{ \bar{x}, \bar{y} \} ) = \mathcal{K}_{a+b} (\{ \bar{v}_1, \bar{v}_2 \} \mid \{ \bar{x}, \bar{y} \} + c )$$

(8.15)

Then the sum over partitions in (8.6) is equivalent to a multiple contour integral

$$Z_{a,b} (\vec{t}; \vec{x} \mid \vec{y} ) = \frac{\mathcal{K}_{a+b} (\vec{t} - c, \vec{z} ) \mathcal{g} (\vec{w}, \vec{t} ) \mathcal{f} (\vec{w}, \vec{t} ) \mathcal{f} (\vec{w}, \vec{t} ) \mathcal{g} (\vec{z}, \vec{x} ) \mathcal{g} (\vec{z}, \vec{s} ) }{\mathcal{f} (\vec{z}, \vec{x} ) \mathcal{f} (\vec{z}, \vec{s} ) \mathcal{\Delta}_b (\vec{z} ) \mathcal{\Delta}_b (\vec{z} ) } \times \mathcal{d} \vec{z} .$$  

(8.16)

Using (A.2) we recast (8.16) as

$$Z_{a,b} (\vec{t}; \vec{x} \mid \vec{y} ) = \frac{\mathcal{K}_{a+b} (\vec{x} - c, \vec{\bar{z}} ) \mathcal{g} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{h} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{g} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{g} (\vec{\bar{z}}, \vec{\bar{x}} ) \mathcal{g} (\vec{\bar{z}}, \vec{\bar{s}} ) }{\mathcal{f} (\vec{\bar{z}}, \vec{\bar{x}} ) \mathcal{f} (\vec{\bar{z}}, \vec{\bar{s}} ) \mathcal{\Delta}_b (\vec{\bar{z}} ) \mathcal{\Delta}_b (\vec{\bar{z}} ) } \times \mathcal{d} \vec{\bar{z}} .$$  

(8.17)

Setting $\eta = (\vec{t}, \vec{y} + c )$ now, we obtain

$$Z_{a,b} (\vec{t}; \vec{x} \mid \vec{y} ) = \frac{\mathcal{K}_{a+b} (\vec{x} - c, \vec{\bar{z}} ) \mathcal{g} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{h} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{g} (\vec{\bar{w}}, \vec{\bar{t}} ) \mathcal{g} (\vec{\bar{z}}, \vec{\bar{x}} ) \mathcal{g} (\vec{\bar{z}}, \vec{\bar{s}} ) }{\mathcal{f} (\vec{\bar{z}}, \vec{\bar{x}} ) \mathcal{f} (\vec{\bar{z}}, \vec{\bar{s}} ) \mathcal{\Delta}_b (\vec{\bar{z}} ) \mathcal{\Delta}_b (\vec{\bar{z}} ) } \times \mathcal{d} \vec{\bar{z}} .$$  

(8.18)

Now we can evaluate the integral by the residues outside the integration contours. All of them are collected in the product $\mathcal{h} (\vec{z}, \vec{\eta} - 2c )$. Taking the residues in the points $\vec{\eta} = 2c$ and using $h (x - 2c, y) = - h (y, x)$ we immediately arrive at (8.4).

Similarly, starting with integral representation (8.10) one can obtain sum formula (8.5).

### 9. Conclusion

In this paper we have derived a sum formula for the scalar product of Bethe vectors in models with $\mathfrak{g}(2|1)$ symmetry. We considered the case of generic Bethe vectors. This means that the Bethe parameters are generic complex numbers. However, one can also use this sum formula if the Bethe parameters obey some constraints. In some of these particular cases the sums over
partitions can be taken explicitly, leading eventually to determinant representations for the scalar products. In the second part of this paper we will consider these particular cases in more detail. We will show that if part of the Bethe parameters satisfy Bethe equations, then the sum formula in the $\text{gl}(2|1)$-based models reduces to a single determinant.

To conclude this paper we would like to mention that our results can be applied to models with $\text{gl}(1|2)$ symmetry as well. Indeed, due to an isomorphism between Yangians $Y(\text{gl}(1|2))$ and $Y(\text{gl}(2|1))$ (see [13]), it is enough to make the replacements $\{\tilde{u}^c, \tilde{v}^B\} \rightarrow \{v^c, \bar{v}^B\}$ and $a \leftrightarrow b$ in the sum formula. The obtained expression describes the scalar product of Bethe vectors in $\text{gl}(1|2)$-based models.

Acknowledgments

The work of A L has been funded by Russian Academic Excellence Project 5-100 and by joint NASU–CNRS project F14-2016. The work of S P was supported in part by RFBR grant 16-01-00562-a. N A S was supported by grants RFBR-15-31-20484-mol-a-ved and RFBR-14-01-00860-a.

Appendix A. Properties of DWPF

The DWPF $K_n(x|y)$ is a symmetric function of $x_1, \ldots, x_n$ and a symmetric function of $y_1, \ldots, y_n$. It behaves as $1/x_i$ (resp. $1/y_i$) as $x_i \to \infty$ (resp. $y_i \to \infty$) with other variables fixed. It has simple poles at $x_i = y_i$. It follows directly from definition (2.13) that the DWPF possesses the following properties:

$K_{n+m}((\tilde{x}, \tilde{z} - c)|\{\tilde{y}, \tilde{z}\}) = K_{n+m}((\tilde{x}, \tilde{z})|\{\tilde{y}, \tilde{z} + c\}) = (-1)^m K_n(\tilde{x}|\tilde{y}), \quad \#\tilde{z} = m,$

(A.1)

and

$K_n(\tilde{x} - c|\tilde{y}) = K_n(\tilde{x}|\tilde{y} + c) = (-1)^{n|c} K_n(\tilde{y}|\tilde{x}),\quad f(\tilde{y}, \tilde{x}).$

(A.2)

Appendix B. Sum over partitions as a contour integral

Proposition B.1. Let $\tilde{w} = \{w_1, \ldots, w_N\}$ be a set of complex numbers. Let $\mathcal{F}(\tilde{z})$ be a function of $n$ variables $z_1, \ldots, z_n$ ($n \leq N$). Assume that $\mathcal{F}(\tilde{z})$ is a symmetric function of $\tilde{z}$ and that it is holomorphic with respect to each $z_j$ within a domain containing the points $\tilde{w}$. Define

$$\langle \mathcal{F} \rangle = \frac{1}{(2\pi i)^n} \oint_{\tilde{w}} \frac{g(\tilde{z}, \tilde{w}) \, d\tilde{z}}{\Delta_0(\tilde{z}) \Delta_n(\tilde{z})^2} \mathcal{F}(\tilde{z}).$$

(B.1)

Here $d\tilde{z} = dz_1, \ldots, dz_n$ and the integration contour for every $z_j$ surrounds the points $\tilde{w}$ in the anticlockwise direction. We assume that there are no other singularities of the integrand within the integration contours. Then

$$\langle \mathcal{F} \rangle = \sum g(\tilde{w}_1, \tilde{w}_1) \mathcal{F}(\tilde{w}_1),$$

(B.2)

where the sum is taken over partitions $\tilde{w} \Rightarrow \{\tilde{w}_1, \tilde{w}_2\}$ such that $\#\tilde{w}_1 = n$. 
Proof. We use induction over $n$. For $n = 1$ the statement of the proposition is obvious. Suppose that it is valid for some $n - 1$. Then splitting $z = \{ z_n, z_{n-1} \}$, we obtain
\[
\mathcal{F} = \frac{1}{(2\pi i)^n} \int_{\partial D} \frac{g(z_n, w) g(z_{n-1}, w) \mathcal{F}(\{ z_n, z_{n-1} \})}{\Delta_n(z_n) \Delta_{n-1}(z_{n-1}) g(z_n, z_{n-1}) g(z_{n-1}, z_n)} dz_n dz_{n-1},
\]
where the sum is taken over partitions $\tilde{w} \Rightarrow \{ \tilde{w}_1, \tilde{w}_2 \}$ such that $\#\tilde{w} = n - 1$. Performing the integration over $z_n$ we find
\[
\mathcal{F} = \frac{1}{n} \sum g(\tilde{w}_1, \tilde{w}_2) \mathcal{F}(\{ \tilde{w}_1, \tilde{w}_2 \}),
\]
where we obtain an additional partition $\tilde{w}_3 \Rightarrow \{ \tilde{w}_1, \tilde{w}_2, \tilde{w}_3 \}$ with $\#\tilde{w}_3 = 1$. Substituting in (B.4) $\tilde{w}_3 = \{ \tilde{w}_1, \tilde{w}_2, \tilde{w}_3 \}$ and setting there $\tilde{w}_3 = w_0$ we arrive at
\[
\mathcal{F} = \frac{1}{n} \sum g(\tilde{w}_0, \tilde{w}_3) \mathcal{F}(\tilde{w}_0).
\]
Now the sum over partitions is organized as follows. First we have the partitions $\tilde{w} \Rightarrow \{ \tilde{w}_0, \tilde{w}_1 \}$ with $\#\tilde{w}_0 = n$, and then we have the additional partition $\tilde{w}_1 \Rightarrow \{ \tilde{w}_1, \tilde{w}_2 \}$ with $\#\tilde{w}_1 = 1$. Obviously, the sum over the later partition gives $n$, and we obtain the statement of the proposition.

Note that if $n > N$ in (B.1), then $\mathcal{F} = 0$.

Appendix C. Summation formulas

Lemma C.1. Let $\xi, \tilde{\alpha}$ and $\tilde{\beta}$ be sets of complex variables with $\#\alpha = n, \#\beta = m$, and $\#\xi = n + m$. Then
\[
\sum K_n(\xi) K_m(\beta) \xi g(\xi, \alpha) f(\xi, \beta) = (-1)^n f(\xi, \tilde{\alpha}) K_{n+m}(\{ \tilde{\alpha} - \beta, \beta \}) \xi.
\]
The sum is taken with respect to all partitions of the set $\xi$ into subsets $\tilde{\xi}$ and $\xi_{\tilde{\beta}}$ with $\#\tilde{\xi} = n$ and $\#\xi_{\tilde{\beta}} = m$.

The proof of this lemma can be found in [25].

Lemma C.2. For any set of functions $\phi_k(\beta)$, $k = 1, \ldots, n + m$, let
\[
\Phi_{n+m}(\beta) = \Delta_{n+m}(\beta) \det_{n+m} \phi_k(\beta),
\]
where $\tilde{\beta} = \{ \beta_1, \ldots, \beta_{n+m} \}$. Then
\[
\Phi_{n+m}(\beta) = \Delta_n(\beta_1) \det_{k=1, \ldots, n} \phi_k(\beta_1) \cdot \Delta_m(\beta_{n+1}) \det_{k=1, \ldots, m} \phi_k(\beta_m) \cdot g(\beta_n, \beta_m)
\]
where $\Phi_k(\tilde{\xi})$ is built on the functions $\phi_k$, $k = 1, \ldots, n$, while $\Phi_m(\tilde{\xi})$ is built on the functions $\phi_{n+k}$, $k = 1, \ldots, m$.

Proof. Developing the determinant in (C.2) over the first $n$ columns via the Laplace formula we obtain
\[ \Phi_{n+m}(\tilde{\beta}) = \Delta_{n+m}(\tilde{\beta}) \sum_{k=1}^{m} (-1)^{\hat{x}_k} \det_{k=1}^{n} \phi_{k}(\beta) \det_{k=1}^{n+m} \phi_{k+i}(\beta_i), \quad \text{(C.4)} \]

where the sum is taken over partitions \( \tilde{\beta} \Rightarrow \{ \beta_1, \beta_2 \} \) such that \( \#\beta_1 = n \). The sign \( P_{n;k} \) is the parity of a permutation mapping the union \( \{ \beta_1, \beta_2 \} \) into the naturally ordered set \( \hat{\beta} \). One can get rid of this sign presenting \( \Delta_{n+m}(\tilde{\beta}) \) as follows

\[ \Delta_{n+m}(\tilde{\beta}) = (-1)^{\hat{x}_1} \Delta_n(\hat{\beta}_1) \Delta_m(\hat{\beta}_2) \delta(\beta_1, \beta_1). \quad \text{(C.5)} \]

Substituting (C.5) into (C.4) we immediately arrive at (C.3). \( \square \)

We use several particular cases of (C.3) in the core of the paper. Let

\[ \phi_{k}(\beta) = \frac{g(\beta, x_k)}{h(\beta, x_k)}, \quad k = 1, \ldots, n; \]
\[ \phi_{k+i}(\beta) = \frac{g(\beta, y_k)}{h(\beta, y_k)}, \quad k = 1, \ldots, m, \quad \text{(C.6)} \]

where \( x, y, \tilde{t}, \) and \( s \) are some sets of parameters. Then the matrix elements \( \phi_{k}(\beta) \) coincide with the entries \( J_{\beta k} \) (6.3). Hence, we obtain for \( J_{\beta k}(x; y, \tilde{t}, s|\tilde{\beta}) \) (6.2)

\[ J_{\beta k}(x; y, \tilde{t}, s|\tilde{\beta}) = \Delta_n(\tilde{\beta}_1) \Delta_m(\tilde{\beta}_2) \sum_{\beta_1} \Delta_n(\beta) \Delta_m(\beta) \delta(\beta_1, \beta_1) \]
\[ \times \Delta_n(\hat{\beta}) \det_{\beta_1} \left( \frac{g(\beta, x_k)}{h(\beta, x_k)} \right) g(\beta, \tilde{\beta}_1). \quad \text{(C.7)} \]

Now we use the definition of the DWPF

\[ \Delta_n(\tilde{\beta}) \Delta_n(\hat{\beta}) \det_{\beta_1} \left( \frac{g(\beta, x_k)}{h(\beta, x_k)} \right) = K_n(\beta_1 | x), \quad \text{(C.8)} \]

and an explicit expression (7.2) for the Cauchy determinant \( \det_{\beta k} g(\beta_1, y_k) \). Substituting these expressions into (C.7) we find

\[ J_{\beta k}(x; y, \tilde{t}, s|\tilde{\beta}) = \sum_{\beta_1} K_n(\beta_1 | x) \cdot g(\beta, \tilde{\beta}_1), \quad \text{(C.9)} \]

which coincides with (6.4).

Another example used in the text is

\[ \phi_{k}(\beta) = g(\beta, x_k), \quad k = 1, \ldots, n; \]
\[ \phi_{k+i}(\beta) = g(\beta, y_k), \quad k = 1, \ldots, m. \quad \text{(C.10)} \]

Then using an explicit representation of the Cauchy determinant (7.2) we have

\[ \Phi_{n+m}(\tilde{\beta}) = g(\tilde{\beta}, \tilde{x}) g(\tilde{\beta}, \tilde{y}) \Delta_{n+m}(\tilde{x}, \tilde{y}). \quad \text{(C.11)} \]

On the other hand, it follows from (C.3) that

\[ g(\tilde{\beta}, \tilde{x}) g(\tilde{\beta}, \tilde{y}) \Delta_{n+m}(\tilde{x}, \tilde{y}) = \sum_{\beta_1} \Delta_n(\beta_1) \det_{k=1}^{n} g(\beta_1, x_k) \cdot \Delta_m(\hat{\beta}) \det_{k=1}^{m} g(\beta_2, y_k) \cdot g(\beta_1, \hat{\beta}_1). \quad \text{(C.12)} \]
Multiplying \((C.12)\) by \(\Delta'_{\lambda}(\vec{x})\) and \(\Delta'_{\mu}(\vec{y})\) and using \((7.2)\) we arrive at

\[
\sum g(\beta_1, \vec{x}) g(\beta_2, \vec{y}) g(\vec{\beta}_1, \vec{\beta}_2) = \frac{g(\vec{\beta}, \vec{x}) g(\vec{\beta}, \vec{y})}{g(x, y)},
\]

(C.13)

where the sum is taken with respect to the partitions of the set \(\vec{\beta}\) into subsets \(\vec{\beta}_1\) and \(\vec{\beta}_2\) with \(\#\vec{\beta}_1 = n\) and \(\#\vec{\beta}_2 = m\).

**Appendix D. Reduction properties of \(J_{n,m}\)**

Consider a function \(J_{n+1,m}(\{x, z'\}; y|\vec{f}; s|\vec{\beta}, z)\) defined by \((6.2)\). Let \(z' \to z\). Then the matrix element \(g(w, z')/h(w, z)\) becomes singular and the determinant reduces to the product of this singular element and the corresponding minor. After elementary algebra we obtain

\[
\lim_{z' \to z} \frac{1}{g(z, z')} J_{n+1,m}(\{x, z'\}; y|\vec{f}; s|\vec{\beta}, z) = g(\vec{\beta}, z) g(z, x) J_{n,m}(\vec{x}; y|\vec{f}; s|\vec{\beta}).
\]

(D.1)

Similarly, if we consider the function \(J_{n,m+1}(\{\tilde{y}, z'\}|\vec{f}; s|\vec{\beta}, z)\) in the limit \(z' \to z\), then the matrix element \(g(z, z') h(z, \vec{f})/h(z, s)\) becomes singular. The determinant again reduces to the product of this singular element and the corresponding minor, and we find

\[
\lim_{z' \to z} \frac{1}{g(z, z')} J_{n,m+1}(\{\tilde{y}, z'\}|\vec{f}; s|\vec{\beta}, z) = g(z, \vec{\beta}) g(\tilde{y}, z) \frac{h(z, \vec{f})}{h(z, s)} J_{n,m}(\vec{x}; y|\vec{f}; s|\vec{\beta}).
\]

(D.2)

Equations (D.1) and (D.2) obviously could be generalized to the case where \(x\) and \(z\) are respectively replaced with the sets \(\tilde{x}\) and \(\tilde{z}\) such that \(#\tilde{x} = #\tilde{z} = \rho \geq 1\). Then

\[
\lim_{z' \to z} \frac{1}{g(\tilde{x}, \tilde{z}')} J_{n+\rho,m}(\{x, z'\}; y|\vec{f}; s|\vec{\beta}, z) = g(\vec{\beta}, \tilde{x}) g(\tilde{x}, x) J_{n,m}(\tilde{x}; y|\vec{f}; s|\vec{\beta}),
\]

(D.3)

and

\[
\lim_{z' \to z} \frac{1}{g(\tilde{x}, \tilde{z}')} J_{n+\rho,m}(\{\tilde{y}, z'\}|\vec{f}; s|\vec{\beta}, z) = g(\tilde{x}, \tilde{z}) g(\tilde{y}, \tilde{z}) \frac{h(\tilde{z}, \vec{f})}{h(\tilde{z}, \vec{s})} J_{n,m}(\tilde{x}; y|\vec{f}; s|\vec{\beta}).
\]

(D.4)

One more obvious reduction is

\[
J_{n,m}(\tilde{x}; y|\vec{f}; s|\vec{\beta}) = J_{n,m}(\tilde{x}; y|\vec{f}; s|\vec{\beta}).
\]

(D.5)

**References**

[1] Faddeev L D, Sklyanin E K and Takhtajan L A 1979 Quantum inverse problem. I Theor. Math. Phys. 40 688–706

[2] Faddeev L D and Takhtajan L A 1979 The quantum method of the inverse problem and the Heisenberg XYZ model Usp. Math. Nauk. 34 13

Faddeev L D and Takhtajan L A 1979 The quantum method of the inverse problem and the Heisenberg XYZ model Russ. Math. Surv. 34 11 (Engl. transl.)

[3] Faddeev L D 1998 Les Houches Lectures Quantum Symmetries ed A Connes et al (Amsterdam: North Holland) p 149

[4] Kulish P P 1981 Classical and quantum inverse problem method and generalized Bethe ansatz Physica D 3 246–57
[5] Kulish P P and Reshetikhin N Yu 1981 Generalized Heisenberg ferromagnet and the Gross-Neveu model Zh. Eksp. Teor. Fiz. 80 214–28
Kulish P P and Reshetikhin N Yu 1981 Generalized Heisenberg ferromagnet and the Gross-Neveu model Sov. Phys. JETP 53 108–14 (Engl. transl.)

[6] Kulish P P and Reshetikhin N Yu 1983 Diagonalization of GL(N) invariant transfer matrices and quantum N-wave system (Lee model) J. Phys. A: Math. Gen. 16 L591–6

[7] Kulish P P and Sklyanin E K 1980 On the solution of the Yang-Baxter equation Zap. Nauchn. Semin. LOMI 95 129–60
Kulish P P and Sklyanin E K 1982 On the solution of the Yang-Baxter equation J. Sov. Math. 19 1596–620 (Engl. transl.)

[8] Kulish P P and Reshetikhin N Yu 1982 GL(3)-invariant solutions of the Yang-Baxter equation and associated quantum systems Zap. Nauchn. Sem. POMI 120 92–121
Kulish P P and Reshetikhin N Yu 1982 GL(3)-invariant solutions of the Yang-Baxter equation and associated quantum systems J. Sov. Math. 34 1948–71 (Engl. transl.)

[9] Kulish P P and Sasaki R 1993 Covariance properties of reflection equation algebras Prog. Theor. Phys. 89 741–62

[10] Damaskinskii E V and Kulish P P 2000 Symmetries connected with Yang-Baxter and reflection equations Zap. Nauchn. Semin. POMI 269 180–92
Damaskinskii E V and Kulish P P 2003 Symmetries connected with Yang-Baxter and reflection equations J. Math. Sci. 115 1986–93 (Engl. transl.)

[11] Avan J, Kulish P P and Rollet G 2011 Reflection K-matrices related to Temperley-Lieb R-matrices Theor. Math. Phys. 169 1530–8

[12] Kulish P P 1985 Integrable graded magnets Zap. Nauchn. Sem. LOMI 145 140–63
Kulish P P Integrable graded magnets J. Sov. Math. 35 2648–62 (Engl. transl.)

[13] Pakulyak S, Ragoucy E and Slavnov N A 2016 Bethe vectors for models based on super-Yangian Y(gl(m|n)) arXiv:1604.02311

[14] Hutsalyuk A, Lyashik A, Pakuliak S, Ragoucy E and Slavnov N A 2016 Multiple actions of the monodromy matrix in gl(2|1)-invariant integrable models arXiv:1605.06419

[15] Korepin V E 1982 Calculation of norms of Bethe wave functions Commun. Math. Phys. 86 391–418

[16] Izergin A G and Korepin V E 1984 The quantum inverse scattering method approach to correlation functions Commun. Math. Phys. 94 67–92

[17] Izergin A G 1987 Partition function of the six-vertex model in a finite volume Dokl. Akad. Nauk SSSR 297 331–3
Izergin A G 1987 Partition function of the six-vertex model in a finite volume Sov. Phys. Dokl. 32 878–9 (Engl. transl.)

[18] Reshetikhin N Yu 1986 Calculation of the norm of Bethe vectors in models with SU(3)-symmetry Zap. Nauchn. Sem. LOMI 150 196–213
Reshetikhin N Yu 1989 Calculation of the norm of Bethe vectors in models with SU(3)-symmetry J. Math. Sci. 46 1694–706 (Engl. transl.)

[19] Pakulyak S Z, Ragoucy E and Slavnov N A 2014 Scalar products in models with the GL(3) trigonometric R-matrix: general case Theor. Math. Phys. 180 795–814

[20] Slavnov N A 2015 Scalar products in GL(3)-based models with trigonometric R-matrix. Determinant representation J. Stat. Mech. P03019

[21] Slavnov N A 1989 Calculation of scalar products of wave functions and form factors in the framework of the algebraic Bethe ansatz Theor. Math. Phys. 79 502–8

[22] Kitanine N, Maillet J M and Terras V 1999 Form factors of the XXZ Heisenberg spin-1/2 finite chain Nucl. Phys. B 554 647–78

[23] Belliard S, Pakuliak S, Ragoucy E and Slavnov N A 2012 The algebraic Bethe ansatz for scalar products in SU(3)-invariant integrable models J. Stat. Mech. P10017

[24] Slavnov N A 2016 Multiple commutation relations in the models with gl(2|1) symmetry arXiv:1604.05343

[25] Belliard S, Pakuliak S, Ragoucy E and Slavnov N A 2012 Highest coefficient of scalar products in SU(3)-invariant integrable models J. Stat. Mech. P09003

[26] Foerster A and Karowski M 1993 Algebraic properties of the Bethe ansatz for an spl(2, 1)-supersymmetric t-J model Nucl. Phys. B 396 611–38

[27] Göhmann F and Murakami S 1998 Fermionic representations of integrable lattice systems J. Phys. A 31 7729–50