Записки научных
семинаров ПОМИ
Том 494, 2020 г.

V. S. Kapitonov, A. G. Pronko

SIX-VERTEX MODEL AS A GRASSMANN INTEGRAL, ONE-POINT FUNCTION, AND THE ARCTIC ELLIPSE

Abstract. We formulate the six-vertex model with domain wall boundary conditions in terms of an integral over Grassmann variables. Relying on this formulation, we propose a method of calculation of correlation functions of the model for the case of the weights satisfying the free-fermion condition. We consider here in details the one-point correlation function describing the probability of a given state on arbitrary edge of the lattice, or, – polarization. We show that in the thermodynamic limit, performed such that the lattice is scaled to the square of unit side length, this function exhibits the "arctic ellipse" phenomenon, in agreement with previous studies on random domino tilings of Aztec diamonds: it approaches its limiting values outside of an ellipse inscribed into this square, and takes continuously intermediate values inside the ellipse. We derive also scaling properties of the one-point function in the vicinity of an arbitrary point of the arctic ellipse and in the vicinities of the points where the ellipse touches the boundary.

§1. Introduction

The present interest in the study of 2D lattice models of classical statistical mechanics is in particularly aimed at unveiling their thermodynamic properties when fixed (rather than periodic) boundary conditions are imposed. An important example here is the six-vertex model with domain wall boundary conditions [1–3], which was originally introduced in the context of the problem of calculation of correlation functions of 1D quantum solvable models (for a review, see [4]). Thereafter it was discovered that the model is also closely related with some interesting objects in enumerative combinatorics (see, e.g., [5] and references therein). Most interestingly, the model is known to exhibit spatial separation of phases in the thermodynamic limit, which is the effect due to both of the ice-rule and the peculiar choice of the boundary conditions [6].

Key words and phrases: Vertex models, lattice fermions, Grassmann variables, domain wall boundary conditions, coherent states, correlation functions, arctic ellipse phenomenon, Airy kernel.

Phase separation phenomena are very well known to be an attribute of random tilings of planar finite regions. A famous example, closely related to the present study, is provided by domino tilings of the region called "Aztec diamond" where configurations exhibit "arctic circle" phenomenon: in statistically dominating configurations dominoes are frozen (i.e., ordered) outside of a circle inscribed into the Aztec diamond, and temperate (i.e., disordered) inside it [7,8]. For dominoes certainly weighted according to their orientations, or, "biased", the circle deforms into an ellipse [9]. Similar phenomena are also known for rhombus tilings of a hexagon [10, 11], and more generally, for dimer models with a boundary [12, 13].

As the domino tilings of Aztec diamonds can be mapped to the six-vertex model with domain wall boundary conditions at its free-fermion point [14], phase separation phenomena can be expected at the six-vertex model side too, and not only at the free-fermion point. This is also reasoned by obtained influence of boundary conditions on the value of free energy per site in the six-vertex model with generic weights [15, 16]; some analytical aspects were also discussed in [17, 18]. The presence of the phase separation phenomena in six-vertex model with domain boundary conditions for the whole disordered and anti-ferroelectric regimes were demonstrated numerically in [19, 20].

Motivated by the problem of the phase separation phenomena in the six-vertex model with domain wall boundary conditions, a series of papers was devoted to calculation of its correlation functions in the last decade. Among correlation functions computed up to date are boundary one [21, 22] and two-point functions [23, 24], and certain bulk, though non-local, correlation function, called emptiness formation probability [25]. This last correlation function permitted to obtain an analytical expression for the arctic curve (i.e., frozen boundary of the limit shape, in the language of random surfaces) in the model [26–28].

To obtain more sophisticated information, such as density of local states or averaged height-function (i.e., the whole limit shape, rather than just its boundary) it is necessary to have an access to local bulk correlation functions. A simplest such correlation function is the one-point function describing the edge-state probability, or, – polarization. Although some steps were made towards solution of this problem in full settings recently [29, 30], it still remains open.

In view of this, it seems interesting to consider first the same problem for a simple but nevertheless interesting special case, in which the six-vertex

model weights are restricted to satisfy the free-fermion condition. Indeed, because of lack of bijection between configurations of domino tilings and six-vertex model (see, e.g., [14,31]), it is useful to gain access to quantities of interest directly in the framework of the later. At the same time, one can expect that scaling properties of correlations are somehow universal, and the results obtained previously for domino tilings are indeed reproduced in the thermodynamic limit.

The present paper is therefore devoted to treating the free-fermion case in the six-vertex model with domain wall boundary conditions. We focus our attention here on the one-point correlation function which describes the edge-state probability. We obtain various representations for this function and recover, in the thermodynamic limit, the arctic ellipse phenomenon (separation of order and disorder), known from previous studies on domino tilings. We also discuss finite size effects in the vicinities of the phase separation curve and points where the disordered region touches the boundary.

The presentation of the paper is made as much as possible to be self-contained. Our treatment of the free-fermion case of the six-vertex model is based on a direct reformulation of the model in terms of fermions–fermionization, rather than on the standard for this model framework of the quantum inverse scattering method. The fermionization is applicable for generic weights and arbitrary boundary conditions. As a result, the partition function is written in terms of an integral over Grassmann variables, resembling a fermionic 2D lattice field theory with a four-fermion interaction term. At the free-fermion point this term vanishes. The construction, with the explicit use of the domain wall boundary conditions, is explained in Sect. 2.

Using this formulation, we consider next in Sect. 3 calculations for the free-fermion model. We develop here an approach based on evaluation of powers of involved matrices in an explicit form, without any diagonalization procedure. This makes it possible to compute correlation functions in the coordinate representation directly. We obtain various formulas for the one-point function, including a representation in terms of a double contour integral, suitable for an analysis in the thermodynamic limit.

We perform the saddle-point analysis of this integral in Sect. 4. In the scaling limit where the initial lattice is mapped to the unit square, the one-point function demonstrates the "arctic ellipse" phenomenon, namely, it acquires its limiting values in the exterior of an ellipse inscribed into the

square and takes intermediate values in the interior of it. We also show that for large but finite lattices the one-point function at the vicinity of the arctic ellipse coincides with the density of eigenvalues of the Gaussian unitary ensemble at the edge of the spectrum, as it should. We also investigate similarly the one-point function in vicinity of the points where the ellipse touches the boundary.

This paper contains results obtained by the authors in the period 2003–2013 and it was almost in its final form, when the first author died suddenly in February 2014. Since then, this text was shared among colleagues as a draft manuscript. Inspired by many requests, the second author have finally decided to publish it, and henceforth it is presented here in its original form, with minimal modifications. In particular, this explains lack of citations in the list of references dated after 2013.

§2. Fermionization of the model

Under the term fermionization we assume here that the partition function is represented as an integral over Grassmann variables. We start with giving basic definitions related to the model. Using the formalism of diagonal-to-diagonal transfer matrices, we express first the partition function as some quantum expectation value in the Hilbert space of a finite number of fermionic degrees of freedom. We show that this expectation value can be written in terms of a fermionic transfer matrix. We invoke the notion of the Grassmann coherent states to switch from the operator formulation of the partition function to that in terms of an integral over Grassmann variables, thus performing the fermionization.

2.1. Basic definitions. We consider the six-vertex model on a finite square lattice formed by intersection of N vertical and N horizontal lines ($N \times N$ lattice). We recall that configurations of the model can be depicted by placing arrows pointed along edges of the lattice, with the condition that at each lattice vertex must income and outgo exactly two arrows. The domain wall boundary conditions mean that all arrows on external edges of our $N \times N$ lattice are fixed such that on the horizontal edges all arrows are outgoing while on the vertical ones they are incoming, see Fig. 1.

To each of six typical arrow configurations around a vertex is assigned a Boltzmann weight. We use standard assumption that these weights are invariant under simultaneous reversal of all four arrows, and denote the

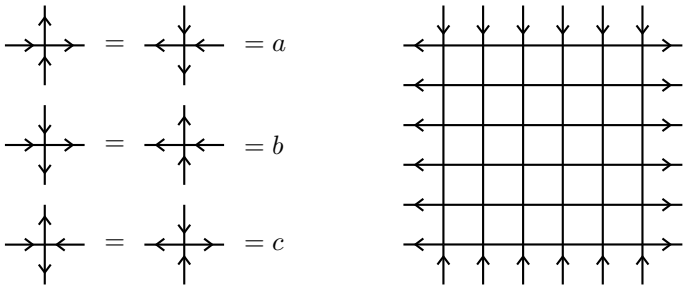

Figure 1. Definition of the model: the six allowed vertex configurations with their Boltzmann weights, and $N \times N$ lattice with the domain wall boundary conditions, $N = 6$.

three different weights as a, b, and c. In this section we assume that they are arbitrary.

For description of the states on the edges it is often more convenient to use binary variables (i.e., taking values $0, 1$) rather than arrows. Let such a variable takes value 0 if the arrow is pointing upward or right, and value 1 if the arrow is pointing downward of left. Consider a single vertex and let μ, ν, ν' and μ' denote variables assigned to the left, bottom, top, and right edges of the vertex, respectively. Let $w(\mu, \nu, \nu', \mu')$ denotes the Boltzmann weight corresponding to this vertex, then

$$w(0,0,0,0) = w(1,1,1,1) = a,$$
$$w(0,1,1,0) = w(1,0,0,1) = b,$$
$$w(0,1,0,1) = w(1,0,1,0) = c,$$

and $w(\mu, \nu, \nu', \mu') = 0$, if $\mu + \nu \neq \nu' + \mu'$.

The partition function of the six-vertex model with domain wall boundary conditions then can be defined as

$$Z = \sum_{\{\mu\},\{\nu\}} \prod_{i,j=1}^{N} w(\mu_{i,j}, \nu_{i,j}, \nu_{i,j+1}, \mu_{i+1,j}) \Big|_{\substack{\mu_{1,*}=\nu_{*,N+1}=1 \\ \mu_{N+1,*}=\nu_{*,1}=0}}. \tag{2.1}$$

Here sum is taken over values of the variables assigned to the internal edges of the $N \times N$ lattice, and the star means that the subscript must run over all possible values.

In discussion of the one-point correlation function which describes the probability of a given state on an edge, it will be convenient for us to

use labeling of the lattice vertices by the Cartesian coordinates (m, n), where $m, n = 1, \dots, N$, with the origin $(1, 1)$ corresponding to the bottom left vertex. The one-point function at arbitrary edge of the lattice can be defined as

$$G_{\mathcal{A}}(m, n) = \langle \chi_{\mathcal{A}}(m, n) \rangle , \qquad \mathcal{A} = \uparrow, \leftarrow, \downarrow, \rightarrow,$$

where the averaging is performed with respect to the Gibbs measure induced by (2.1) and $\chi_{\mathcal{A}}(m, n)$ is the characteristic function of the arrow $\mathcal{A}$ outgoing from the vertex (m, n),

$$\chi_{\mathcal{A}}(m, n) = \begin{cases} 1 & \text{if the arrow } \mathcal{A} \text{ outgoes from the vertex } (m, n) \\ 0 & \text{otherwise.} \end{cases}$$

For instance,

$$G_{\rightarrow}(m, n) = \frac{1}{Z} \sum_{\{\mu\},\{\nu\}} \delta(\mu_{m+1,n}, 0) \times \prod_{i,j=1}^{N} w(\mu_{i,j}, \nu_{i,j}, \nu_{i,j+1}, \mu_{i+1,j}) \Big|_{\substack{\mu_{1,*}=\nu_{*,N+1}=1 \\ \mu_{N+1,*}=\nu_{*,1}=0}}, \quad (2.2)$$

where $\delta(\mu, \mu')$ is the Kronecker symbol,

$$\delta(\mu, \mu') = \begin{cases} 1 & \text{if } \mu = \mu' \\ 0 & \text{if } \mu \neq \mu'. \end{cases}$$

Evidently, we have the relations

$$G_{\rightarrow}(m, n) + G_{\leftarrow}(m + 1, n) = 1, \qquad G_{\uparrow}(m, n) + G_{\downarrow}(m, n + 1) = 1.$$

Using the symmetries both of the lattice with domain wall boundary conditions and the vertex weights under reflections with respect to diagonals (see Fig. 1), we also have the relations

$$G_{\rightarrow}(m, n) = G_{\downarrow}(n, m + 1) = G_{\uparrow}(N - n + 1, N - m).$$

These relations mean that it is sufficient to consider the edge-state probability only for, e.g., horizontal edges. Below we consider function (2.2) and use notation $G(m, n) = G_{\rightarrow}(m, n)$. Furthermore, from the four relations above for this function we have the following self-consistency relation:

$$G(m, n) = 1 - G(N - m, N - n + 1). \quad (2.3)$$

This relation can also be obtained by rotating the lattice through 180° (since such rotation is just a composition of two reflections). Obviously, relation (2.3) can be seen as a kinematic restriction, since it arises merely due to global symmetries of the system; in the periodic boundary conditions case it would be replaced by the condition of a translational invariance of the edge-state probability.

In the remaining parts of this section we discuss reformulation of the partition function in terms of integral over Grassmann variables, postponing further discussion of the one-point function to Sect. 3.3.

2.2. Operator formulation. We start with considering the model in the framework of the diagonal-to-diagonal transfer matrix formalism, also known as the light-cone lattice approach. As it was noticed in [32], in this formalism one can easily apply the Jordan-Wigner transformation to spin-1/2 operators to obtain an equivalent description in terms of canonical fermion operators.

To employ the diagonal-to-diagonal transfer matrix formalism, it is useful to look at the $N\times N$ lattice as being obtained by attaching of $2N$ zigzag lines of proper lengths, and to enumerate these lines consecutively as shown in Fig. 2. To each of these lines we assign a vector space $\mathbb{C}^2$. Consider the vector space $\mathcal{H} = (\mathbb{C}^2)^{\otimes 2N}$ where the jth copy of $\mathbb{C}^2$ corresponds to the jth zigzag line. Then to the vertices of the $N \times N$ lattice we can assign operators X_j, $j = 1, \dots, 2N-1$, acting in $\mathcal{H}$, such that each operator X_j acts non-trivially only in the jth and $(j+1)$th copies of $\mathbb{C}^2$.

Let v_0, v_1 form a basis in $\mathbb{C}^2$,

$$v_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

For the basis vectors of the space $\mathcal{H}$, which will play the role of the quantum space of states, we will employ the "bra-ket" notation

$$|\Psi_{\mu_1,\mu_2,\dots,\mu_{2N}}\rangle = v_{\mu_1} \otimes v_{\mu_2} \otimes \cdots \otimes v_{\mu_{2N}}, \qquad \mu_1, \dots, \mu_{2N} = 0, 1.$$

We thus can define the operators X_j, $j = 1, \dots, 2N-1$, by the formula

$$\langle \Psi_{\mu_1,\dots,\mu_{2N}} | X_j | \Psi_{\mu'_1,\dots,\mu'_{2N}} \rangle = w(\mu_j, \mu_{j+1}, \mu'_j, \mu'_{j+1}) \prod_{\substack{l=1 \\ l\neq j,j+1}}^{2N} \delta(\mu_l, \mu'_l).$$

Here $w(\mu, \nu, \nu', \mu')$ is the Boltzmann weight defined above.

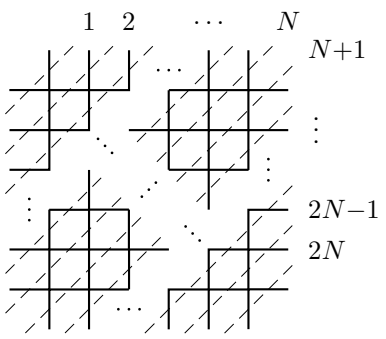

Figure 2. The $2N$ zigzag lines forming the $N \times N$ lattice: each zigzag line is formed by the set of edges crossed by the same dashed line.

In terms of the spin-1/2 operators,

$$\sigma_j^\alpha = I^{\otimes(j-1)} \otimes \sigma^\alpha \otimes I^{\otimes(2N-j)}, \qquad \alpha = x, y, z, \tag{2.4}$$

where

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and I denotes 2×2 unity matrix, the X-operators read

$$X_j = \frac{a+c}{2} + \frac{b}{2}\left(\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y\right) + \frac{a-c}{2}\,\sigma_j^z \sigma_{j+1}^z. \tag{2.5}$$

The partition function of the six-vertex model with domain wall boundary conditions can be written in the form

$$Z = \langle \Lambda | \Xi | \Lambda \rangle, \tag{2.6}$$

where the vector $|\Lambda\rangle$ accounts the boundary conditions of the top and right boundaries,

$$|\Lambda\rangle = |\Psi \underbrace{1, \ldots, 1}_{N} \underbrace{0, \ldots, 0}_{N}\rangle,$$

and, similarly, the vector $\langle\Lambda|$ accounts those of the left and bottom boundaries. The operator Ξ describes the transfer over the whole lattice, from

the top right corner to the bottom left one, and reads:

$$\Xi = X_N(X_{N-1}X_{N+1})\cdots(X_1X_3\cdots X_{2N-3}X_{2N-1})\cdots \times (X_{N-1}X_{N+1})X_N. \quad (2.7)$$

In this formula, each group of operators is a diagonal-to-diagonal transfer matrix, which corresponds to a set of vertices along the NW-SE diagonal direction (see Fig. 2). Namely, the rightmost operator, X_N, corresponds to the top right vertex of the lattice, the operator $X_{N-1}X_{N+1}$ corresponds to the second (from the top right corner) diagonal row of vertices, and so on, down to the bottom left vertex, to which the operator X_N corresponds.

It is useful to note that since $X_jX_k = X_kX_j$ for $|j-k| \geqslant 2$, the X-operators in the expression for Ξ can also be grouped along the vertical or horizontal lines of the lattice. For example, grouping along the vertical lines, we can write

$$\Xi = (X_N\cdots X_2X_1)\cdots(X_{2N-2}\cdots X_NX_{N-1})(X_{2N-1}\cdots X_{N+1}X_N). \quad (2.8)$$

In this expression, the order of the groups of X-operators coincides with the order of the vertical lines to which they correspond. The order of the X-operators within each group, from right to left, corresponds to running over the vertices of a vertical line, from top to bottom.

Now we can apply the Jordan-Wigner transformation and express all operators in terms of canonical fermion operators, which we denote by f_j and f_j^*, $j = 1,\dots,2N$. They satisfy the anti-commutation relations

$$f_jf_k^* + f_k^*f_j = \delta_{jk}, \qquad f_jf_k + f_kf_j = 0, \qquad f_j^*f_k^* + f_k^*f_j^* = 0.$$

Using the notation $\sigma_j^{\pm} = (1/2)(\sigma_j^x \pm \mathrm{i}\sigma_j^y)$, we define them as

$$f_j = \sigma_1^z\cdots\sigma_{j-1}^z\sigma_j^+, \qquad f_j^* = \sigma_1^z\cdots\sigma_{j-1}^z\sigma_j^-. \quad (2.9)$$

The operators f_j^* and f_j can be seen as the creation and annihilation operators, respectively, of the jth fermion, with respect to the vacuum vector $|\Omega\rangle = |\Psi_{0,\dots,0}\rangle$. For the basis vectors of the space $\mathcal{H}$ we have

$$|\Psi_{\mu_1,\mu_2,\dots,\mu_{2N}}\rangle = (f_1^*)^{\mu_1}(f_2^*)^{\mu_2}\cdots(f_{2N}^*)^{\mu_{2N}}|\Omega\rangle, \quad (2.10)$$

and, correspondingly, the bra-vectors read

$$\langle\Psi_{\mu_1,\mu_2,\dots,\mu_{2N}}| = \langle\Omega|(f_{2N})^{\mu_{2N}}\cdots(f_2)^{\mu_2}(f_1)^{\mu_1}. \quad (2.11)$$

As a result, the partition function is expressed in terms of the algebra of canonical fermions, by (2.6) and (2.7) (or (2.8)), where

$$\langle\Lambda| = \langle\Omega|f_Nf_{N-1}\cdots f_1, \qquad |\Lambda\rangle = f_1^*f_2^*\cdots f_N^*|\Omega\rangle, \quad (2.12)$$

and the X-operators are given by the expression

$$X_j = a - (a-c)\left(f_j^* f_j + f_{j+1}^* f_{j+1} - 2 f_j^* f_j f_{j+1}^* f_{j+1}\right) + b(f_j^* f_{j+1} + f_j^* f_{j+1}), \quad (2.13)$$

which follows from (2.4), (2.5), and (2.9).

2.3. Fermionic transfer matrix. Our aim now will be rewrite the partition function in the form, typical for the row-to-row transfer matrix formalism. In particular, we express it as a matrix element of the Nth power of certain operator, which can be written as a supertrace of a product of local operators.

We first introduce the permutation operators of fermions, which can be defined by the relations

$$P_{jk} f_j P_{jk} = f_k, \qquad P_{jk} f_j^* P_{jk} = f_k^*, \qquad P_{jk}|\Omega\rangle = |\Omega\rangle,$$

together with the standard relations of the permutation group, i.e.,

$$P_{jk} P_{lk} P_{jk} = P_{lj}, \qquad (P_{jk})^2 = 1.$$

These operators have the following explicit form:

$$P_{jk} = 1 - f_j^* f_j - f_k^* f_k + f_j^* f_k + f_k^* f_j. \quad (2.14)$$

Given permutation operators, we can consider the cyclic shift operator of $2N$ fermions,

$$C = P_{12} P_{23} \cdots P_{2N-1,2N},$$

which has the property $C^{2N} = 1$.

Let us now consider the partition function, given by (2.6), focusing first attention on the operator Ξ, written in the form (2.8). Using

$$X_{j+1} = C X_j C^{-1},$$

we can write each group of factors in (2.8) as follows:

$$X_{N+j-1} \cdots X_{j+1} X_j = C^j \left(X_N \cdots X_2 X_1\right) C^{-j}.$$

This yields

$$\Xi = T^N C^{-N},$$

where we have introduced the operator

$$T = X_N \cdots X_2 X_1 C. \quad (2.15)$$

Now, introducing the vector

$$|\Lambda'\rangle := C^N |\Lambda\rangle = f_{N+1}^* f_{N+2}^* \cdots f_{2N}^* |\Omega\rangle,$$

where the vector $|\Lambda\rangle$ is given in (2.12), we can rewrite the partition function in the form

$$Z = \langle \Lambda | T^N | \Lambda' \rangle. \tag{2.16}$$

To proceed further, we introduce one more set of local operators, R_{jk}, where $j, k = 1, \dots, 2N$ and $j \neq k$. We first introduce the operators R_{jk} with $k = j + 1$, by

$$R_{j,j+1} := X_j P_{j,j+1}.$$

Acting with the permutation operators we can extend the above definition on all other values of indices,

$$R_{jk} = P_{k,j+1} R_{j,j+1} P_{k,j+1} \qquad (k \neq j, j+1).$$

Due to (2.13) and (2.14), the operators R_{jk} thus defined has the following explicit form

$$R_{jk} = a + c(f_j^* f_k + f_k^* f_j) + (b - a)(f_j^* f_j + f_k^* f_k) - 2b\, f_j^* f_j f_k^* f_k. \tag{2.17}$$

In this expression one can easily recognize the so-called fermionic R-matrix [33, 34]. Note that we have in fact obtained here (2.17) simply from the expression for the X-operators, first using the Jordan-Wigner transformation, and next acting with the fermion permutation operators (cf. also [32]).

Let us now come back to the operator (2.15). Using the definition of the R-operators given above, we can rewrite the operator T in the form

$$T = (R_{N,N+1} R_{N-1,N+1} \cdots R_{1,N+1}) \times (P_{N+1,N+2} P_{N+2,N+3} \cdots P_{2N-1,2N}). \tag{2.18}$$

Noting that the second group of factors here, involving only the permutation operators, is just the cyclic shift operator over the last N fermions (out of total $2N$), we can use the fact that its Nth power is equal to one, and therefore we can simplify all such factors when considering the operator T^N, that yields

$$T^N = (R_{N,N+1} R_{N-1,N+1} \cdots R_{1,N+1})\, (R_{N,N+2} R_{N-1,N+2} \cdots R_{1,N+2}) \cdots \times (R_{N,2N} R_{N-1,2N} \cdots R_{1,2N})\,. \tag{2.19}$$

This representation, when substituted in (2.16), provides the fermionic version of the usual row-to-row transfer matrix formulation of the partition function, cf. [1–3]. In (2.19), each operator $R_{j,N+k}$ corresponds to the vertex lying at intersection of the jth horizontal line, counted from the top, and the kth vertical line, counted from the left.

In what follows we rely on the representation (2.16) with the operator T given in the form typical for a transfer matrix, i.e., as a (super)trace of some product of operators. Indeed, (2.18) can be obtained as a result of evaluation of the supertrace with respect to an auxiliary, or, "zeroth", fermion degree of freedom,

$$T = \mathrm{str}_0 \left(P_{2N,0} P_{2N-1,0} \cdots P_{N+1,0} R_{N,0} R_{N-1,0} \cdots R_{1,0}\right). \tag{2.20}$$

The supertrace can be defined as

$$\mathrm{str}_0(\dots) = \langle \Omega_0 | \dots | \Omega_0 \rangle - \langle \Omega_0 | f_0 \dots f_0^* | \Omega_0 \rangle,$$

where $|\Omega_0\rangle$ denotes the vacuum state of the auxiliary fermion. Equivalence of (2.18) and (2.20) can be seen by moving the operator $P_{2N,0}$ in (2.20) over all the operators to the right and using the fact that $\mathrm{str}_0 P_{2N,0} = 1$.

2.4. Grassmann integral representation. To represent (2.16) as an integral over Grassmann variables, we will exploit the technique of the Grassmann coherent states. For a detailed exposition on the Grassmann variables and integration over them see [35, 36]; the Grassmann coherent states are exposed, e.g., in [37, 38] (see also references therein), where also all the sufficient information on the Grassmann integrals can be found.

First of all. we recall that Grassmann variables are classical variables (not operators), which anti-commute between themselves. The standard requirement in their definition is that they also anti-commute with the fermion creation and annihilation operators. Given $2N$ fermion degrees of freedom in our operators, we associate Grassmann variables $\eta_1, \dots, \eta_{2N}$ with the ket-state $|\eta\rangle := |\eta_1, \dots, \eta_{2N}\rangle$, and Grassmann variables $\eta_1^*, \dots, \eta_{2N}^*$ with the bra-state $\langle \eta^*| := \langle \eta_1^*, \dots, \eta_{2N}^*|$.

The defining property of Grassmann coherent states is that they are eigenstates of the fermion annihilation and creation operators,

$$f_j |\eta\rangle = \eta_j |\eta\rangle, \qquad \langle \eta^* | f_j^* = \langle \eta^* | \eta_j^*. \tag{2.21}$$

We define the Grassmann coherent states as follows:

$$|\eta\rangle = \exp\left\{\sum_{j=1}^{2N} f_j^* \eta_j\right\} |\Omega\rangle, \qquad \langle \eta^*| = \langle \Omega| \exp\left\{\sum_{j=1}^{2N} \eta_j^* f_j\right\}. \tag{2.22}$$

Their scalar product is

$$\langle \eta^* | \eta \rangle = \exp\left\{\sum_{j=1}^{2N} \eta_j^* \eta_j\right\}. \tag{2.23}$$

Note that we assume that η_j and η_j^*, $j = 1, \dots, 2N$, are independent variables (i.e., no complex involution is introduced in the Grassmann algebra). We also use, in comparison with [37, 38], a different normalization of the states; our choice is more suitable for our purposes below.

A important role in calculations is played by the unity decomposition

$$1 = \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j) \, \exp\left\{ - \sum_{j=1}^{2N} \eta_j^* \eta_j \right\} |\eta\rangle\langle\eta^*|. \tag{2.24}$$

We use the usual rules for the integrals over Grassmann variables: an integral over a Grassmann variable is defined as a derivative (from the left) and multiple integrals are defined as repeated ones. In writing integrals we follow the convention that the order of differentials coincides with the order of differentiations. We often use below a formula for a $2M$-tuple (where M varies, depending on a situation) Gaussian Grassmann integral

$$\int \prod_{j=1}^{M} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j) \, \exp\left\{ - \sum_{j,k=1}^{M} \eta_j^* \mathcal{A}_{jk} \, \eta_k + \sum_{j=1}^{M} \xi_j^* \eta_j + \sum_{j=1}^{M} \eta_j^* \xi_j \right\}$$
$$= \left(\det_{1 \leqslant j,k \leqslant M} \mathcal{A}_{jk} \right) \exp\left\{ \sum_{j,k=1}^{M} \xi_j^* (\mathcal{A}^{-1})_{jk} \, \xi_k \right\}, \tag{2.25}$$

where $\xi_1, \dots, \xi_M, \xi_1^*, \dots, \xi_M^*$ are Grassmann variables; the matrix $\mathcal{A}$ is assumed to be a nondegenerate.

Since (2.24) is important for what follows, it is useful to outline a proof. We first note that (2.24) essentially implies that the right-hand side acts identically on all the basis vectors (2.10) and (2.11). Next we note that the Grassmann coherent states (2.22) are in fact generating functions of all basis vectors; hence it is suffice to show that a Grassmann coherent state

is reproduced. Finally, acting, for example, on the ket-state $|\xi\rangle$, we have

$$\begin{aligned}
&\int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j)\, \exp\left\{ -\sum_{j=1}^{2N} \eta_j^* \eta_j \right\} |\eta\rangle\langle\eta^*|\xi\rangle \\
&\quad = \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j)\, \exp\left\{ -\sum_{j=1}^{2N} \eta_j^* \eta_j + \sum_{j=1}^{2N} \eta_j^* \xi_j \right\} |\eta\rangle \\
&\quad = \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j)\, \exp\left\{ -\sum_{j=1}^{2N} \eta_j^* \eta_j + \sum_{j=1}^{2N} \eta_j^* \xi_j + \sum_{j=1}^{2N} f_j^* \eta_j \right\} |\Omega\rangle \\
&\quad = \exp\left\{ \sum_{j=1}^{2N} f_j^* \xi_j \right\} |\Omega\rangle = |\xi\rangle,
\end{aligned}$$

where on the first step we used (2.23), and on the third one we used (2.25), with $M = 2N$, $\mathcal{A} = I$, and $\xi_j^* = f_j^*$. Calculations in the case of the bra-state are similar.

Let us now consider the expression for the partition function, (2.16). Inserting the unity decomposition between the vectors and the operator T^N, we have the expression

$$\begin{aligned}
Z = \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j \mathrm{d}\xi_j^* \mathrm{d}\xi_j)\, \exp\left\{ -\sum_{j=1}^{2N} \eta_j^* \xi_j - \sum_{j=1}^{2N} \xi_j^* \eta_j \right\} \\
\times \langle\Lambda|\xi\rangle\langle\eta^*|T^N|\eta\rangle\langle\xi^*|\Lambda'\rangle. \quad (2.26)
\end{aligned}$$

From (2.12) and (2.22), we have:

$$\langle\Lambda|\eta\rangle = \xi_N \cdots \xi_2 \xi_1, \qquad \langle\eta^*|\Lambda'\rangle = \xi_{N+1}^* \xi_{N+2}^* \cdots \xi_{2N}^*.$$

Writing the product of these two expressions as $\prod_{j=1}^{N} \xi_j \xi_{N+j}^*$ and also taking into account that $\prod_{j=1}^{2N} \mathrm{d}\xi_j^* \mathrm{d}\xi_j = \prod_{j=1}^{N} \mathrm{d}\xi_{j+N} \mathrm{d}\xi_j^* \mathrm{d}\xi_{j+N}^* \mathrm{d}\xi_j$, we first evaluate in (2.26) the integrals over the variables ξ_j and ξ_{j+N}^*, $j = 1, \ldots, N$, and next

over the variables ξ_j^* and ξ_{j+N}, $j = 1, \dots, N$, thus obtaining

$$\begin{aligned} Z &= \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j \mathrm{d}\xi_j^* \mathrm{d}\xi_j) \prod_{j=1}^{N} (\xi_j \xi_{N+j}^*) \\ &\qquad \times \exp\left\{ -\sum_{j=1}^{2N} \xi_j^* \eta_j - \sum_{j=1}^{2N} \eta_j^* \xi_j \right\} \langle \eta^* | T^N | \eta \rangle \\ &= \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j) \prod_{j=1}^{N} (\mathrm{d}\xi_{j+N} \mathrm{d}\xi_j^*) \\ &\qquad \times \exp\left\{ -\sum_{j=1}^{N} \xi_j^* \eta_j - \sum_{j=1}^{N} \eta_{j+N}^* \xi_{j+N} \right\} \langle \eta^* | T^N | \eta \rangle \\ &= \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j) \prod_{j=1}^{N} (\eta_j \eta_{N+j}^*) \, \langle \eta^* | T^N | \eta \rangle. \end{aligned} \tag{2.27}$$

Hence, we need only to write $\langle \eta^* | T^N | \eta \rangle$ as a Grassmann integral.

To derive the Grassmann integral formula for this matrix element, we can insert the unity decompositions between all the factors T in T^N, thus reducing the task to finding such a formula for $\langle \eta^* | T | \eta \rangle$. In computing of this matrix element, a crucial role plays the fact the operator T is represented as a supertrace, see (2.20). For a convenience, let us introduce functions which explicitly describe functional dependence of the operators $P_{j,k}$ and $R_{j,k}$ on the fermion operators, as given by (2.14) and (2.17), respectively, by defining them as $P(f_j^*, f_j; f_k^*, f_k) = P_{jk}$ and $R(f_j^*, f_j; f_k^*, f_k) = R_{jk}$. Then, by using (2.21), the matrix element $\langle \eta^* | T | \eta \rangle$ can be written in the form

$$\begin{aligned} \langle \eta^* | T | \eta \rangle = \exp\left\{ \sum_{j=1}^{2N} \eta_j^* \eta_j \right\} \\ \times \mathrm{str}_0 \left\{ P(\eta_{2N}^*, \eta_{2N}; f_0^*, f_0) \cdots P(\eta_{N+1}^*, \eta_{N+1}; f_0^*, f_0) \right. \\ \left. \times R(\eta_N^*, \eta_N; f_0^*, f_0) \cdots R(\eta_1^*, \eta_1; f_0^*, f_0) \right\}. \end{aligned} \tag{2.28}$$

Recalling that $\mathrm{str}_0(\dots) = \langle \Omega_0 | \dots | \Omega_0 \rangle - \langle \Omega_0 | f_0 \dots f_0^* | \Omega_0 \rangle$, where $|\Omega_0\rangle$ is the vacuum of the auxiliary fermion, it is fairly easy to see that the supertrace

can be written as the following integral over a pair of Grassmann variables:

$$\mathrm{str}_0(\dots) = \int \mathrm{d}\phi^* \mathrm{d}\phi \, \mathrm{e}^{-\phi^*\phi} \, \langle \Omega_0 | \mathrm{e}^{\phi^* f_0} (\dots) \mathrm{e}^{f_0^* \phi} | \Omega_0 \rangle.$$

Since in (2.28) the Grassmann variables can be seen as external variables, we can now use the unity decomposition in its one-mode version, for the auxiliary fermion. In this way one arrives at the following representation

$$\begin{aligned}\langle \eta^* | T | \eta \rangle = \exp \left\{ \sum_{j=1}^{2N} \eta_j^* \eta_j \right\} \\ \times \int \prod_{j=1}^{2N} (\mathrm{d}\phi_j^* \mathrm{d}\phi_{j+1}) \exp \left\{ - \sum_{j=1}^{2N} \phi_j^* \phi_{j+1} + \sum_{j=1}^{2N} \phi_j^* \phi_j \right\} \\ \times \prod_{j=1+N}^{2N} P(\eta_j^*, \eta_j; \phi_j^*, \phi_j) \prod_{j=1}^{N} R(\eta_j^*, \eta_j; \phi_j^*, \phi_j),\end{aligned}$$

where $\phi_{2N+1} := \phi_1$. The P- and R-functions can be represented as exponentials,

$$\begin{aligned} P(\eta^*, \eta; \phi^*, \phi) &= 1 - \eta^*\eta - \phi^*\phi + \phi^*\eta + \eta^*\phi \\ &= \exp\left\{ -\eta^*\eta - \phi^*\phi + \phi^*\eta + \eta^*\phi \right\}, \\ R(\eta^*, \eta; \phi^*, \phi) &= a + c(\eta^*\phi + \phi^*\eta) + (b-a)(\eta^*\eta + \phi^*\phi) - 2b\, \eta^*\eta\, \phi^*\phi \\ &= a \, \exp\left\{ -\eta^*\eta - \phi^*\phi + V(\eta^*, \eta; \phi^*, \phi) \right\}, \end{aligned}$$

where the function $V(\eta^*, \eta; \phi^*, \phi)$ is given by

$$\begin{aligned} V(\eta^*, \eta; \phi^*, \phi) = \frac{b}{a}(\eta^*\eta + \phi^*\phi) + \frac{c}{a}(\eta^*\phi + \phi^*\eta) \\ - \frac{a^2 + b^2 - c^2}{a^2} \eta^*\eta\, \phi^*\phi. \end{aligned} \quad (2.29)$$

Taking into account that $\prod_{j=1}^{2N} \mathrm{d}\phi_j^* \mathrm{d}\phi_{j+1} = -\prod_{j=1}^{2N} \mathrm{d}\phi_j^* \mathrm{d}\phi_j$, we thus arrive at following Grassmann integral formula:

$$\langle \eta^* | T | \eta \rangle = -a^N \int \prod_{j=1}^{2N} (\mathrm{d}\phi_j^* \mathrm{d}\phi_j) \exp\left\{ -\sum_{j=1}^{2N} \phi_j^* \phi_{j+1} + \sum_{j=1}^{N} V(\eta_j^*, \eta_j; \phi_j^*, \phi_j) + \sum_{j=N+1}^{2N} \phi_j^* \eta_j + \sum_{j=N+1}^{2N} \eta_j^* \phi_j \right\}. \quad (2.30)$$

Note that this integral is Gaussian, and hence it can be evaluated. However, it is useful to keep it temporary in the present form; we will use the evaluated form in next section.

The further derivation is straightforward. Using (2.24) and substituting (2.30), we arrive at the following representation for the partition function in terms of a 2D lattice Grassmann integral:

$$\begin{aligned} Z = (-1)^N a^{N^2} \int \prod_{i=1}^{N} \prod_{j=1}^{2N} (\mathrm{d}\eta_{ij}^* \mathrm{d}\eta_{ij} \mathrm{d}\phi_{ij}^* \mathrm{d}\phi_{ij}) \prod_{j=1}^{N} (\eta_{1,j} \eta_{N,N+j}^*) \\ \times \exp\left\{ -\sum_{i=1}^{N} \sum_{j=1}^{2N} \phi_{ij}^* \phi_{i,j+1} - \sum_{i=1}^{N-1} \sum_{j=1}^{2N} \eta_{ij}^* \eta_{i+1,j} \right. \\ + \sum_{i,j=1}^{N} V(\eta_{ij}^*, \eta_{ij}; \phi_{ij}^*, \phi_{ij}) \\ \left. + \sum_{i=1}^{N} \sum_{j=N+1}^{2N} \phi_{ij}^* \eta_{ij} + \sum_{i=1}^{N} \sum_{j=N+1}^{2N} \eta_{ij}^* \phi_{ij} \right\}. \quad (2.31) \end{aligned}$$

Here $\phi_{i,2N+1} := \phi_{i,1}$, $i = 1, \ldots, N$. The expression in (2.31) can be simplified by evaluating the integrals over the variables with the second subscript taking values $j = N+1, \ldots, 2N$. Such evaluation can be done recursively using that, e.g., $\int \exp\{-\eta^* \phi\} \mathrm{d}\phi = \eta^*$, and that single Grassmann variable standing in the integrand can be treated as the delta-function, $\eta^* = \delta(\eta^*)$. As a result, we readily obtain

$$Z = a^{N^2} \int \prod_{i,j=1}^{N} (\mathrm{d}\eta_{ij}^* \mathrm{d}\eta_{ij} \mathrm{d}\phi_{ij}^* \mathrm{d}\phi_{ij}) \prod_{k=1}^{N} (\eta_{1,k} \phi_{N-k+1,N}^*) \exp\{S_{\mathrm{SV}}\}, \quad (2.32)$$

where we have denoted

$$S_{\mathrm{SV}} = -\sum_{i=1}^{N}\sum_{j=1}^{N-1}\phi^*_{ij}\phi_{i,j+1} - \sum_{i=1}^{N-1}\sum_{j=1}^{N}\eta^*_{ij}\eta_{i+1,j} + \sum_{i,j=1}^{N} V(\eta^*_{ij},\eta_{ij};\phi^*_{ij},\phi_{ij}). \quad (2.33)$$

Formula (2.32) essentially solves the problem of fermionization of the six-vertex model. In (2.33) each term has a simple relation with the elements of the $N \times N$ lattice. The first and the second term in (2.33) are related to the vertical and the horizontal edges (links) of the lattice, respectively, while the third one is related to the vertices; the index i labels the vertical lines (enumerating them from right to left) and the index j the horizontal ones (from top to bottom). Note that (2.32) can also be regarded as the Grassmann integral form of writing for (2.16), with the operator T^N given by (2.19).

We conclude our discussion of the fermionization, by noting that (2.32) can also be written as an integral of some exponential, i.e., in the form, canonical for a partition function,

$$Z = a^{N^2}\int \prod_{i,j=1}^{N}(\mathrm{d}\eta^*_{ij}\mathrm{d}\eta_{ij}\mathrm{d}\phi^*_{ij}\mathrm{d}\phi_{ij}) \prod_{k=1}^{N}(\mathrm{d}\xi^*_k\mathrm{d}\xi_k)\,\exp\{S_{\mathrm{SV}}+S_{\mathrm{DW}}\}, \quad (2.34)$$

where

$$S_{\mathrm{DW}} = \sum_{k=1}^{N}\xi^*_k\eta_{1,k} - \sum_{k=1}^{N}\phi^*_{N-k+1,N}\xi_k.$$

In (2.34), the term S_{DW} is responsible for accounting domain wall boundary conditions, by imposing constraints on the variables of the six-vertex model action S_{SV}. The auxiliary variables ξ^*_k and ξ_k, entering here only S_{DW}, can be seen as the corresponding Lagrange multipliers.

§3. The free-fermion model

In this section we calculate explicitly the partition and one-point functions of the model at its free-fermion point. Recall that the free-fermion condition means that the weights of the six-vertex model satisfy the relation

$$a^2 + b^2 = c^2. \quad (3.1)$$

This condition leaves a one-parameter freedom in the choice of the weights. At certain stage of calculations in this section, we switch to the following parameterization

$$\frac{a}{c} = \sqrt{1-\alpha}, \qquad \frac{b}{c} = \sqrt{\alpha}. \tag{3.2}$$

Here α is some parameter, $0 < \alpha < 1$.

3.1. Evaluation of the integrals. Within the Grassmann integral formalism developed in the previous section, the free-fermion condition guarantees vanishing of the quartic terms in the action $S_{\rm SV}$, see (2.29) and (2.33). Hence, the six-vertex model in this case is mapped to a free-fermion theory on a finite 2D lattice. Basing on the Grassmann integral (2.34), the generating functional for correlation functions can be constructed and, since all integrations are Gaussian, formally evaluated. However, this route, standard for a free-fermion theory, encounters some complications due to the presence of the constraints, induced by the domain wall boundary conditions. Since our goal here is just limited to the derivation of the one-point function, we have found that the calculations along this approach are too cumbersome; we will present them elsewhere.

Instead, we use here a different method of treating the Grassmann integral representations of the previous Section, which exploits certain property of the operator T, valid only for the free-fermion model. In fact, the whole approach appears to be rather general and it can be applied for computing some multi-point correlation functions as well. One of the features of this approach is that to compute the correlation functions one has to invert some matrix just of size $N \times N$ (rather than of size $(2N^2+N)\times(2N^2+N)$, in the case of the integral (2.34)). Moreover, this task turns out to be trivial, due to factorization of this matrix as the product of a lower-triangular, diagonal, and upper-triangular matrices.

Our starting point is that in the free-fermion model, as it can be easily seen from (2.30) and (2.25), the operator T possesses the property that its matrix element between the two Grassmann coherent states is Gaussian:

$$\langle \eta^* | T | \eta \rangle = a^N \exp\{\eta^* \mathcal{T} \eta\}. \tag{3.3}$$

Here $\mathcal{T}$ denotes a $2N \times 2N$ matrix and we have introduced the shorthand notation $\eta^* \mathcal{T} \eta := \sum\limits_{j,l} \eta_j^* \mathcal{T}_{jl} \eta_l$; for a later convenience, we also adopt here the convention that $2N \times 2N$ matrices are denoted by calligraphic letters (except the identity matrix, for which the standard notation I will be

used). We obtain the matrix $\mathcal{T}$ below; in our present discussion its explicit form is irrelevant.

The crucial property of the operators whose matrix elements between the Grassmann coherent states are Gaussian, like in (3.3), is that the matrix elements of their products are also Gaussian. Namely, let operators $T_1, \dots, T_m$ be such that $\langle \eta^* | T_k | \eta \rangle = \varrho_k \exp\{\eta^* \mathcal{T}_k \eta\}$, $k = 1, \dots, m$, where $\mathcal{T}_k$ are some matrices and ϱ_k are some normalization constants (numbers). Then,

$$\langle \eta^* | T_1 \cdots T_m | \eta \rangle = \prod_{k=1}^{m} \varrho_k \cdot \exp\{\eta^* \mathcal{T}_1 \cdots \mathcal{T}_m \eta\}. \tag{3.4}$$

Indeed, let us consider the case $m = 2$. Inserting the unity decomposition (2.24), we obtain

$$\begin{aligned} \langle \eta^* | T_1 T_2 | \eta \rangle &= \int \prod_{j=1}^{2N} (\mathrm{d}\phi_j^* \mathrm{d}\phi_j) \exp\left\{-\phi^* I \phi\right\} \langle \eta^* | T_1 | \phi \rangle \langle \phi^* | T_2 | \eta \rangle \\ &= \varrho_1 \varrho_2 \int \prod_{j=1}^{2N} (\mathrm{d}\phi_j^* \mathrm{d}\phi_j) \exp\left\{-\phi^* I \phi + \eta^* \mathcal{T}_1 \phi + \phi^* \mathcal{T}_2 \eta\right\} \\ &= \varrho_1 \varrho_2 \exp\left\{\eta^* \mathcal{T}_1 \mathcal{T}_2 \eta\right\}. \end{aligned} \tag{3.5}$$

At the last step we used (2.25) with $\mathcal{A} = I$, $\xi_j = (\mathcal{T}_2 \eta)_j$, and $\xi_j^* = (\eta^* \mathcal{T}_1)_j$. Since in (2.25) the external variables ξ_j and ξ_j^* are arbitrary, the matrices $\mathcal{T}_1$ and $\mathcal{T}_2$ here may also be arbitrary. Applying induction in the number of operators, or repeatedly using (3.5), we arrive to (3.4).

Due to (3.4), from (3.3) we thus have:

$$\langle \eta^* | T^m | \eta \rangle = a^{mN} \exp\{\eta^* \mathcal{T}^m \eta\}, \qquad m \in \mathbb{N}. \tag{3.6}$$

For example, the partition function can be computed in terms of the Nth power of the matrix $\mathcal{T}$, simply using (2.27). Indeed, let us define $N \times N$ matrices A_m, B_m, C_m, and D_m to be 2×2 block entries of $2N \times 2N$ matrix being the mth power ($m = 1, 2, \dots$) of the matrix $\mathcal{T}$:

$$\mathcal{T}^m = \begin{pmatrix} A_m & B_m \\ C_m & D_m \end{pmatrix}. \tag{3.7}$$

Then, setting $m = N$ in (3.6) and substituting it in (2.27), we obtain

$$\begin{aligned} Z &= a^{N^2} \int \prod_{j=1}^{2N} (\mathrm{d}\eta_j^* \mathrm{d}\eta_j) \prod_{j=1}^{N} (\eta_j \eta_{N+j}^*) \, \exp\{\eta^* \mathcal{T}^N \eta\} \\ &= a^{N^2} \int \prod_{j=1}^{N} (\mathrm{d}\eta_{N+j} \mathrm{d}\eta_j^*) \exp\left\{ \sum_{j,l=1}^{N} \eta_j^* (B_N)_{jl} \eta_{N+l} \right\} \\ &= a^{N^2} \det B_N. \end{aligned} \tag{3.8}$$

Below we show that the matrix B_N is given as a product of a lower-triangular, diagonal, and upper-triangular matrices; hence, its determinant (and the inverse matrix) can be easily evaluated.

The above formulas can be directly generalized on the case of correlation functions. Consider, for example, the correlation functions, which describe probabilities of the same state (e.g., right arrow) on given horizontal edges; the simplest example of such correlation function is provided by the one-point function (2.2). In the operator formulation of the previous section, all these correlation functions can defined by suitably inserting the local fermion vacuum state projectors, i.e., the operators $f_k f_k^*$, between the powers of the operator T. Their matrix elements in the Grassmann coherent states are

$$\begin{aligned} \langle \eta^* | f_k f_k^* | \eta \rangle &= \langle \eta^* | (1 - f_k^* f_k) | \eta \rangle \\ &= (1 - \eta_k^* \eta_k) \langle \eta^* | \eta \rangle \\ &= \exp\{\eta^* \mathcal{J}_k \eta\}, \end{aligned} \tag{3.9}$$

where the matrix $\mathcal{J}_k$ differs from the identity matrix just in single entry, with 0 instead of 1 at the kth position on the diagonal,

$$(\mathcal{J}_k)_{ij} = \delta_{ij} - \delta_{ik}\delta_{jk}. \tag{3.10}$$

The Gaussian form of (3.9) implies that using (3.4) one can reduce calculation of the correlation functions to evaluation of the integrals, similar to (3.8).

In this way, the correlation functions can be expressed as as some $N \times N$ determinants. In particular, in the one-point function case such determinant involves the matrix B_N plus some matrix of rank one. Since the matrix B_N has a peculiar factorized structure and admits explicit inversion, this determinant can be easily evaluated. Below in this section we provide the details of this calculation.

3.2. Calculation of the matrices. Here we compute matrices A_m, B_m, C_m, and D_m for $m = 1, \dots, N$ explicitly. We show also that the matrix B_N is equal to a product of a lower-triangular, diagonal, and upper-triangular matrices.

Let us compute the matrix $\mathcal{T}$. Imposing the condition (3.1) and performing the change of the integration variables $\phi_{j+1} \mapsto \phi_j$, one can write (2.30) as follows:

$$\langle \eta^* | T | \eta \rangle = a^N \exp\{\eta^* \mathcal{Y} \eta\} \times \int \prod_{j=1}^{2N} (\mathrm{d}\phi_j^* \mathrm{d}\phi_j) \exp\{-\phi^*(I - \mathcal{C}\mathcal{Y})\phi + \phi^* \mathcal{C}\mathcal{X}\eta + \eta^* \mathcal{X}\phi\}. \quad (3.11)$$

Here $\mathcal{C}$ is the $2N \times 2N$ cyclic shift matrix

$$\mathcal{C} = \begin{pmatrix} & & & 1 \\ & & & \\ & & & \\ & 2N-1 & & \end{pmatrix},$$

and $\mathcal{X}$, $\mathcal{Y}$ are the diagonal matrices

$$\mathcal{X} = \begin{pmatrix} \gamma I_N & & \\ & & I_N \end{pmatrix}, \qquad \mathcal{Y} = \begin{pmatrix} \beta I_N & & \\ & & \end{pmatrix},$$

where the following notations are used

$$\beta = \frac{b}{a}, \qquad \gamma = \frac{c}{a}.$$

The subscript of I indicates the size of the identity matrix. Evaluation of the integral in (3.11) by (2.25) gives us formula (3.3), where

$$\mathcal{T} = \mathcal{Y} + \mathcal{X}(I - \mathcal{C}\mathcal{Y})^{-1} \mathcal{C}\mathcal{X}.$$

The matrix $\mathcal{C}\mathcal{Y}$ is lower-triangular, and, moreover, nilpotent, $(\mathcal{C}\mathcal{Y})^{N+1} = 0$, hence $(I - \mathcal{C}\mathcal{Y})^{-1}$ can be easily evaluated. In this way one can compute all entries of the matrix $\mathcal{T}$.

For the matrices A_1, B_1, C_1, and D_1, defined as blocks of the matrix $\mathcal{T}$, see (3.7), the following holds. The matrix A_1 is lower-triangular, with the entries

$$(A_1)_{ij} = \begin{cases} 0 & i < j, \\ \beta & i = j, \\ \gamma^2 \beta^{i-j-1} & i > j. \end{cases}$$

The matrix B_1 has nonzero entries only in the last column, while the matrix C_1 has those only in the first row:

$$(B_1)_{ij} = \gamma\beta^{i-1}\delta_{j,N}, \qquad (C_1)_{ij} = \gamma\beta^{N-j}\delta_{1,i}.$$

The matrix D_1 is

$$D_1 = \begin{pmatrix} & \beta^N \\ I_{N-1} & \end{pmatrix}.$$

A peculiar structure of these matrices allows us recursively compute matrices A_m, B_m, C_m, and D_m for $m = 2, \dots, N$. Let us introduce column vectors $\vec{v}$ and $\vec{u}$, such that $\vec{v}$ denotes the last column of the matrix B_1, and $\vec{u}^T$ coincides with the first row of the matrix C_1,

$$v_i = \gamma\beta^{i-1}, \qquad u_i = \gamma\beta^{N-i} = v_{N-i+1}.$$

Consider case $m = 2$. Using $A_2 = A_1^2 + B_1C_1$, and observing that $B_1C_1 = 0$, we find that $A_2 = A_1^2$. Using $B_2 = A_1B_1 + B_1D_1$, we find that the matrix B_2 has last two columns nonzero, the $(N-1)$th column is given by $\vec{v}$, and the Nth one by $A_1\vec{v}$. Using $C_2 = C_1A_1 + D_1C_1$, we find that the matrix C_2 has two first rows nonzero, the first is given by $\vec{u}^TA_1$, and the second one by $\vec{u}^T$. Using $D_2 = D_1^2 + C_1B_1$, we obtain

$$D_2 = \begin{pmatrix} & \beta^N & N\gamma^2\beta^{N-1} \\ & & \beta^N \\ I_{N-2} & & \end{pmatrix}.$$

Similarly, one can consider the cases $m = 3$, $m = 4$, etc. From these considerations it turns out that a general result can be easily guessed, and, in fact, proved by induction.

Namely, for $m = 1, \dots, N$, the $N \times N$ matrices A_m, B_m, C_m, and D_m, defined by (3.7), are given by the following formulas. The matrix A_m is just mth power of the matrix A_1,

$$A_m = A_1^m. \tag{3.12}$$

The matrix B_m has its last m columns nonzero,

$$(B_m)_{ij} = \begin{cases} 0 & j = 1, \dots, N-m, \\ \left(A_1^{j-N+m-1}\vec{v}\right)_i & j = N-m+1, \dots, N. \end{cases}$$

The matrix C_m has its first m rows nonzero,

$$(C_m)_{ij} = \begin{cases} \left(\vec{u}^TA_1^{m-i}\right)_j & i = 1, \dots, m \\ 0 & i = m+1, \dots, N. \end{cases}$$

The matrix D_m itself has the following block structure

$$D_m = \begin{pmatrix} & S_m \\ I_{N-m} & \end{pmatrix},$$

where the nontrivial block, S_m, of size $m \times m$, is an upper-triangular matrix with the entries of the transpose of A_N,

$$(S_m)_{ij} = (A_N)_{ji}, \qquad i, j = 1, \ldots, m.$$

Below we discuss some other representations and properties of these matrices.

First of all, we mention that these matrices can be conveniently written in terms of lower- and upper-triangular matrices E_- and E_+, respectively, which have entries

$$(E_\pm)_{ij} = \delta_{i,j\mp 1}.$$

These matrices are nilpotent, $E_\pm^N = 0$. Using this property, we can use for the matrix A_m formula (3.12) where the matrix A_1 is given by

$$A_1 = \beta I + \gamma^2 E_- \sum_{k=0}^{N-2} \beta^k E_-^k = \frac{\beta I + E_-}{I - \beta E_-}. \tag{3.13}$$

Other matrices can be written as follows

$$\begin{aligned} B_m &= B_N E_+^{N-m}, \\ C_m &= E_+^{N-m} C_N, \\ D_m &= E_-^m + A_N^T E_+^{N-m}, \end{aligned} \tag{3.14}$$

where the matrices B_N and C_N can be defined as matrices with entries $(B_N)_{ij} = \left(A_1^{j-1}\vec{v}\right)_i$ and $(C_N)_{ij} = \left(\vec{u}^T A_1^{N-i}\right)_j$, respectively. Note, that if we introduce the matrix of transposition of N elements, $\Omega_{ij} = \delta_{i,N-j+1}$, then, due to the relations $\vec{u} = \Omega\vec{v}$ and $A_N^T = \Omega A_N \Omega$, we can connect matrices B_N and C_N by the relation $C_N = \Omega B_N^T \Omega$; similar relation also holds for B_m and C_m.

Let us now consider more attentively the matrix B_N. Using (3.13) in $(B_N)_{ij} = \left(A_1^{j-1}\vec{v}\right)_i$, we obtain, by direct calculation of its entries and making use of standard relations for hypergeometric series, the expression

$$(B_N)_{ij} = \gamma\beta^{i+j-2} \sum_{p=0}^{\min(i-1,j-1)} \binom{i-1}{p}\binom{j-1}{p}\left(\frac{\gamma^2}{\beta^2}\right)^p, \tag{3.15}$$

where the standard notation for the binomial coefficient is used (recall that $\binom{z}{k} = \frac{z(z-1)\cdots(z-k+1)}{k!}$ if k is a positive integer, $\binom{z}{k} = 1$, if $k = 0$, and $\binom{z}{k} = 0$ if k is a negative integer). This formula implies that B_N can be represented as a product of lower-triangular, diagonal, and upper-triangular matrices. Indeed, let us introduce $N \times N$ matrices J_-, J_0, and J_+, with entries

$$(J_-)_{ij} = j\,\delta_{i,j+1}, \qquad (J_0)_{ij} = \left(j - \frac{1}{2}\right)\delta_{ij}, \qquad (J_+)_{ij} = i\,\delta_{i,j-1}.$$

For an arbitrary parameter z, we have

$$\left(\mathrm{e}^{zJ_-}\right)_{ij} = \binom{i-1}{j-1} z^{i-j}, \qquad \left(\mathrm{e}^{zJ_+}\right)_{ij} = \binom{j-1}{i-1} z^{j-i}. \tag{3.16}$$

Thus formula (3.15) can be written in as the following matrix product:

$$B_N = \mathrm{e}^{\beta J_-}\,\gamma^{2J_0}\,\mathrm{e}^{\beta J_+}. \tag{3.17}$$

Since formula (3.17) is crucial for subsequent calculations, it is useful to mention here that it can also be obtained without a direct calculation of entries of the matrix B_N (i.e., without (3.15)), but just using the following commutation relations

$$J_0 J_\pm = J_\pm (J_0 \mp I).$$

Matrices $E_\pm$ admit the following representations

$$E_+ = \left(\tfrac{1}{2}\,I + J_0\right)^{-1} J_+, \qquad E_- = J_- \left(\tfrac{1}{2}\,I + J_0\right)^{-1}.$$

We also have

$$\mathrm{e}^{zJ_\pm} J_0 \mathrm{e}^{-zJ_\pm} = J_0 \pm zJ_\pm,$$

and therefore

$$\mathrm{e}^{zJ_\pm} E_\pm \mathrm{e}^{-zJ_\pm} = \frac{E_\pm}{I \pm zE_\pm}. \tag{3.18}$$

From the last formula follows, after comparing it with (3.13), that the matrix A_1 can also be written in the form

$$A_1 = \mathrm{e}^{\beta J_-}\left(\beta I + \gamma^2 E_-\right)\mathrm{e}^{-\beta J_-}. \tag{3.19}$$

Coming back to matrix B_N, we see that since $(B_N)_{ij} = (A_1^{j-1}\vec{v})_i$, formula (3.19) suggests us to consider the matrix $\exp\{-\beta J_-\}B_N$. Using (3.16),

and taking into account that vector $\vec{v}$ can be written as $\vec{v} = \exp\{\beta J_-\}\vec{w}$, where $w_i = \gamma\delta_{1,i}$, for entries of this matrix we have the expression

$$\left(\mathrm{e}^{-\beta J_-} B_N\right)_{ij} = \gamma\left[\left(\beta I + \gamma^2 E_-\right)^{j-1}\right]_{i,1} = \binom{j-1}{i-1}\beta^{j-1}\gamma^{2i-1}.$$

Again using (3.16), one can find that the last expression exactly coincides with $\left(\gamma^{2J_0}\mathrm{e}^{\beta J_+}\right)_{ij}$, that proves (3.17).

As a simple application of (3.17), we obtain

$$\det B_N = \gamma^{N^2},$$

and hence from (3.8) we recover the well-known result (see, e.g., [14]) for the partition function:

$$Z = a^{N^2}\gamma^{N^2} = c^{N^2}.$$

3.3. The one-point function. Applying the formalism of fermionic operators of Sect. 2.2 to the one-point function $G(m,n) = G_{\rightarrow}(m,n)$, defined in (2.2), we can write it as the following matrix element

$$G(m,n) = Z^{-1}\langle\Lambda|T^m f_{N-n+1} f^*_{N-n+1} T^{N-m}|\Lambda'\rangle.$$

The analogue of the Grassmann integral formula (2.27) for this matrix element reads

$$G(m,n) = Z^{-1}\int\prod_{j=1}^{2N}(\mathrm{d}\eta^*_j\mathrm{d}\eta_j)\prod_{j=1}^{N}(\eta_j\eta^*_{N+j}) \times\langle\eta^*|T^m f_{N-n+1} f^*_{N-n+1} T^{N-m}|\eta\rangle. \quad (3.20)$$

Using (3.4) and (3.9), we have

$$\langle\eta^*|T^m f_{N-n+1} f^*_{N-n+1} T^{N-m}|\eta\rangle = a^{N^2}\exp\left\{\eta^*\mathcal{T}^m\mathcal{J}_{N-n+1}\mathcal{T}^{N-m}\eta\right\}.$$

Recalling that n can take values $1,\dots,N$, we consider here the matrix $\mathcal{J}_k$ only for $k = 1,\dots,N$. This matrix, see (3.10), has the block structure

$$\mathcal{J}_k = \begin{pmatrix} I_N - \Pi_k & \\ & I_N \end{pmatrix}, \qquad k = 1,\dots,N,$$

where Π_k denotes the $N\times N$ diagonal matrix with single nonzero entry, equal to one, standing at the kth position:

$$(\Pi_k)_{ij} = \delta_{ik}\delta_{jk}.$$

Hence, for $k = 1, \dots, N$, the matrix $\mathcal{T}^m \mathcal{J}_k \mathcal{T}^{N-m}$ has the form

$$\mathcal{T}^m \mathcal{J}_k \mathcal{T}^{N-m} = \mathcal{T}^N - \begin{pmatrix} A_m \Pi_k A_{N-m} & A_m \Pi_k B_{N-m} \\ C_m \Pi_k A_{N-m} & C_m \Pi_k B_{N-m} \end{pmatrix}.$$

Evaluation of the Grassmann integral in (3.20) yields

$$G(m,n) = a^{N^2} Z^{-1} \det \left(B_N - A_m \Pi_{N-n+1} B_{N-m} \right).$$

Since Π_k is a rank one matrix, we have

$$\begin{aligned} G(m,n) &= 1 - \operatorname{tr} \left(B_N^{-1} A_m \Pi_{N-n+1} B_{N-m} \right) \\ &= 1 - \left(B_{N-m} B_N^{-1} A_m \right)_{N-n+1, N-n+1} \\ &= \left(B_m B_N^{-1} A_{N-m} \right)_{nn}. \end{aligned}$$

Here, at the last step we have exploited relation (2.3).

To progress further, let us make use of explicit forms of the matrices involved. Using (3.12) and (3.14), we get

$$G(m,n) = \left(B_N E_+^{N-m} B_N^{-1} A_1^{N-m} \right)_{nn}.$$

Substituting (3.17) and (3.19) into this expression, using relation (3.18) and the relation

$$z^{J_0} E_\pm z^{-J_0} = z^{\mp 1} E_\pm,$$

and also switching to the parameter $\alpha = \beta^2/\gamma^2$ entering the parameterization of the free-fermion model weights (3.2), we can write

$$\begin{aligned} G(m,n) &= \left(\mathrm{e}^{\beta J_-} \gamma^{2J_0} \mathrm{e}^{\beta J_+} E_+^{N-m} \mathrm{e}^{-\beta J_+} \gamma^{-2J_0} (\beta I + \gamma^2 E_-)^{N-m} \mathrm{e}^{-\beta J_-} \right)_{nn} \\ &= \left(\mathrm{e}^{\beta J_-} \left(\frac{E_+}{\gamma^2 I + \beta E_+} \right)^{N-m} \left(\beta I + \gamma^2 E_- \right)^{N-m} \mathrm{e}^{-\beta J_-} \right)_{nn} \\ &= \left(\mathrm{e}^{\alpha J_-} \left(\frac{E_+}{I + E_+} \right)^{N-m} (I + E_-)^{N-m} \, \mathrm{e}^{-\alpha J_-} \right)_{nn}. \end{aligned} \tag{3.21}$$

The last formula shows that the one-point function is some polynomial in the variable α.

To derive an explicit expression for this polynomial, let us consider the matrices in the last expression. Using

$$\left(\left(\frac{E_+}{I + E_+} \right)^p \right)_{ij} = (-1)^{i-j+p} \binom{j-i-1}{j-i-p}, \tag{3.22}$$

and

$$\big((I+E_-)^p\big)_{ij} = \binom{p}{i-j}, \tag{3.23}$$

where $p = 0, 1, \ldots, N-1$, we find that the entries of the matrix being the product of these two matrices are

$$\begin{aligned}\left(\left(\frac{E_+}{I+E_+}\right)^p (I+E_-)^p\right)_{ij} &= \sum_{k=1}^{i-k+p} \binom{k-i-1}{k-i-p}\binom{p}{k-j} \\ &= \sum_{q=0}^{N-p-i} (-1)^q \binom{p+q-1}{q}\binom{p}{p+q+i-j} \\ &= (-1)^{N-p-i}\binom{j-i-1}{N-i-p}\binom{N-i}{j-i}. \end{aligned} \tag{3.24}$$

In (3.24) the summation is done due to the formula

$$\sum_{q=0}^{r}(-1)^q \binom{p}{p+q-s}\binom{p+q-1}{p-1} = (-1)^r \binom{s-1}{r}\binom{p+r}{s}, \tag{3.25}$$

in which we set $r = N-p-i$ and $s = j-i$. Formula (3.25), in turn, can be seen as a special case of the Pfaff-Saalschütz summation formula for a terminating ${}_3F_2$-series,

$${}_3F_2\left(\begin{matrix}-r,\, a,\, b \\ c,\, a+b-c-r+1\end{matrix}\,\middle|\, 1\right) = \frac{(c-a)_r (c-b)_r}{(c)_r (c-a-b)_r},$$

upon setting $a = p$, $b = -s$, and taking the limit $c \to -r$.

Finally, using the expression for entries of the matrix $\exp\{\alpha J_-\}$, see (3.16), and collecting all terms of the same powers in α, we obtain

$$\begin{aligned} G(m,n) = (-1)^{m+n} \sum_{s=0}^{N-1}\sum_{r=0}^{m-1} &\binom{s-1}{r}\binom{N-m+r}{s}\binom{n-1}{m-r-1} \\ &\times \binom{m-r+s-1}{n-1}(-\alpha)^s. \end{aligned} \tag{3.26}$$

Note that, in this expression the summation over r can be extended to the value $N-1$, so that the coordinates m, n enter only the coefficients of the polynomial.

Representation for the one-point function (3.26) is interesting since it is possibly the simplest expression that can be derived. It can be used, for example, for numerical plots of the one-point function, see Fig. 3. We also

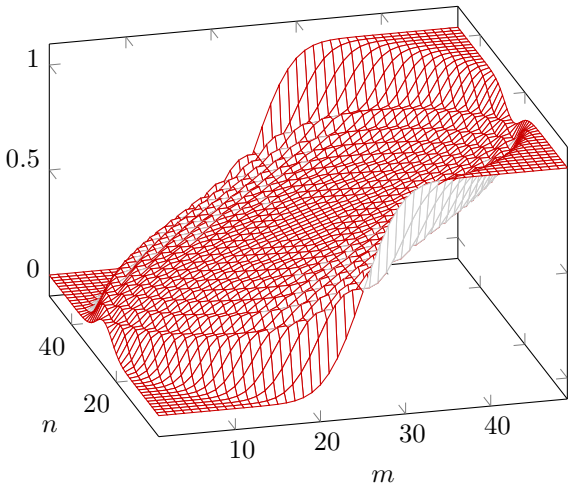

Figure 3. Plot of the one-point function $G(m,n)$, $N=50$, $\alpha=1/2$.

mention that (3.26) had been used by the authors of [19] in their testing of numerical simulations against exact results ([19], Ref. [13] therein).

Similar calculations of the one-point function can also performed, e.g., for the five-vertex model at its free-fermion point with the boundary conditions at which the model's configurations are in one-to-one correspondence with boxed plane partitions [39, 40].

3.4. Double integral formula. In addition to (3.26), the one-point function can be written as a double contour integral. Such a representation appears useful since it can used to study the one-point function in thermodynamic limit (i.e., in the limit of large lattice). We discuss this limit in the next section.

To derive such a representation we first rewrite the one-point function as a sum of the product of two Jacobi polynomials. Recall that for arbitrary real values of the parameters μ and ν the Jacobi polynomials $P_\ell^{(\mu,\,\nu)}(x)$ can be defined as follows:

$$P_\ell^{(\mu,\,\nu)}(x)=\sum_{k=0}^{\ell}\binom{\mu+\ell}{\ell-k}\binom{\mu+\nu+\ell+k}{k}\left(\frac{x-1}{2}\right)^k.$$

Using (3.16) and (3.22), we can write

$$\begin{aligned}\left(\mathrm{e}^{\alpha J_-}\left(\frac{E_+}{I+E_+}\right)^p\right)_{ij} &= \sum_{k=1}^{j-p}(-1)^{k-j+p}\binom{i-1}{k-1}\binom{j-k-1}{j-k-p}\alpha^{i-k}\\ &= \alpha^{i-j+p}\sum_{k=0}^{j-p-1}\binom{i-1}{j-p-k-1}\binom{k+p-1}{k}(-\alpha)^k\\ &= \alpha^{i-j+p}\,P_{j-p-1}^{(i-j+p,\,p-i)}\,(1-2\alpha)\,.\end{aligned}$$

Similarly, using (3.23), we have

$$\begin{aligned}\left((I+E_-)^p\mathrm{e}^{-\alpha J_-}\right)_{ij} &= \sum_{k=j}^{i}\binom{p}{i-k}\binom{k-1}{j-1}(-\alpha)^{k-j}\\ &= \sum_{k=0}^{i-j}\binom{p}{i-j-k}\binom{k+j-1}{k}(-\alpha)^k\\ &= P_{i-j}^{(p-i+j,\,p-i)}\,(1-2\alpha)\,.\end{aligned}$$

Hence, from (3.21) we obtain the following representation:

$$\begin{aligned}G(m,n) = \sum_{j=0}^{\min(m-1,N-n)} &\alpha^{n-m+j}\,P_{m-1-j}^{(n-m+j,\,N-m-n)}\,(1-2\alpha)\\ &\times\,P_{N-n-j}^{(n-m+j,\,-N+m+n-1)}\,(1-2\alpha)\,.\end{aligned}\tag{3.27}$$

Note that, since $\alpha\in[0,1]$, the argument of the Jacobi polynomials takes values in the interval $[-1,1]$.

Now, using the Rodrigues formula for Jacobi polynomials, and applying the Cauchy theorem, we can write them in terms of Schläfli integral:

$$P_\ell^{(\mu,\,\nu)}(x) = \frac{(-1)^\ell}{2^{\ell+1}\pi\mathrm{i}}(1-x)^{-\mu}(1+x)^{-\nu}\oint_{C_x}\frac{(1-t)^{\mu+\ell}(1+t)^{\nu+\ell}}{(t-x)^{\ell+1}}\,\mathrm{d}t.\tag{3.28}$$

Here C_x denotes a simple, closed, counterclockwise-oriented contour enclosing the point $t=x$, and lying in its small vicinity.

Using (3.28) and making change of the variable

$$t\mapsto z = -\frac{t-1+2\alpha}{\alpha(t+1)}$$

for the first polynomial in (3.27) we obtain

$$P_{m-j-1}^{(n-m+j,\ N-m+n)}(1-2\alpha) = \frac{1}{2\pi\mathrm{i}} \oint_{C_0} \frac{(z+1)^{n-1}}{(\alpha z+1)^{N-m} z^{m-j}} \mathrm{d}z.$$

A similar formula for the second polynomial can be obtained by replacing here $N, m, n \mapsto N', m', n'$, where $N' = N+1$, $m' = N-n+1$, $n' = N-m+1$.

Substituting the obtained integral representations for the Jacobi polynomials in (3.27) yields the following expression for the one-point function:

$$G(m,n) = \frac{\alpha^{m-n}}{(2\pi\mathrm{i})^2} \oint_{C_0} \mathrm{d}z_1 \oint_{C_0} \frac{(z_1+1)^{n-1}(z_2+1)^{N-m}}{(\alpha z_1+1)^{N-m} z_1^m (\alpha z_2+1)^n z_2^{N-n+1}} \times \frac{1-(\alpha z_1 z_2)^{\min(m,N-n+1)}}{1-\alpha z_1 z_2}\, \mathrm{d}z_2. \quad (3.29)$$

Obviously, in this formula the term $(\alpha z_1 z_2)^{\min(m,N-n+1)}$ does not contribute into the integral, since it cancels a pole at the origin, in either z_1 or z_2. Hence this term can be dropped. The final expression, after the change of the integration variables $z_1 \mapsto 1/z_1$ and $z_2 \mapsto z_2/\alpha$, reads

$$G(m,n) = \frac{1}{(2\pi\mathrm{i})^2} \oint_{C_\infty} \mathrm{d}z_1 \oint_{C_0} \frac{\Phi(z_1)}{\Phi(z_2)(z_1+1)z_2(z_1-z_2)}\, \mathrm{d}z_2. \quad (3.30)$$

Here

$$\Phi(z) = \frac{(z+1)^n z^{N-n}}{(z+\alpha)^{N-m}},$$

and C_∞ denotes the contour surrounding the point $z=\infty$, but it is clockwise-oriented (i.e., in (3.30) the both contours are oriented counterclockwise around the origin). We also mention that the order of integrals in (3.30) can be changed, as it can be done in (3.29).

§4. One-point function in the thermodynamic limit

In this section we study the one-point function $G(m,n)$ in the thermodynamic limit, $N \to \infty$, which we will regard as a continuous, or scaling, limit such that the $N \times N$ lattice is mapped onto the unit square $[0,1] \times [0,1]$. Let us set

$$x := \frac{m}{N}, \qquad y := \frac{n}{N}, \qquad x, y \in [0,1]. \quad (4.1)$$

Taking into account that the one-point function $G(m,n)$ depends on N as a parameter, $G(m,n) = G(m,n;N)$, we define the thermodynamic limit one-point function $g(x,y)$ as the result of the limit

$$g(x,y) = \lim_{N\to\infty} G([xN],[yN];N).$$

Our aim below is to compute $g(x,y)$ and to discuss approximate expressions for the one-point function $G(m,n)$, which are valid for large but finite values of N.

4.1. The arctic ellipse. We shall consider the limit $N \to \infty$ for the one-point function by applying the saddle point method to (3.30). It is useful to rewrite this formula in the form

$$G(m,n) = \frac{1}{(2\pi\mathrm{i})^2} \oint_{C_\infty} \mathrm{d}z_1 \oint_{C_0} \frac{\exp\{NS(z_1,z_2)\}}{(z_1+1)z_2(z_1-z_2)}\,\mathrm{d}z_2, \tag{4.2}$$

where $S(z_1,z_2) = F(z_1) - F(z_2)$ and

$$F(z) = y\ln(z+1) + (1-y)\ln z - (1-x)\ln(z+\alpha). \tag{4.3}$$

To apply the saddle-point method we start with determining zeros of the first derivative the function $F(z)$ governing the large N behavior of the exponent of the integrand. From (4.3) we find

$$F'(z) = \frac{xz^2 + (x-y+\alpha)z + (1-y)\alpha}{z(z+1)(z+\alpha)} = x\frac{(z-w_+)(z-w_-)}{z(z+1)(z+\alpha)}, \tag{4.4}$$

where the roots $w_\pm$ are functions of the coordinates, $w_\pm = w_\pm(x,y)$, and also of the parameter α,

$$w_\pm = \frac{y-x-\alpha \pm \sqrt{D(x,y)}}{2x}. \tag{4.5}$$

The discriminant $D(x,y)$ can be written as

$$D(x,y) = \alpha(1-\alpha)\left[\frac{(x+y-1)^2}{1-\alpha} + \frac{(x-y)^2}{\alpha} - 1\right]. \tag{4.6}$$

Thus. there are four saddle-points $(z_1,z_2) = (w_i,w_j)$, $i,j = +,-$, for the double integral saddle-point problem in our case. The saddle-point analysis will significantly depends on whether $w_\pm$ are real or complex, that is $D(x,y) > 0$ or $D(x,y) < 0$ in (4.5).

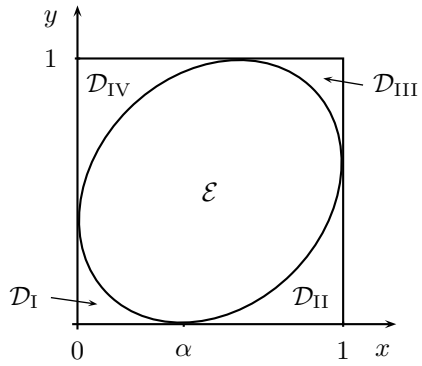

Figure 4. The ellipse $D(x,y)=0$ and five regions of the unit square.

Clearly, the condition $D(x,y)=0$ corresponds to the points $(x,y)\in[0,1]\times[0,1]$ which form an ellipse inscribed into the square. The parameter α describes a ratio of the radii of the ellipse; in particular, $\alpha=1/2$ corresponds to the circle.

The ellipse divides the unit square on five regions, see Fig. 4. Namely, there is a region in the interior of the ellipse, $\mathcal{E}$, where $D(x,y)<0$, and four regions in the exterior of it, $\mathcal{D}_{\rm I}$, $\mathcal{D}_{\rm II}$, $\mathcal{D}_{\rm III}$, and $\mathcal{D}_{\rm IV}$, where $D(x,y)>0$. Below we prove that if the point (x,y) lies in one of the regions $\mathcal{D}_{\rm I}$, $\mathcal{D}_{\rm II}$, $\mathcal{D}_{\rm III}$, $\mathcal{D}_{\rm IV}$, then the function $g(x,y)$ is equal to a constant, equal to 0 or 1 everywhere in the region, and if $(x,y)\in\mathcal{E}$ then the function $g(x,y)$ is a continuous function, taking intermediate values between 0 and 1. Since the values of $g(x,y)$ equal to 0 or 1 can be interpreted as freezing of local states, the ellipse $D(x,y)=0$, by analogy with the terminology used in domino tilings, can be called arctic ellipse.

To compute function $g(x,y)$, we have to know regions of values for the roots $w_\pm$ as functions of the coordinates x,y. Clearly, if $(x,y)\in\mathcal{E}$ then w_+ and w_- are complex conjugate; in this case we define w_+ (respectively, w_-) such that w_+ lies in the upper half-plane (lower half-plane). If $(x,y)\in\mathcal{D}_{\rm I}\cup\mathcal{D}_{\rm II}\cup\mathcal{D}_{\rm III}\cup\mathcal{D}_{\rm IV}$ then $w_\pm$ are real, and $w_+>w_-$. Furthermore, to each of these corner regions corresponds an interval of the real axis where both w_+ and w_- take values,

$$(x,y)\in\mathcal{D}_i\leftrightarrow w_\pm\in l_i,\qquad i={\rm I},\dots,{\rm IV}. \tag{4.7}$$

The intervals l_i $(i = \mathrm{I}, \ldots, \mathrm{IV})$ are

$$l_{\mathrm{I}} := (-\infty, -1), \quad l_{\mathrm{II}} := (-1, -\alpha), \quad l_{\mathrm{III}} := (-\alpha, 0), \quad l_{\mathrm{IV}} := (0, \infty). \tag{4.8}$$

Obviously, the case of coincidence of the roots, $w_+ = w_-$, corresponds to the points of the ellipse $D(x, y) = 0$ itself, and, besides, the values $w_\pm = -1, -\alpha, 0, \infty$ correspond to the touching points of the ellipse with the bottom, right, top, and left sides of the square, respectively.

4.2. The function $g(x, y)$. The specifics of integral (4.2) consists in the presence of the factor $(z_1 - z_2)^{-1}$ in the integrand. In large N analysis of the integral there may arise contributions due to deformation of the initial contours of integration over the variables z_1 and z_2 to the corresponding saddle-point ones. Indeed, if one regards (4.2) as a single integral, e.g., over the variable z_1, then the integrand is given by some function having a branch cut in the complex plane of the integration variable, which coincides with the integration contour of the variable z_2. Deformation of the initial contours to the saddle-point ones therefore implies that in this single integral one has to deform both the contour and the branch cut of the integrand. As a result, in applying the saddle-point method to this integral one has to take into account contributions which are due to intersection of the integration contour with the branch cut of the integrand.

Let us denote by Γ_1 and Γ_2 the saddle-point contours of the integration variables z_1 and z_2, respectively. Restrictions on these contours and their particular choice are given below. The integral in (4.2), after deformation of the contours can be written as

$$G(m, n) = \frac{1}{(2\pi\mathrm{i})^2} \oint_{\Gamma_1} \mathrm{d}z_1 \oint_{\Gamma_2} \frac{\exp\{NS(z_1, z_2)\}}{(z_1 + 1)z_2(z_1 - z_2)} \, \mathrm{d}z_2 + \frac{1}{2\pi\mathrm{i}} \int_{\Gamma_{12}} \frac{1}{(z + 1)z} \, \mathrm{d}z, \tag{4.9}$$

where we have used the fact that $S(z, z) = 0$. Here the contour Γ_{12} is a variety of all intersection points of the integration contours occurred in deforming of the initial contours to the contours Γ_1 and Γ_2. In the aforementioned single integral interpretation, the contour Γ_{12} can be defined as a portion of the contour Γ_2 which have been intersected by the contour of the variable z_1 in the course of its deformation to the contour Γ_1.

The second term in (4.9) determines the function $g(x, y)$. Indeed, it can be shown that the first term in (4.9), as usual for the saddle-point method,

vanishes as $N \to \infty$, while the second term is totally independent of N. In general, it can be nonzero. An explicit expression for the function $g(x, y)$ is therefore determined by the contour Γ_{12}. This contour, in turn, is uniquely determined by the saddle-point contours Γ_1 and Γ_2.

Remind that, according to the general concepts of the saddle-point method (see, e.g., [41]), the saddle-point contours must satisfy the so-called minimax principle. Namely, the contours Γ_1 and Γ_2 in our case must be chosen as those providing a minimum on the variety of all possible contours equivalent to the initial ones, for the maximum of the function $\operatorname{Re} S(z_1, z_2)$ on the set of points of the contours where z_1 and z_2 take their values. In practice this means that Γ_1 and Γ_2 must be chosen such that they pass through the saddle-point (or several saddle-points) in which the quantity $\operatorname{Re} S(z_1, z_2)$ approach its minimum in the set of all saddle-points.

Let us consider first the case $(x, y) \in \mathcal{E}$. In this case $\operatorname{Re} S(z_1, z_2) = 0$ for $(z_1, z_2) = (w_i, w_k)$ $(i, k = \pm)$, that is, the minimum of the function $\operatorname{Re} S(z_1, z_2)$ on the set of all saddle-points is approached on all these points. Hence both contours Γ_1 and Γ_2 must pass through the points w_+ and w_-, and moreover, since $\partial^2_{z_1} S(z_1, z_2)|_{z_1=z_2} = -\partial^2_{z_2} S(z_1, z_2)|_{z_1=z_2}$, they must be perpendicular to each other at these points. Obviously, deformation of the initial contours to such contours generate a nontrivial contour Γ_{12}, which can be chosen as a portion of the contour Γ_2 with the endpoints w_- and w_+, see Fig. 5.

Taking into account that, according to the standard saddle-point results, the first term in (4.9) is estimated as $O(1/N)$ (where the integral in vicinity of the points $z_1 = z_2 = w_\pm$ has to be understood in the sense of its principal value), we thus obtain that the function $g(x, y)$ is given by the second term,

$$g(x,y) = \frac{1}{2\pi\mathrm{i}} \int\limits_{\Gamma_{12}} \frac{1}{(z+1)z}\,\mathrm{d}z = \frac{1}{2\pi\mathrm{i}} \big[\ln z - \ln(z+1)\big]\Big|_{w_-}^{w_+} = \frac{\varphi(x,y)}{\pi}. \quad (4.10)$$

Here $\varphi(x, y)$ is the angle between the lines connecting the point w_+ (or the point w_-) with the endpoints of the interval $[-1, 0]$, see Fig. 5. An elementary calculation yields

$$g(x,y) = \frac{1}{\pi} \operatorname{arccot}\left(-\frac{\alpha(x+y-1) + (1-\alpha)(x-y)}{\sqrt{-D(x,y)}}\right), \qquad (x,y) \in \mathcal{E}, \quad (4.11)$$

where the arc-cotangent function takes values in the interval $[0, \pi]$.

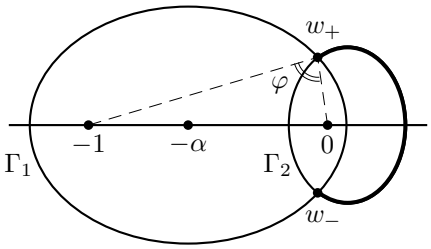

Figure 5. The saddle-point contours Γ_1 and Γ_2 for the case of $(x, y) \in \mathcal{E}$. The contour Γ_{12}, the portion of the contour Γ_2, is shown in bold. Also shown the angle $\varphi = \varphi(x, y)$ which determines the function $g(x, y)$, see (4.10).

Let us now consider the case of $(x, y) \in \mathcal{D}_i$ $(i = \mathrm{I}, \ldots, \mathrm{IV})$. Recall that in this case $w_\pm$ are real, $w_- < w_+$, and they take values described by (4.7) and (4.8). The minimum of the function $\operatorname{Re} S(z_1, z_2)$ on the set of saddle-points is approached at single point: $(z_1, z_2) = (w_-, w_+)$ for $\mathcal{D}_\mathrm{I}$ and $\mathcal{D}_\mathrm{III}$, and $(z_1, z_2) = (w_+, w_-)$ for $\mathcal{D}_\mathrm{II}$ and $\mathcal{D}_\mathrm{IV}$. At the points $z_1 = w_\mp$ and $z_2 = w_\pm$ the contours Γ_1 and Γ_2 must be chosen perpendicular to the real axis.

In deforming the initial contours to the saddle-point ones in the present case we have to satisfy certain restrictions on these contours. Namely, both Γ_1 and Γ_2 have to be simple closed contours, the contour Γ_1 must surround the point $z_1 = -\alpha$, while the contour Γ_2 must surround the point $z_2 = 0$, but not the point $z_2 = -1$. Equivalently, the contour Γ_2 may be chosen surrounding the point $z_2 = -1$, but not the point $z_2 = 0$, in which case it has to be clockwise oriented. A simple analysis shows that all the conditions imposed on the saddle-point contours can be satisfied and the contour Γ_{12} is uniquely defined. Namely, for the regions $\mathcal{D}_\mathrm{I}$ and $\mathcal{D}_\mathrm{IV}$ this contour is null, while for the regions $\mathcal{D}_\mathrm{II}$ and $\mathcal{D}_\mathrm{III}$ the contour Γ_{12} coincides with Γ_2, see Fig. 6.

Since the first term in (4.9) can be estimated as $O\bigl(\mathrm{e}^{-N|S(w_+,w_-)|}\bigr)$, the function $g(x, y)$ is again given by the second term, as expected, for which $\Gamma_{12} = \emptyset$, if $(x, y) \in \mathcal{D}_\mathrm{I} \cup \mathcal{D}_\mathrm{IV}$, and $\Gamma_{12} = C_0$, if $(x, y) \in \mathcal{D}_\mathrm{II} \cup \mathcal{D}_\mathrm{III}$. Hence

$$g(x, y) = \begin{cases} 0 & \text{if } (x, y) \in \mathcal{D}_\mathrm{I} \cup \mathcal{D}_\mathrm{IV}, \\ 1 & \text{if } (x, y) \in \mathcal{D}_\mathrm{II} \cup \mathcal{D}_\mathrm{III}. \end{cases} \tag{4.12}$$

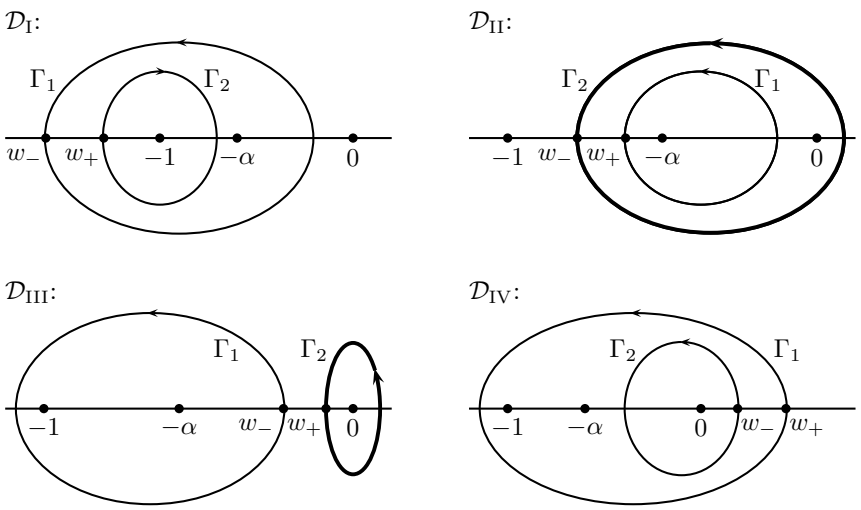

Figure 6. The saddle-point contours Γ_1 and Γ_2 for $(x,y) \in D_i$ $(i = \mathrm{I},\dots,\mathrm{IV})$. The contour Γ_{12} is shown in bold: it coincides with Γ_2 for $(x,y) \in \mathcal{D}_{\rm II} \cup \mathcal{D}_{\rm III}$ and null for $(x,y) \in \mathcal{D}_{\rm I} \cup \mathcal{D}_{\rm IV}$.

It is interesting to note, that the result obtained above for the function $g(x,y)$ in terms of the angle φ, see (4.10), turns out also valid for the points $(x,y) \in \mathcal{D}_i$ $(i = \mathrm{I},\dots,\mathrm{IV})$; indeed, $\varphi = \pi$, if $w_\pm$ take values at the interval $(-1,0)$, and $\varphi = 0$ otherwise, i.e., if $w_\pm \in (-\infty,-1) \cup (0,\infty)$, see (4.7) and (4.8).

4.3. Vicinity of the ellipse. Having established the thermodynamic limit expression for the one-point function, let us now consider fluctuations of local states near the arctic ellipse over the limiting value for a finite system. Our aim is to derive an approximate expression for the one-point function $G(m,n)$ valid for large, but finite N, and coordinates (m,n) taking values is some small (in comparison with N) vicinity of the phase separation curve, i.e., near the arctic ellipse. We show that for the size of this vicinity of order $N^{1/3}$ (that corresponds $N^{-2/3}$ in terms of coordinates (x,y)) our one-point function coincides with that of the Gaussian unitary ensemble at the edge of the spectrum.

We shall consider here the case of vicinity of the portion of the ellipse separating regions $\mathcal{D}_{\rm IV}$ and $\mathcal{E}$ (see Fig. 4); calculations for other portions are essentially similar.

We first introduce a parametrization for coordinates which is suitable in vicinity of the ellipse. The points of the indicated portion of the ellipse we denote as (x_0, y_0), i.e., we have $D(x_0, y_0) = 0$, where the function $D(x, y)$ is given by (4.6). We have relations

$$x_0 + y_0 - 1 = \sqrt{1-\alpha}\cos\phi, \qquad y_0 - x_0 = \sqrt{\alpha}\sin\phi,$$

where ϕ is some parameter. It will be convenient here to parameterize the parameter α as

$$\alpha = \sin^2\lambda, \qquad \lambda \in [0, \pi/2]. \tag{4.13}$$

Then for all points (x_0, y_0) we have the following parameterization

$$x_0 = \cos^2\left(\frac{\phi+\lambda}{2}\right), \qquad y_0 = \cos^2\left(\frac{\phi-\lambda}{2}\right), \qquad \phi \in [\lambda, \pi-\lambda].$$

The points (x, y), lying on a normal to the ellipse which intersects it at the point (x_0, y_0), can be parameterized as

$$x = x_0 + \frac{\sin(\phi-\lambda)}{\sqrt{1-\cos 2\phi\cos 2\lambda}}t, \qquad y = y_0 - \frac{\sin(\phi+\lambda)}{\sqrt{1-\cos 2\phi\cos 2\lambda}}t, \tag{4.14}$$

where t is introduced such that $|t|$ has the meaning of the distance between the points (x, y) and (x_0, y_0); note that $t > 0$, if $(x, y) \in \mathcal{E}$, and $t < 0$, if $(x, y) \in \mathcal{D}_{\rm IV}$. Below we will be interested in a double scaling limit, $N \to \infty$ and $t \to 0$, in which N and t approach their limits in some consistent way.

Before addressing this problem, we first consider behavior of the function $g(x, y)$ in vicinity of the arctic ellipse, that corresponds to $N \to \infty$ with t small but fixed. Substituting (4.13) and (4.14) into (4.6), for the discriminant $D(x, y)$, as t is small, we obtain

$$D(x, y) = -\sin 2\lambda\big(1 - \cos 2\phi\cos 2\lambda\big)^{1/2}t + O(t^2). \tag{4.15}$$

Consider the case $t > 0$, i.e., $(x, y) \in \mathcal{E}$. Taking into account that $\alpha(x + y - 1) + (1-\alpha)(x - y) = -\sin\lambda\cos\lambda\sin(\phi-\lambda) + O(t)$, for the function $g(x, y)$, from (4.11), we obtain

$$g(x, y) \sim \frac{2\big(1 - \cos 2\phi\cos 2\lambda\big)^{1/4}}{\pi\sqrt{\sin 2\lambda}\sin(\phi-\lambda)}\sqrt{t} \qquad (t > 0). \tag{4.16}$$

In the case $t < 0$, i.e., $(x, y) \in \mathcal{D}_{\rm IV}$, for this function we have $g(x, y) = 0$ for all values of t, see (4.12). Hence the function $g(x, y)$ has a square root

singularity at the arctic ellipse. It is to be mentioned, however, that this is valid for all points of the considered portion of the ellipse, except the point $\phi = \lambda$ (at which the coefficient in (4.16) has a pole), corresponding to the point $(x, y) = (1 - \alpha, 1)$, at which the function $g(x, y)$ has a different kind of singularity (see discussion in Sect. 4.4).

Now we turn to our main problem here, namely, we compute an approximate expression for the one-point function $G(m, n)$ valid at large but finite N, and coordinates $m = Nx$ and $n = Ny$ with x and y given by (4.14), at small t. One can expect that the square root singularity of the one-point function appears somehow smoothed, as t scaled accordingly with N, as both approach their limiting values, $N \to \infty$ and $t \to 0$. The results below are valid in vicinity of all points our portion of the ellipse, except vicinity of the points which are too close to the boundary, that is, far enough from the points $\phi = \lambda$ and $\phi = \pi - \lambda$.

For small t, application of the saddle-point method to the integral (4.2) is based on an observation that $w_\pm \to w_0$ as $t \to 0$, where w_0 is real and $w_0 \in (0, \infty)$, since we are considering vicinity of the arctic ellipse separating the regions $\mathcal{E}$ and $\mathcal{D}_{\mathrm{IV}}$, see (4.7). From (4.5) we have $w_0 = (y_0 - x_0 - \alpha)/2x_0$, or, in terms of the parameters ϕ and λ,

$$w_0 = \frac{\sin\lambda \sin\left(\frac{\phi-\lambda}{2}\right)}{\cos\left(\frac{\phi+\lambda}{2}\right)}. \tag{4.17}$$

More precisely, for $w_\pm$, due to (4.15), we have

$$w_\pm = w_0 \pm \mathrm{i}\frac{(\sin 2\lambda)^{1/2}(1 - \cos 2\lambda \cos 2\phi)^{1/4}}{2\cos^2\left(\frac{\phi+\lambda}{2}\right)}\sqrt{t} + O(t).$$

The second equality in (4.4) yields

$$F'(w_0) = \frac{2(1 - \cos 2\lambda \cos 2\phi)^{1/2} \cot\frac{1}{2}(\phi + \lambda)}{\sin 2\lambda \sin(\phi - \lambda)} t + O(t^2).$$

Evaluating higher derivatives of the function $F(z)$ at the point $z = w_0$, we obtain

$$F''(w_0) = O(t^2), \qquad F'''(w_0) = \frac{16\cos^4\left(\frac{\phi+\lambda}{2}\right)}{\sin^2 2\lambda \sin(\phi - \lambda)} \cot\left(\frac{\phi + \lambda}{2}\right) + O(t),$$

and also we find that all quantities $F^{(r)}(w_0)$ for $r \geqslant 4$ are nonsingular as $t \to 0$. Thus, since $F''(w_0) = 0$, $F'''(w_0) \neq 0$ as $t \to 0$, and since the function $F(z)$ is expandable in a Taylor series at $z = w_0$, we are in

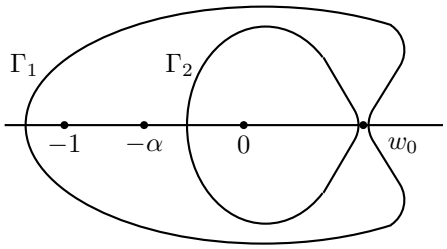

Figure 7. The saddle-point contours Γ_1 and Γ_2 for the case of vicinity (for t of order $N^{-2/3}$) of the portion of the arctic ellipse separating regions $\mathcal{E}$ and $\mathcal{D}_{\mathrm{IV}}$.

a standard situation of the single variable saddle-point method, for each variable z_1 and z_2, where two saddle-points join into a single point and the leading term of asymptotics, for $t = O(N^{-2/3})$ as $N \to \infty$, is given by Airy function (see, e.g., [41,42]). In our case, the saddle-point contours Γ_1 and Γ_2 are shown in Fig. 7. Note that Γ_1 and Γ_2 can be chosen such that they do not intersect each other in the vicinity of the point w_0, as shown in the picture.

An approximate representation for the one-point function is based on the use of the expression

$$NF(z) \approx NF(w_0) + \sigma s + s^3/3,$$

where s and σ are finite, and neglected terms tend to zero as $N \to \infty$. Here, s is a new integration variable

$$s = \Omega N^{1/3}(z - w_0), \quad \Omega = \left[\frac{8\cos^4\left(\frac{\phi+\lambda}{2}\right)}{\sin^2 2\lambda \sin(\phi - \lambda)} \cot\left(\frac{\phi+\lambda}{2}\right) \right]^{1/3}, \quad (4.18)$$

and σ is a new parameter related to the distance to the arctic ellipse,

$$\sigma = \frac{2^{2/3}\big(1 - \cos 2\phi \cos 2\lambda\big)^{1/2}}{(\sin 2\lambda)^{1/3}\big[\sin(\phi - \lambda)\sin(\phi + \lambda)\big]^{2/3}} N^{2/3}\, t.$$

Making change of the integration variables $z_1 \mapsto s_1$ and $z_2 \mapsto s_2$ in virtue (4.18), we obtain the following expression for the one-point function for

large N and finite σ:

$$G(m,n) \approx \frac{1}{(2\pi \mathrm{i})^2 (w_0+1) w_0 \, \Omega N^{1/3}} \times \int\limits_{\mathcal{L}_-} \mathrm{d}s_2 \int\limits_{\mathcal{L}_+} \frac{\exp\left\{\sigma s_1 + \frac{1}{3}s_1^3 - \sigma s_2 - \frac{1}{3}s_2^3\right\}}{s_1 - s_2} \, \mathrm{d}s_1.$$

Here, the contour $\mathcal{L}_+$ comes from the infinity with $\arg s_1 = \pi/3$ at the vicinity of the origin with $\operatorname{Re} s_1 > 0$ and goes to infinity with $\arg s_1 = -\pi/3$, while the contour $\mathcal{L}_-$ comes from infinity with $\arg s_2 = -2\pi/3$ at the vicinity of the origin with $\operatorname{Re} s_2 < 0$ and goes to infinity with $\arg s_2 = 2\pi/3$. Since for all points of these contours $\operatorname{Re}(s_1 - s_2) > 0$, we can use the formula

$$\frac{\exp\{\sigma(s_1 - s_2)\}}{s_1 - s_2} = \int\limits_{-\infty}^{\sigma} \exp\{v(s_1 - s_2)\} \, \mathrm{d}v.$$

Changing the order of integrals (due to their absolute convergence), and using explicit expressions for w_0 and Ω, see (4.17) and (4.18), we finally obtain

$$G(m,n) \approx \left[\frac{2\sin(\phi+\lambda)}{N \sin 2\lambda \sin^2(\phi-\lambda)}\right]^{1/3} \int\limits_{-\infty}^{\sigma} [\mathrm{Ai}(-v)]^2 \mathrm{d}v, \tag{4.19}$$

where $\mathrm{Ai}(v)$ is Airy function (see, e.g., [42]). The integral in (4.19) can be evaluated (see, e.g., [43]):

$$\int\limits_{-\infty}^{\sigma} [\mathrm{Ai}(-v)]^2 \mathrm{d}v = [\mathrm{Ai}'(-\sigma)]^2 + \sigma [\mathrm{Ai}(-\sigma)]^2. \tag{4.20}$$

Expression (4.19) describes an expected smoothening of the one-point function. Indeed, since $[\mathrm{Ai}'(-\sigma)]^2 + \sigma[\mathrm{Ai}(-\sigma)]^2 \propto \sigma^{1/2}/\pi$ as $\sigma \to \infty$, the expression in (4.16) follows from (4.19) as $N \to \infty$, at t small and kept fixed.

The function standing in the right hand side of (4.20) is a particular case, at $u = v = -\sigma$, of the Airy kernel

$$K(u,v) = \frac{\mathrm{Ai}(u)\mathrm{Ai}'(v) - \mathrm{Ai}'(u)\mathrm{Ai}(v)}{u - v}.$$

This kernel is known as describing correlation functions of the Gaussian unitary ensemble at the edge of the spectrum [44] (for a review, see, e.g., [45]). In particular, $K(-\sigma,-\sigma)$ is the one-point function, or the density of random matrix eigenvalues. Thus, the expression in (4.19) tell us, that in the vicinity of the arctic ellipse the one-point function of the six-vertex model at its free-fermion point coincides, modulo a factor, with the one-point function of the Gaussian unitary matrix ensemble at the edge of the spectrum. This result agrees with previous results on domino tilings, see [31, 46].

4.4. Vicinity of the touching points. Let us now derive an approximate expression for the one-point function $G(m,n)$ in vicinities of the points where the disordered region touches the boundary. As in the case of the vicinity of the arctic ellipse considered above, a characteristic distance in the number of lattice site here will be of order $N^{1/3}$, or, in terms of the scaled coordinates (x,y), of order $N^{-2/3}$. It is to be mentioned that near the touching point a different scaling is also possible, namely, with $O(1)$ distance away from the boundary and $O(N^{1/2})$ distance along it [47, 48]. We limit ourselves here only to treating the $O(N^{1/3})$ case of scaling near the touching points, since it naturally complements our discussion of the vicinity of the ellipse above.

To keep contact with our previous discussion, we will consider vicinity of those contact points, which belong to the region $\mathcal{D}_{\mathrm{IV}}$, i.e., the points $(x,y)=(1-\alpha,1)$ and $(x,y)=(0,\alpha)$. The results for other two points, $(x,y)=(\alpha,0)$ and $(x,y)=(1,1-\alpha)$ can be derived essentially similarly, and they can also be recovered from those for the first two, respectively, in virtue of the relation (2.3).

We start with considering the function $g(x,y)$. From the discussion above it is clear that it has a singular behavior at the touching points. Let us consider first the case of the point $(x,y)=(1-\alpha,1)$. We parameterize the coordinates in a small vicinity of this point as $(x,y)=(1-\alpha-\rho\cos\mu, 1-\rho\sin\mu)$ where $\mu\in[0,\pi]$ and ρ is positive. As ρ is small,

$$D(x,y) = -4\alpha(1-\alpha)\rho\sin\mu + O(\rho^2). \tag{4.21}$$

Substituting this into (4.11), we get

$$g(x,y) = \frac{1}{\pi}\operatorname{arccot}\left(\frac{\cos\mu-(1-2\alpha)\sin\mu}{2\sqrt{\alpha(1-\alpha)\sin\mu}}\sqrt{\rho}+O(\rho^{3/2})\right)$$

and, since $\operatorname{arccot}\varepsilon = \pi/2 - \varepsilon + O(\varepsilon^3)$, as $\varepsilon \to 0$, expanding in power series in ρ yields

$$g(x,y) = \begin{cases} 0 & \mu = 0, \\ \dfrac{1}{2} + \dfrac{(1-2\alpha)\sin\mu - \cos\mu}{2\pi\sqrt{(1-\alpha)\alpha\sin\mu}}\,\sqrt{\rho} + O(\rho^{3/2}) & 0 < \mu < \pi, \\ 1 & \mu = \pi. \end{cases} \quad (4.22)$$

Thus, at this point the function $g(x,y)$ has a step-wise behavior; this explains the singularity of the coefficient in (4.16) at $\phi = \lambda$.

In a small vicinity of the point $(x,y) = (0,\alpha)$, we parameterize the coordinates as $(x,y) = (\rho\sin\mu, \alpha - \rho\cos\mu)$; for ρ small the discriminant is again given by (4.21), but the function $g(x,y)$ now reads

$$\begin{aligned} g(x,y) &= \frac{1}{\pi}\operatorname{arccot}\left(\sqrt{\frac{\alpha(1-\alpha)}{\rho\sin\mu}} + O(\rho^{1/2})\right) \\ &= \frac{1}{\pi}\sqrt{\frac{\rho\sin\mu}{\alpha(1-\alpha)}} + O(\rho^{3/2}), \end{aligned} \quad (4.23)$$

and, in fact, $g(x,y) = 0$ exactly for $\mu = 0, \pi$, as we know from (4.12). Thus we recover (4.16) at $\phi = \pi - \lambda$, as expected, with $t := \rho\sin\mu$.

Consider now the problem of deriving an approximate expression for $G(n,m)$ for the point (m,n) close to one of the touching points as N is large but finite. This task is very similar to that in the case of the vicinity of the arctic ellipse solved above, but it differs both in how the saddle-point approximation actually works and in a specific choice of variables suitable for calculations in the double scaling limit in question.

In comparison with the case of vicinity of the arctic ellipse, in the case of vicinity of a touching point there exists a large integer which can be used instead of N. Namely, such integer, which we denote M, can be interpreted as the number of lattice sites from the point (m,n) to the boundary; below we give a precise definition of M. At the same time, there exists a small positive parameter, denote it ε, such that $M\varepsilon$ is finite, which somehow is related to the distance of the point (m,n) from the touching point in terms of the scaled coordinates; we shall below see that ε is in fact proportional to the square root of this distance.

To be specific, we consider here in detail calculations for the case of the touching point $(x,y) = (1-\alpha, 1)$. As above, we parameterize coordinates in a small vicinity of this point as $(x,y) = (1 - \alpha - \rho\cos\mu, 1 - \rho\sin\mu)$

and assume that $\mu \neq 0, \pi$ since only the case of $(x, y) \in \mathcal{E}$ is actually interesting; more precisely, in our calculations below it is assumed that $|\cot \mu|$ is bounded. When ρ is small, the numbers $w_\pm$ (solutions of the equation $F'(z) = 0$, see (4.5)) are complex conjugate (since $(x, y) \in \mathcal{E}$) and moreover close to each other and to the point $z = 0$ (see the discussion after (4.8)). Namely, using (4.21), we have

$$w_\pm = \pm \mathrm{i}\varepsilon + O(\rho), \tag{4.24}$$

where

$$\varepsilon = \sqrt{\frac{\alpha}{1-\alpha}\, \rho \sin \mu}. \tag{4.25}$$

The meaning of (4.24) becomes clear if we come back to (3.30) and recall that the function $\Phi(z)$ contains the factor z^{N-n}. Then (4.24) implies that we deal with the situation when the saddle-points are close to each other and to a zero, for z_1, or to a pole, for z_2, of a large order, the value of which is not fixed, but grows together with the large parameter of the integral (i.e., together with N). Denoting this order by $M := N - n$, we thus have the following relation between M, N, and ε:

$$M = N\rho \sin \mu = \frac{1-\alpha}{\alpha} N\varepsilon^2. \tag{4.26}$$

The saddle-point approximation to the double integral in this case is constructed by zooming in a small vicinity of the point $z = 0$ in the both integrals. Indeed, the function $F(z)$ possesses a singular Taylor expansion at points $z = w_\pm$ since its derivatives at these points are estimated, as $\varepsilon \to 0$, by $F^{(k)}(w_\pm) \sim \varepsilon^{-k+2}$ $(k = 2, 3, \dots)$. These estimates govern the scaling into the vicinity of the point $z = 0$ to be of the form $z = \varepsilon s$, where s must be kept finite. For the function $F(z)$, we have then the following representation

$$NF(\varepsilon s) = f_0 + M\left(\frac{s^2}{2} + \log s\right) - M\varepsilon\left(\frac{1+\alpha}{3\alpha}\, s^3 + \frac{\alpha + \cot \mu}{\alpha}\, s\right) + \sum_{k=2}^{\infty} M\varepsilon^k f_{k+2}(s), \tag{4.27}$$

where $f_k(s)$ are some polynomials of degree k in s, and f_0 denotes all s-independent terms (so, e.g., $f_k(0) = 0$). As M is large, the saddle-point approximation admits $M\varepsilon$, being the coefficient of the third term in (4.27),

to be a finite quantity. Due to (4.26), in such case $N\varepsilon^3$ is finite too, that fixes the whole scale at which our approximation is applicable:

$$M \sim N^{1/3}, \qquad \varepsilon \sim N^{-1/3}. \tag{4.28}$$

The saddle-point contours Γ_1 and Γ_2 are to be chosen lying in small vicinities of the points $z_1 = 0$ and $z_2 = 0$ such that Γ_1, as passing through the points $-\mathrm{i}\varepsilon$ and $+\mathrm{i}\varepsilon$, crosses the real axis perpendicularly, while Γ_2 goes parallel to the real axis below (from left to right), passing through the point $-\mathrm{i}\varepsilon$, and above (from right to left), passing through the point $+\mathrm{i}\varepsilon$. After zooming into the vicinity of the origin in the both integrals, and ignoring possible exponentially small corrections, we obtain the following approximate representation for the one-point function:

$$G(m,n) \approx \frac{1}{(2\pi\mathrm{i})^2} \left[\int_{-\mathrm{i}-\infty}^{-\mathrm{i}+\infty} + \int_{\mathrm{i}+\infty}^{\mathrm{i}-\infty} \right] \mathrm{d}s_2 \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\exp\{NS(\varepsilon s_1, \varepsilon s_2)\}}{(1+\varepsilon s_1) s_2 (s_1 - s_2)} \, \mathrm{d}s_1 + \frac{1}{\pi} \operatorname{arccot} \varepsilon. \tag{4.29}$$

Now the standard saddle-point method can be applied to this expression, with $NS(\varepsilon s_1, \varepsilon s_2) = NF(\varepsilon s_1) - NF(\varepsilon s_2)$ where $F(\varepsilon s)$ is given by (4.27), to obtain an $1/M$ expansion; we just point out that at the saddle-points $(s_1, s_2) = (\pm\mathrm{i}, \pm\mathrm{i})$ the singularities have to be understood in the sense of the principal values. As a result, for the the one-point function in the vicinity of the touching point $(x, y) = (1-\alpha, 1)$ we obtain

$$G(m,n) = \frac{1}{2} + \varepsilon \frac{1 - 2\alpha - \cot\mu}{2\pi\alpha} + \frac{(-1)^M}{4\pi M} \sin\left(2M\varepsilon \frac{1 - 2\alpha - 3\cot\mu}{3\alpha} \right) + O(M^{-2}) \tag{4.30}$$

and we remind that M, ε, and $\cot\mu$ are related to m, n, and N by

$$M = N - n, \qquad \varepsilon = \sqrt{\frac{(1-\alpha)M}{\alpha N}}, \qquad \cot\mu = \frac{(1-\alpha)N - m}{M}.$$

In (4.30), the first two terms coincides with the thermodynamic limit expansion (4.22). Third term is of the same magnitude as the second term, see (4.28), and for large but finite values of N this term describes a finite-size correction. Due to (4.26), as $N \to \infty$, at ε small and fixed, it vanishes. The meaning of each term in the approximate expression in (4.30) is also get clarified when the one-point function is plotted near the touching point

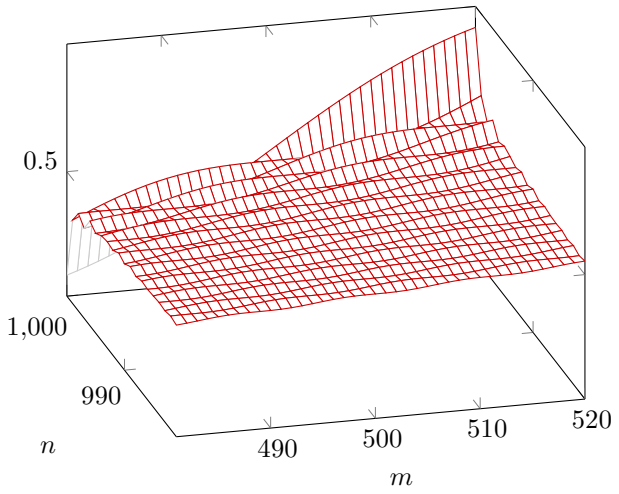

Figure 8. Plot (for $N = 1000$ and $\alpha = 1/2$) of exact formula (3.26) for the one-point function $G(m, n)$ near the touching point at the top boundary. An approximate expression is given by (4.30).

at the top boundary for sufficiently large N, an example for $N = 1000$ and $\alpha = 1/2$ is shown in Fig. 8.

We finish here by outlining the case of the touching point $(x, y) = (0, \alpha)$. As in deriving (4.23), we parameterize the coordinates in a small vicinity of this point as $(x, y) = (\rho \sin \mu, \alpha - \rho \cos \mu)$. The parameter ε is again defined by (4.25), but now $M = m$ and so (4.26) holds. The only difference with the previous calculation is that in this case the integral similar to (4.29) acquires an extra overall factor ε and the leading saddle-point contribution describes a next to the leading order correction. The result reads:

$$G(m,n) = \frac{\varepsilon}{\pi\alpha}\left\{1 - \frac{(-1)^M}{4M}\cos\left(2M\varepsilon\frac{1-2\alpha-3\cot\mu}{3\alpha}\right) + O(M^{-2})\right\}. \quad (4.31)$$

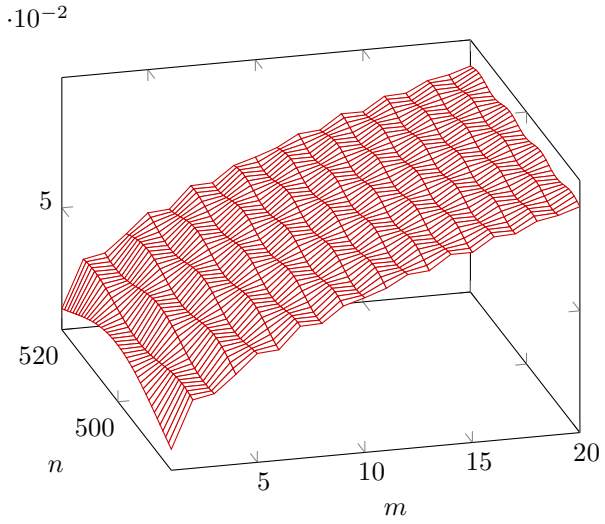

Figure 9. Plot (for $N = 1000$ and $\alpha = 1/2$) of exact expression (3.26) for the one-point function $G(m, n)$ near the touching point at the left boundary. An approximate expression is given by (4.31).

Here, M, ε, and $\cot\mu$ are related to m, n, and N by

$$M = m, \qquad \varepsilon = \sqrt{\frac{(1-\alpha)M}{\alpha N}}, \qquad \cot\mu = \frac{\alpha N - n}{M}.$$

Fig. 9 shows plot of the one-point function (at $N = 1000$ and $\alpha = 1/2$) near the touching point at the left boundary, where approximate expression (4.31) is applicable.

§5. Conclusion

The purpose of this paper is to outline a fermionic approach to the arctic ellipse phenomenon in the six-vertex model. Namely, using the formulation of the model in terms of fermions (Sect. 2) we show that the one-point function at the free-fermion point can be straightforwardly computed (Sect. 3) that, in turn, allows us to derive the arctic ellipse phenomenon (Sect. 4). There are open problems, which can be addressed in relation with each of these steps.

First, in Sect. 2 we have derived a representation for the partition function in terms of Grassmann integral, resembling a 2D lattice fermion theory. One can wonder about which field theory it corresponds. Answering this question may appear relevant in obtaining results beyond the free-fermion point.

Second, in Sect. 3 we have evaluated the integrals essentially considering matrix elements of the fermionic transfer matrix, T, between Grassmann coherent states. Can this procedure be generalized on the case of the generic weights (i.e., generic Δ)?

Last, in Sect. 4 we have studied the one-point function in the thermodynamic limit and derived approximate expression in the $N^{1/3}$ vicinity of the arctic ellipse. In the general situation, for the points of the ellipse far away from the boundary, we have found that the one-point function is described by the Airy-kernel, while near the points where the arctic ellipse touches the boundary we have obtained instead some relatively simple expressions. It would be very interesting to understand these results from the point of view of the random matrices theory. Indeed, it is well-known that in the random matrix models the Airy-kernel describes the double scaling limit near the edge of the eigenvalue density. At the same time, the case of the points where the arctic ellipse touches the boundary corresponds to a very peculiar situation in the random matrix model, namely, where the edge of the spectrum exactly coincides with a hard-wall, and, moreover, the matrix model is considered with a discrete spectrum (since the one-point function, an analog of the eigenvalue density, is constrained). A similar situation arises for the six-vertex considered on a domain with a cut-off corner, such that the size of this corner is macroscopically large and hits the arctic ellipse [49].

Acknowledgements

The authors are indebted for interest to the work and stimulating discussions to A. A. Andrianov, P. A. Belov, N. M. Bogoliubov, H. Boos, A. Capelli, F. Colomo, S. E. Derkachev, F. Göhmann, A. Klümper, A. Knitzel, A. V. Kitaev, V. E. Korepin, I. Lyberg, J. Mussardo, Kh. F. Nirov, A. Okounkov, N. Reshetikhin, G. A. P. Ribeiro, O. F. Syljuasen, J. Viti, P. Wiegmann, and M. B. Zvonarev.

References

1. V. E. Korepin, *Calculations of norms of Bethe wave functions*. Commun. Math. Phys. **86** (1982), 391–418.
2. A. G. Izergin, *Partition function of the six-vertex model in the finite volume*. Sov. Phys. Dokl. **32** (1987), 878–879.
3. A. G. Izergin, D. A. Coker, V. E. Korepin, *Determinant formula for the six-vertex model*. J. Phys. A **25** (1992), 4315–4334.
4. V. E. Korepin, N. M. Bogoliubov, A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions*, Cambridge University Press, Cambridge, 1993.
5. D. M. Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, Cambridge University Press, Cambridge, 1999.
6. K. Eloranta, *Diamond ice*. J. Stat. Phys. **96** (1999), 1091–1109.
7. H. Cohn, N. Elkies, J. Propp, *Local statistics for random domino tilings of the Aztec diamond*. Duke Math. J. **85** (1996), 117–166.
8. W. Jockush, J. Propp, P. Shor, *Random domino tilings and the arctic circle theorem*, eprint=math.CO/9801068.
9. H. Cohn, R. Kenyon, J. Propp, *A variational priciple for domino tilings*. J. Amer. Math. Soc. **14** (2001), 297–346.
10. N. Destainville, *Entropy and boundary conditions in random rhombus tilings*. J. Phys. A **31** (1998), 6123.
11. H. Cohn, M. Larsen, J. Propp, *The shape of a typical boxed plane partition*. New York J. Math. **4** (1998), 137–165.
12. R. Kenyon, A. Okounkov, *Limit shapes and the complex Burgers equation*. Acta Math. **199** (2007), 263–302.
13. R. Kenyon, A. Okounkov, S. Sheffield, *Dimers and amoebae*. Ann. Math. **163** (2006), 1019–1056.
14. N. Elkies, G. Kuperberg, M. Larsen, J. Propp, *Alternating-sign matrices and domino tilings*. J. Algwebraic Combin. (1992), No. 1, 111–132; 219–234.
15. V. E. Korepin, P. Zinn-Justin, *Thermodynamic limit of the six-vertex model with domain wall boundary conditions*. J. Phys. A **33** (2000), 7053–7066.
16. P. Zinn-Justin, *Six-vertex model with domain wall boundary conditions and one-matrix model*. Phys. Rev. E. **62** (2000), 3411–3418.
17. P. Zinn-Justin, *The influence of boundary conditions in the six-vertex model*,
18. K. Palamarchuk, N. Reshetikhin, *The six-vertex model with fixed boundary conditions*. PoS (Solvay), (2008), 012.
19. O. F. Syljuasen, M. B. Zvonarev, *Monte-Carlo simulations of vertex models*. Phys. Rev. E **70** (2004), 016118.
20. D. Allison, N. Reshetikhin, *Numerical study of the* 6*-vertex model with domain wall boundary conditions*. Ann. Inst. Fourier (Grenoble) **55** (2005), 1847–1869.
21. N. M. Bogoliubov, A. V. Kitaev, M. B. Zvonarev, *Boundary polarization in the six-vertex model*. Phys. Rev. E **65** (2002), 026126.
22. N. M. Bogoliubov, A. G. Pronko, M. B. Zvonarev, *Boundary correlation functions of the six-vertex model*. J. Phys. A **35** (2002), 5525–5541.
23. O. Foda, I. Preston, *On the correlation functions of the domain wall six vertex model*. J. Stat. Mech. Theory Exp. **2004** (2004), P11001.

24. F. Colomo, A. G. Pronko, *On two-point boundary correlations in the six-vertex model with domain wall boundary conditions*. J. Stat. Mech. Theory Exp. **2005** (2005), P05010.
25. F. Colomo, A. G. Pronko, *Emptiness formation probability in the domain-wall six-vertex model*. Nucl. Phys. B **798** (2008), 340–362.
26. F. Colomo, A. G. Pronko, *The limit shape of large alternating-sign matrices*. SIAM J. Discrete Math. **24** (2010), 1558–1571.
27. F. Colomo, A. G. Pronko, *The arctic curve of the domain-wall six-vertex model*. J. Stat. Phys. **138** (2010), 662–700.
28. F. Colomo, A. G. Pronko, P. Zinn-Justin, *The arctic curve of the domain-wall six-vertex model in its anti-ferroelectric regime*. J. Stat. Mech. Theory Exp. **2010** (2010), L03002.
29. K. Motegi, *Boundary correlation functions of the six and nineteen vertex models with domain wall boundary conditions*. Physica A **390** (2011), 3337–3347.
30. F. Colomo, A. G. Pronko, *An approach for calculating correlation functions in the six-vertex model with domain wall boundary conditions*. Theor. Math. Phys. **171**, No. 2 (2012), 254–270.
31. P. L. Ferrari, H. Spohn, *Domino tilings and the six-vertex model at its free fermion point*. J. Phys. A **39** (2006), 10297–10306.
32. C. Destri, H. J. de Vega, *Light-cone lattice approach to fermionic theories in 2D: The massive Thirring model*. Nucl. Phys. B **290** (1987), 363–391.
33. Y. Umeno, M. Shiroishi, M. Wadati, *Fermionic R-operator for the fermion chain model*. — J. Phys. Soc. Japan **67**, No. 6 (1998), 1930–1935.
34. F. Göhmann, V. E. Korepin, *Solution of the quantum inverse problem*. — J. Phys. A **33** (2000), 1199–1220.
35. F. A. Berezin, *The method of second quantization. Pure and applied physics. A series of monographs and textbooks*, Academic Press, New York, 1966.
36. F. A. Berezin, *Introduction to Superanalysis, Mathematical Physics and Applied Mathematics*, D. Reidel Publishing Co., Dordrecht, 1987.
37. K. E. Cahill, R. J. Glauber, *Density operators for fermions*. — Phys. Rev. A **59** (1999), 1539–1555.
38. M. Combescure, D. Robert, *Fermionic coherent states*. — J. Phys. A **45** (2012), 244005.
39. V. S. Kapitonov, A. G. Pronko, *The five-vertex model and boxed plane partitions*. — J. Math. Sci. (N. Y.) **158** (2009), 858–867.
40. V. S. Kapitonov, A. G. Pronko, *Weighted enumerations of boxed plane partitions and the inhomogeneous five-vertex model*. — J. Math. Sci. (N. Y.) **192** (2013), 70–80.
41. M. V. Fedoryuk, *The Saddle-Point Method* [in Russian], Nauka, Moskow, 1977.
42. F. W. J. Olver, *Asymptotics and Special Functions*, A. K. Peters Ltd., Wellesley, MA, 1974.
43. A. P. Prudnikov, Y. A. Brychkov, O. I. Marichev, *Integrals and series, Vol. 3: More special functions*, Gordon and Breach Science Publishers, New York, NY, 1990.
44. C. A. Tracy, H. Widom, *Level spacing distributions and the Airy kernel*. — Commun. Math. Phys. **159** (1994), 151–174.
45. M. L. Mehta, *Random Matrices*, 3rd. ed., Elsevier, Amsterdam, 2004.

46. K. Johansson, *The arctic circle boundary and the Airy process.* — Ann. Probab. **33** (2005), 1–30.
47. A. Yu. Okounkov, N. Yu. Reshetikhin, *The birth of a random matrix.* — Mosc. Math. J. **6**, No. 3 (2006), 553–566.
48. K. Johansson, E. Nordenstam, *Eigenvalues of GUE minors.* — Electron. J. Probab. **11** (2006), 1342–1371.
49. F. Colomo, A. G. Pronko, *Third-order phase transition in random tilings.* — Phys. Rev. E **88** (2013), 042125.

Department of Mathematics
St. Petersburg State Technological Institute
(Technical University),
St. Petersburg, Russia

Steklov Mathematical Institute
St. Petersburg, Russia
E-mail: agp@pdmi.ras.ru

Поступило 16 ноября 2020 г.