RECEIVED: June 4, 2019
 REVISED: August 30, 2019
 ACCEPTED: October 17, 2019
 PUBLISHED: October 25, 2019

Continuum limit Tonks-Girardeau matrix elements. Part I. The ground state and the uniform density state

Jarah Evslin, Hosam Mohammed, Hui Liu and Yao Zhou

*Institute of Modern Physics,
 NanChangLu 509, Lanzhou 730000, China
 University of the Chinese Academy of Sciences,
 YuQuanLu 19A, Beijing 100049, China*

E-mail: jarah@impcas.ac.cn, hosam@impcas.ac.cn, liuhui@impcas.ac.cn,
yaozhou@impcas.ac.cn

ABSTRACT: The Tonks-Girardeau model is a quantum mechanical model of N impenetrable bosons in 1+1 dimensions. A Vandermonde determinant provides the exact N -particle wave function of the ground state, or equivalently the matrix elements with respect to position eigenstates. We consider the large N limit of these matrix elements. We present a binning prescription which calculates the leading terms of the matrix elements in a time which is independent of N , and so is suitable for this limit. In this sense, it allows one to solve for the ground state of a strongly coupled continuum quantum field theory in the field eigenstate basis. As examples, we calculate the matrix elements with respect to states with uniform density and also states consisting of two regions with distinct densities.

KEYWORDS: Bethe Ansatz, Integrable Field Theories, Field Theories in Lower Dimensions, Lattice Integrable Models

ARXIV EPRINT: [1906.00683](https://arxiv.org/abs/1906.00683)

Contents

1	Introduction	1
2	The Tonks-Girardeau model	3
2.1	The model and its ground state	3
2.2	A Bethe Ansatz formulation	5
2.3	Bethe equations	5
3	A constant density gas	7
3.1	The exact solution	7
4	The binning procedure	10
4.1	Motivation	10
4.2	Binning in general	10
4.3	Factorization	11
4.4	The two bin case	13
4.5	The rest of this paper	15
5	The constant density matrix element	15
5.1	The master formula	15
5.2	An all even image	18
5.3	Summing over choices of image: arrows	20
5.4	Summing over choices of image: beyond arrows	22
5.5	Summing over choices of image: a systematic approach	24
6	Two densities	25
6.1	When one density vanishes	26
6.1.1	The case: $0 < \alpha \ll 1/2$	26
6.1.2	The case: $1/2 \ll \alpha < 1$	27
6.1.3	The case: $1 < \alpha \ll 3/2$	29
6.2	Equal number of particles at each density	30
6.3	Optimizing the image	32
7	Conclusions	37
A	The decomposition of S_4	38

1 Introduction

A quantum state $|\Psi\rangle$ is completely characterized by its matrix elements $\langle i|\Psi\rangle$ in any basis of the Hilbert space $|i\rangle$. These matrix elements generalize the notion of wave function from quantum mechanics to more general quantum systems. The Coordinate Bethe Ansatz (CBA) [1] computes these matrix elements in the coordinate basis, which consists of eigenvectors of the relevant degrees of freedom. The CBA only exists for systems with finite numbers of degrees of freedom. However, these systems all have a number N , for example the number of lattice sites in a spin chain or the number of particles in the Tonks-Girardeau [2, 3] or Lieb-Liniger [4] models, such that in the formal large N limit one obtains a quantum field theory.

Were it possible to take the large N limit of the CBA, one would, for the first time, have the explicit wave functions for the states of a strongly interacting quantum field theory. This would open exciting possibilities. For example, one could use the equivalence between the high spin XXX spin chain and the $\mathbb{CP}^1$ σ model to understand just how the fractional instanton plasma of [5, 6] is realized in Minkowski space. By studying the difference between the ground state and first excited state wave functions, one may achieve an understanding of the nonperturbative generation of a mass gap [7] in a full quantum field theory with the same concreteness that is achieved in our understanding of the mass gap in the double well quantum mechanics model.

So far, such a large N limit of general wave functions has been lacking because the Bethe Ansatz is complicated. It consists of $N!$ terms, each corresponding to an element of the symmetric group S_N . Many approaches have been developed which simplify the calculation of the matrix elements so that it can be done in a time which is only polynomial in N . But this is still too long for the large N limit. So far, to our knowledge, the large N limit of only one matrix element has ever been computed [8].

In this note we will consider the simplest model which allows a Bethe Ansatz solution, the Tonks-Girardeau model. In this model we will introduce a novel binning approximation scheme for evaluating the matrix elements in the large N limit. Our strategy is essentially to decompose S_N into cosets for each of which the sum can be computed explicitly. For questions of physical interest, like the roles of fractional instantons or unitons in a mass gap, one is interested in matrix elements of Hamiltonian eigenstates with fairly simple coordinate states, or at least with states whose complexities do not themselves depend on N . We will show that for some families of such simple states, our binning prescription allows us to systematically approximate the matrix elements in a time which is independent of N . Therefore we feel that our method can provide a suitable starting point for a large N limit of these states.

While we suspect that a generalization of our method to any Hamiltonian eigenstate is possible, we restrict our attention to the ground state. This choice is largely driven by the fact that the ground state matrix elements are all given by Vandermonde determinants. This allows us to numerically test our approximation scheme for values of N of order 10^4 . We find that the logarithms of the matrix elements can be expanded in powers of $1/N$ and our method allows the leading term in this expansion to be estimated. In some cases we attempt to improve the accuracy of this estimation and we find that we can increase the ac-

Symbol	Description
$\psi(\mathbf{x})$	The wave function for any Hamiltonian eigenstate
$\psi_0(\mathbf{x})$	The real, positive, ground state wave function
L	The size of the circle on which the theory is defined
N	The number of particles
x_j	The position of the j th particle
m_j	A function of x_j written in Coordinate Bethe Ansatz notation
k_j	The j th momentum in the wave function
K_j	A function of k_j written in Coordinate Bethe Ansatz notation
$\mathcal{S}_i$	The i th partition of $[1, N]$ in the domain of the permutation P
F_i	The i th subset of $[1, N]$ in the range of the permutation P
$H_{\{F_i\}}$	The subset of the permutation group S_N that maps $\mathcal{S}_i$ to F_i
$g_{\{F_i\}}$	The monotonic element of $H_{\{F_i\}}$
ψ_F	The contribution to the wave function from a particular F
$\mathcal{F}$	A sequence of 0's and 1's the encodes F
$\uparrow$	Shorthand for the subsequence $\{1, 0\}$ in $\mathcal{F}$
$\downarrow$	Shorthand for the subsequence $\{0, 1\}$ in $\mathcal{F}$
2	Shorthand for the subsequence $\{1, 1\}$ in $\mathcal{F}$
0	Shorthand for the subsequence $\{0, 0\}$ in $\mathcal{F}$
ψ_1	The sum of all ψ_F in which F contains one element from each pair
ψ_2	The sum of all ψ_F in which F consists of arrows, 0's and 2's
α	The density of the x_j in the first bin
ρ	The local density of the x_j

Table 1. List of notation.

curacy as desired at the cost of a longer computation. Critically, the length of computation never depends on N , but only on the complexity of the state and the desired precision.

We begin in section 2 with a review of the Tonks-Girardeau model and its solution using the Bethe Ansatz. Next in section 3 we consider the matrix element of the ground state with the uniform density state. We derive an explicit and exact formula for this matrix element using Vandermonde determinants. In section 4 we present our binning procedure. In section 5 this procedure is applied to calculate a set of improving approximations to the matrix element calculated exactly in section 3. Finally, in section 6, we consider matrix elements with respect to states with two densities. We consider two cases, one in which one density vanishes and another in which the region with each density contains the same number of particles. As this case is more complicated than the single density case, we compute only the leading approximation. We find that it calculates the logarithm of the matrix elements with an error which is roughly independent of N , and in general less than 10%. Table 1 summarizes the notation used in this note.

2 The Tonks-Girardeau model

2.1 The model and its ground state

A Tonks-Girardeau gas is a (1+1)-dimensional, nonrelativistic quantum mechanical model of N impenetrable bosons on a circle of circumference L . The wave function of the N particles satisfies the free nonrelativistic Schrodinger equation

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(\mathbf{x})}{\partial x_j^2} = E\psi \quad (2.1)$$

where

$$\mathbf{x} = \{x_j\}, \quad j \in [1, N], \quad x_j \in [0, L] \quad (2.2)$$

are the coordinates of the N particles. The impenetrability comes from the additional condition

$$\psi(\mathbf{x}) = 0 \text{ if } x_j = x_l \quad (2.3)$$

for any distinct j, l . Periodic boundary conditions are assumed

$$\psi(\mathbf{x}') = \psi(\mathbf{x}) \text{ if } x'_j = x_j \bmod L \quad (2.4)$$

for all j .

Consider the fundamental domain, in which all particle positions are ordered

$$0 < x_1 < \cdots < x_N < L. \quad (2.5)$$

Here no particles touch and so the theory is free. The wave functions in this domain are therefore a superposition of plane waves $e^{2\pi i k x/L}$. The impenetrability condition (2.3) further restricts the space of wave functions on the fundamental domain to the subspace of Slater determinants of N such plane waves

$$e^{2\pi i k_i x/L}, \quad i \in [1, N]. \quad (2.6)$$

In other words, if one defines the matrix

$$M_{ij}(\mathbf{x}) = e^{2\pi i k_i x_j/L} \quad (2.7)$$

then

$$\psi(\mathbf{x}) = \text{Det}(M_{ij}(\mathbf{x})). \quad (2.8)$$

What if the x_i are not ordered? This is a theory of identical bosons, and so the wave function is symmetric under all permutations of the x_i . Therefore the wave function for a general set of x_i is given by first reordering the x_i so that each $x_i < x_{i+1}$ and then using eq. (2.8). As a result of this reordering, the condition imposed by (2.4) on the set of momenta k_i is nontrivial and will be reviewed in subsection 2.3.

Note that if two k_i are equal then the wave function $\psi(\mathbf{x})$ vanishes. The ground state wave function corresponds to lowest values of k^2

$$k_i = i - \frac{N+1}{2}. \quad (2.9)$$

We will now restrict our attention to the choice (2.9).

As the equations (2.1) and (2.3) are linear and real, if $\psi(\mathbf{x})$ is a solution with energy E then so is $\psi^*(\mathbf{x})$. The ground state of the Schrodinger equation is unique up to multiplication by a constant c , as was shown¹ in appendix A of ref. [9]. Therefore $\psi(\mathbf{x}) = c\psi^*(\mathbf{x})$ and so the wave function has a constant phase. As a result $c^{-1/2}\psi$ is real. As it is the ground state, it has no nodes, and so it is always positive or always negative. In the later case, we will multiply it by -1 .

We have shown that $\pm c^{-1/2}\psi(\mathbf{x})$ is real and positive for some choice of sign. Imposing that $|c| = 1$, this is equal to the absolute value of $\psi(\mathbf{x})$

$$\pm c^{-1/2}\psi(\mathbf{x}) = |\psi(\mathbf{x})| = |\text{Det}(M_{ij}(\mathbf{x}))|. \quad (2.10)$$

Adding the same constant C to all of the k_i simultaneously

$$k_i \longrightarrow k_i + C \quad (2.11)$$

has the effect of multiplying the wave function $\psi(\mathbf{x})$ by an x -dependent phase which would change the state (2.8) and in general even change the energy E in (2.1). After changing k_i , $\psi(\mathbf{x})$ would no longer be the ground state and in particular its phase would no longer be constant, and so $\pm c^{-1/2}\psi(\mathbf{x})$ would no longer be equal to $|\psi(\mathbf{x})|$ for any value of c .

More specifically, the shift $k_i \rightarrow k_i + C$ multiplies the j th column of M_{ij} by $e^{2\pi i C x_j / L}$ and so multiplies the determinant $\psi(\mathbf{x})$ by $\prod_j e^{2\pi i C x_j / L}$. The absolute value of the determinant is however unchanged. Therefore if we define

$$\psi_0(\mathbf{x}) = |\psi(\mathbf{x})| = |\text{Det}(M_{ij}(\mathbf{x}))| \quad (2.12)$$

then $\psi_0(\mathbf{x})$ will be invariant under the shift (2.11). At $C = 0$, $\psi_0(\mathbf{x})$ is the ground state wave function. But as $\psi_0(\mathbf{x})$ is independent of C , it will be the ground state wave function at any C . Therefore, we will use (2.12) as our definition of the ground state wave function, and we will choose C to simplify our calculations.

Choosing $C = (N-1)/2$, one obtains

$$k_i = i - 1. \quad (2.13)$$

Now $\psi_0(\mathbf{x})$ is the ground state wave function but $\psi(\mathbf{x})$ is not. Note that the k_i can only be interpreted as momenta if the wave function is $\psi(\mathbf{x})$, defined by eq. (2.8), which it is not.

¹Consider two solutions $f(\mathbf{x})$ and $g(\mathbf{x})$ of (2.1) and (2.3) with the same energy E . Fix some point $\mathbf{x}_0$. Then $h(\mathbf{x}) = g(\mathbf{x}_0)f(\mathbf{x}) - f(\mathbf{x}_0)g(\mathbf{x})$ is another solution with energy E and a zero at $\mathbf{x}_0$. Any solution with a zero is either an excited state or else is everywhere zero. Therefore if $f(\mathbf{x})$ and $g(\mathbf{x})$ are ground states then $h(\mathbf{x}) = 0$ and so $f(\mathbf{x})$ and $g(\mathbf{x})$ are linearly dependent.

Therefore we have not actually modified the momenta, or even the state. On the contrary, $\psi_0(\mathbf{x})$ is proportional to the old $\psi(\mathbf{x})$ obtained from the original momenta (2.9).

With this new choice of k_i , the wave function can be brought into Vandermonde form. Defining the Vandermonde matrix

$$\tilde{M}_{ij} = \eta_j^{(i-1)}, \quad \eta_j = e^{2\pi i x_j / L} \quad (2.14)$$

the wave function is a Vandermonde determinant

$$\psi_0(\mathbf{x}) = \left| \text{Det} \left(\tilde{M}_{ij}(\mathbf{x}) \right) \right| = \prod_{i=1}^{N-1} \prod_{j=i+1}^N |\eta_i - \eta_j|. \quad (2.15)$$

2.2 A Bethe Ansatz formulation

If the x_j or k_i are evenly spaced then the wave function can be brought into Vandermonde form. Whether or not this is the case, the determinant formulas above can all be expanded using the Leibnitz formula for the determinant. We will now provide the expansion of (2.15), the following proceeds identically for (2.8) and (2.12)

$$\psi_0(\mathbf{x}) = \left| \sum_{g \in S_N} (-1)^g \prod_{j=1}^N \tilde{M}_{jP(j)} \right| = \left| \sum_{g \in S_N} (-1)^g \exp \left(\sum_{j=1}^N \frac{2\pi i x_j P(j)}{L} \right) \right|. \quad (2.16)$$

Here g is a permutation of the set of natural numbers $[1, N]$, realized by the bijection $P : [1, N] \rightarrow [1, N]$. The symbol $(-1)^g$ is equal to 1 for an even permutation and (-1) for an odd permutation.

Let us define

$$m_j = x_j \frac{N}{L}, \quad K_j = 2\pi \frac{j}{N}, \quad \Psi_{jk} = 2\pi \theta(j - k) \quad (2.17)$$

where θ is the Heaviside step function. This holds in the case of the ground state, in the case of an excited state a similar derivation using ψ will give the corresponding K_j in terms of the momenta. Then

$$\psi_0(\mathbf{x}) = \left| \sum_{g \in S_N} \exp \left(i \sum_{j=1}^N m_j K_{P(j)} + \frac{i}{2} \sum_{j < k} \Psi_{P(j), P(k)} \right) \right|. \quad (2.18)$$

This is the form of the Coordinate Bethe Ansatz [1], which also describes many more interesting systems. The goal of this paper is to introduce a new method to perform the large N limit of the sum (2.18). The fact that, in the case of the Tonks-Girardeau model, this sum admits a Vandermonde form (2.15) will allow numerical checks of our formulas.

2.3 Bethe equations

For completeness, we will now describe the origin of the Bethe equations in this model. Let us consider an arbitrary state and set of x_i ordered so that

$$0 \leq x_i \leq L, \quad x_i \leq x_{i+1} \quad (2.19)$$

for all $i \in [1, N]$. Going back to eq. (2.8)

$$\begin{aligned}\psi(\mathbf{x}) &= \text{Det} \left(e^{2\pi i k_i x_j / L} \right) \\ &= \sum_{g \in S_N} (-1)^g \exp \left(i \sum_{j=1}^N x_j \frac{2\pi k_{P(j)}}{L} \right) \\ &= \sum_{g \in S_N} \exp \left(i \sum_{j=1}^N x_j \frac{2\pi k_{P(j)}}{L} + \frac{i}{2} \sum_{j < k}^N \Psi_{P(j), P(k)} \right)\end{aligned}\tag{2.20}$$

where Ψ is still given by (2.17).

Let us change the fundamental domain to $[x_i, x_i + L]$ for some fixed $i \in [1, N]$. Such a change of coordinates should not affect the state, and so the wave function must only change by a phase. Let us call the new wave function ψ' . It is still given by eq. (2.20), but now it is understood that all x_j have been shifted to lie in the new fundamental domain

$$x'_j = \begin{cases} x_j + L & \text{if } j < i \\ x_j & \text{if } j \geq i. \end{cases}\tag{2.21}$$

Now recall that (2.20) applies when all x' are ordered, and so the range of the sum $\sum_{j < k}$ needs to be adjusted so that only terms in which $x'_j < x'_k$ are included, yielding

$$\psi'(\mathbf{x}) = \sum_{g \in S_N} \exp \left(i \sum_{j=1}^N x'_j \frac{2\pi k_{P(j)}}{L} + \frac{i}{2} \left(\sum_{0 < j < k}^{i-1} \Psi_{P(j), P(k)} + \sum_{i-1 < j < k}^N \Psi_{P(j), P(k)} + \sum_{k=0}^{i-1} \sum_{j=i}^N \Psi_{P(j), P(k)} \right) \right).\tag{2.22}$$

On the other hand, an infinitesimal downward shift of x_i would put it in the next fundamental domain, where the ordering would yield

$$x''_j = \begin{cases} x_j + L & \text{if } j \leq i \\ x_j & \text{if } j > i \end{cases}\tag{2.23}$$

and so the wave function would become

$$\psi''(\mathbf{x}) = \sum_{g \in S_N} \exp \left(i \sum_{j=1}^N x''_j \frac{2\pi k_{P(j)}}{L} + \frac{i}{2} \left(\sum_{0 < j < k}^i \Psi_{P(j), P(k)} + \sum_{i < j < k}^N \Psi_{P(j), P(k)} + \sum_{k=0}^i \sum_{j=i+1}^N \Psi_{P(j), P(k)} \right) \right).\tag{2.24}$$

By periodicity of the wave function (2.4), $\psi' = \psi''$. However the arguments of the exponentials are different. The xk terms differ when $i = j$, with a difference of $2\pi i k_i$. What about the ψ terms? It is possible but tedious to calculate their difference directly.

A shortcut is to note that, modulo $2\pi i$, they must reproduce the corresponding factors of $(-1)^{g'}$ and $(-1)^{g''}$ where

$$g' = ga, \quad g'' = gb\tag{2.25}$$

and a and b are the permutations which order x' and x'' respectively. For example, a exchanges the $i - 1$ lowest elements with the $N - i + 1$ highest elements. Now the key observation is that

$$b = ac\tag{2.26}$$

where c is just the cyclic permutation which moves i from the beginning to the end. This cyclic permutation is of order N and so $(-1)^c = (-1)^{N-1}$. Therefore, the difference in the Ψ terms in ψ' and ψ'' must just be a factor of $(-1)^{N-1}$, and so moduli $2\pi i$ it is equal to πi if N is even and 0 if N is odd.

The total difference in the arguments of ψ' and ψ'' must vanish moduli 2π , as they are equal. Therefore, adding the contributions from the xk terms and the Ψ terms, we find

$$2\pi k_i = \begin{cases} 2\pi Q_i + \pi & \text{if } N \text{ is even} \\ 2\pi Q_i & \text{if } N \text{ is odd} \end{cases} \quad (2.27)$$

for some integer Q_i . In the case of odd N , this was found in ref. [3]. In the case of even N , it was found in ref. [4]. Recall that the coordinate Bethe Ansatz required us to normalize the momenta as in eq. (2.17)

$$K_i = 2\pi \frac{k_i}{N}. \quad (2.28)$$

With this normalization, (2.27) becomes the Bethe equations

$$NK_i = \begin{cases} 2\pi Q_i + \pi & \text{if } N \text{ is even} \\ 2\pi Q_i & \text{if } N \text{ is odd} \end{cases} \quad (2.29)$$

Note that the usual Ψ terms in the Bethe equations are absent here because Ψ vanishes modulo 2π .

3 A constant density gas

We will begin with the simplest case

$$x_j = j \frac{L}{N} \quad (3.1)$$

corresponding to a constant density gas. In subsection 3.1 we will use the Vandermonde form (2.15) of the ground state wave function to evaluate the ground state wave function $\psi_0(\mathbf{x})$ at this point, or equivalently to find the matrix element

$$\langle \mathbf{x} | \Omega \rangle \quad (3.2)$$

between the ground state wave function $|\Omega\rangle$ and the state $|\mathbf{x}\rangle$ corresponding to a constant density gas (3.1).

Next we will calculate the same matrix element using our binning formalism. We will find that the binning formalism estimates the logarithm of the answer. Successive approximations will be made which calculate this logarithm more and more precisely. The complexity of this calculation is independent of N .

3.1 The exact solution

Let us define the N th root of unity

$$\chi = e^{2\pi i/N}. \quad (3.3)$$

Then the Vandermonde form of the ground state wave function (2.15) may be simplified as follows

$$\begin{aligned}\psi_0(\mathbf{x}) &= \prod_{i=1}^{N-1} \prod_{j=i+1}^N |\chi^i - \chi^j| = \prod_{i=1}^{N-1} \prod_{j=i+1}^N |1 - \chi^{(j-i)}| \\ &= \prod_{k=1}^{N-1} |1 - \chi^k|^{N-k} = \prod_{k=1}^{N-1} |1 - \chi^k|^{N/2} = \left| \prod_{k=1}^{N-1} (1 - \chi^k) \right|^{N/2}.\end{aligned}\quad (3.4)$$

Therefore we need to calculate the term β_N in the absolute value

$$\psi_0(\mathbf{x}) = |\beta_N|^{N/2}, \quad \beta_N = \prod_{k=1}^{N-1} (1 - \chi^k). \quad (3.5)$$

Here β_N is an $(N-1)$ st order polynomial in the χ . Let $\gamma_{N,j}$ be the term with j powers of χ

$$\beta_N = \sum_{j=0}^{N-1} (-1)^j \gamma_{N,j}. \quad (3.6)$$

We will proceed by calculating all of the $\gamma_{N,j}$. As a warm up, let us start with small values of j . Clearly

$$\gamma_{N,0} = 1, \quad \gamma_{N,1} = \sum_{i=1}^{N-1} \chi^i = -1. \quad (3.7)$$

What about $j = 2$?

$$\gamma_{N,2} = \sum_{i_1=1}^{N-2} \sum_{i_2=i_1+1}^{N-1} \chi^{i_1+i_2} = \sum_{i=3}^{2N-3} a_i^{(2)} \chi^i \quad (3.8)$$

where $a_i^{(2)}$ is the number of pairs i_1, i_2 with $1 \leq i_1 < i_2 \leq N-1$ and $i = i_1 + i_2$. Let us defined the reduced summands

$$\tilde{i}_j = i - j, \quad 0 \leq \tilde{i}_1 \leq \tilde{i}_2 \leq N-3, \quad i-3 = \tilde{i}_1 + \tilde{i}_2. \quad (3.9)$$

Now the set $\{\tilde{i}_k\}$ is a two element partition of $i-3$. It is conventional to count the number of nonzero elements in a partition, so then this is a partition of at most two elements. The constants $a_i^{(2)}$ count those partitions of $i-3$ of at most two elements in which each element is bounded by $N-3$. This number is conventionally denoted

$$a_i^{(2)} = p(2, N-3; i-3). \quad (3.10)$$

These functions are known to generate the Gaussian binomial coefficients

$$\sum_{i=0}^{MN} p(M, N; i) \chi^i = \left(\begin{matrix} M+N \\ M \end{matrix} \right)_\chi = \frac{\prod_{j=N+1}^{M+N} (1 - \chi^j)}{\prod_{j=1}^M (1 - \chi^j)}. \quad (3.11)$$

Putting this all together

$$\begin{aligned}
 \gamma_{N,2} &= \sum_{i=3}^{2N-3} p(2, N-3; i-3) \chi^i = \chi^3 \sum_{i=0}^{2(N-3)} p(2, N-3; i) \chi^i = \chi^3 \binom{N-1}{2}_\chi \\
 &= \chi^3 \frac{\prod_{j=N-2}^{N-1} (1 - \chi^j)}{\prod_{j=1}^2 (1 - \chi^j)} = \chi^3 \frac{(1 - \chi^{N-1})(1 - \chi^{N-2})}{(1 - \chi)(1 - \chi^2)} \\
 &= \chi^3 \frac{(1 - \chi^{-1})(1 - \chi^{-2})}{(1 - \chi)(1 - \chi^2)} = (-1)^2 = 1
 \end{aligned} \tag{3.12}$$

where we used the fact that $\chi^N = 1$.

The general case $\gamma_{N,M}$

$$\gamma_{N,M} = \sum_{\{i_j\}} (-1)^{N-1} \chi^{\sum_{j=1}^M i_j}, \quad 1 \leq i_1 < i_2 \dots i_M \leq N-1 \tag{3.13}$$

is no more difficult. Again one may define the coefficients $a_i^{(M)}$ by

$$\gamma_{N,M} = \sum_{i=M(M+1)/2}^{MN-M(M+1)/2} a_i^{(M)} \chi^i. \tag{3.14}$$

These count the allowed values of $\{i_j\}$ or equivalently the reduced

$$\tilde{i}_j = i_j - j, \quad 0 \leq \tilde{i}_1 \leq \tilde{i}_2 \dots \tilde{i}_M \leq N - M - 1, \quad i - \frac{M(M+1)}{2} = \sum_{j=1}^M \tilde{i}_j. \tag{3.15}$$

These are partitions of $i - M(M+1)/2$ consisting of at most M elements such that the maximal value is $N - M - 1$ so

$$a_i^{(M)} = p\left(M, N - M - 1; i - \frac{M(M+1)}{2}\right). \tag{3.16}$$

Finally we may use (3.11) to evaluate $\gamma_{N,M}$

$$\begin{aligned}
 \gamma_{N,M} &= \sum_{i=M(M+1)/2}^{MN-M(M+1)/2} p\left(M, N - M - 1; i - \frac{M(M+1)}{2}\right) \chi^i \\
 &= \chi^{\frac{M(M+1)}{2}} \sum_{i=0}^{2(N-\frac{M(M+1)}{2})} p(M, N - M - 1; i) \chi^i = \chi^{\frac{M(M+1)}{2}} \binom{N-1}{M}_\chi \\
 &= \chi^{\frac{M(M+1)}{2}} \frac{\prod_{j=N-M}^{N-1} (1 - \chi^j)}{\prod_{j=1}^M (1 - \chi^j)} = \frac{\prod_{j=1}^M (1 - \chi^{-j}) \chi^j}{\prod_{j=1}^M (1 - \chi^j)} = (-1)^M.
 \end{aligned} \tag{3.17}$$

We then find

$$\beta_N = \sum_{j=0}^{N-1} (-1)^j \gamma_{N,j} = \sum_{j=0}^{N-1} 1 = N \tag{3.18}$$

and so the matrix element is

$$\psi_0(\mathbf{x}) = N^{N/2}. \quad (3.19)$$

It may seem odd that the matrix element grows with N , but using our conventions

$$\langle \Omega | \Omega \rangle = \int d\mathbf{x} |\psi(\mathbf{x})|^2 = N! L^N \quad (3.20)$$

and the norm of the $|\mathbf{x}\rangle$ states is infinite. To obtain probabilities from these matrix elements, of course one would need to normalize the states. This would not be difficult, but our goal is just to calculate the matrix elements.

4 The binning procedure

4.1 Motivation

The dramatic simplifications in the above calculation occur in part because in eq. (2.17) m_j and K_j are linear in j while Ψ_{jk} is quite simple. The linearity of m_j results from the fact that we consider the matrix element with respect to a constant density distribution, that of K_j from the fact that we consider the ground state and the simplicity of Ψ_{jk} is a consequence of the impenetrability of the particles in this model. The simplicity of Ψ_{jk} occurs for all matrix elements in the model, and that of K_j occurs for all matrix elements of the ground state $|\Omega\rangle$. However, to evaluate the ground state wave function with general particle positions $\mathbf{x}$ one needs to relax the linearity of m_j or equivalently of x_j .

We are interested in the limit of a large number N of particles. In this limit, it makes sense to consider a local density of particles. We do not know whether the wave function will be supported on configurations for which the local density is smooth. Nonetheless, the matrix elements with respect to smooth configurations are the most interesting for physical applications. Our goal is to devise a method to compute such matrix elements.

Our wave function is symmetric under permutations of the x_j [3] and so without loss of generality we may order the x_j so that $x_{j+1} > x_j$. Then if the density is smooth, on sufficiently small scales the x_j will be linearly distributed in j to any desired precision for j in a sufficiently small interval. Therefore we expect simplifications to occur on such small scales. Our strategy will be to decompose the wave function into factors which depend only on such small ranges, to evaluate these factors and then to reassemble them to obtain an approximate wave function.

4.2 Binning in general

More concretely, our procedure for evaluating the sum (2.18) is as follows:

Step 1. Divide the ordered set $\{x_j\}$ into q intervals $\mathcal{S}_i$ called bins. Let n_i be the cardinality of $\mathcal{S}_i$, so that

$$\mathcal{S}_i \subset [1, N], \quad |\mathcal{S}_i| = n_i, \quad [1, N] = \coprod_i \mathcal{S}_i, \quad \mathcal{S}_i \cap \mathcal{S}_j = \emptyset \text{ for } i \neq j \quad (4.1)$$

Step 2. For each bin $\mathcal{S}_i$, choose a disjoint subset

$$F_i \subset [1, N], \quad |F_i| = n_i, \quad [1, N] = \coprod_i F_i, \quad F_i \cap F_j = \emptyset \text{ for } i \neq j \quad (4.2)$$

which also has n_i elements. Recall that each $g \in S_N$ corresponds to a map $P : [1, N] \rightarrow [1, N]$. Let $H_{\{F_i\}} \subset G$ consist of all elements $g \in S_N$ such that the corresponding map satisfies

$$P(\mathcal{S}_i) = F_i \quad (4.3)$$

for all i .

Step 3. Calculate the terms in eq. (2.18) corresponding to $H_{\{F_i\}} \subset S_N$. In other words, sum over only those permutations which fix $P(\mathcal{S}_i) = F_i$.

Step 4. Consider all possible sets of images $\{F_i\}$ and sum over the result of step 3 for each summand.

This procedure, if carried out exactly, will produce the exact wave function. However step 4 is computationally difficult and so we will only perform this sum approximately, yielding an approximate answer.

4.3 Factorization

In this subsection we will consider general energy eigenstates. In other words, k in eq. (2.7) will not be fixed. This means that our wave function ψ will not necessarily have a constant phase, and so we are not justified in taking its absolute value. Therefore our wave function will be ψ and not ψ_0 . We will study step 3 in this setting.

Step 4 states that the wave function ψ can be written as a sum over cosets of summands ψ_F

$$\begin{aligned} \psi(\mathbf{x}) &= \sum_{\{F_i\}} \psi_{\{F_i\}}(\mathbf{x}) \\ \psi_{\{F_i\}}(\mathbf{x}) &= \sum_{g \in H_{\{F_i\}}} \exp \left(i \sum_{j=1}^N m_j K_{P(j)} + \frac{i}{2} \sum_{j < k}^N \Psi_{P(j), P(k)} \right) \\ &= \sum_{g \in H_{\{F_i\}}} (-1)^g \exp \left(i \sum_{j=1}^N m_j K_{P(j)} \right) \end{aligned} \quad (4.4)$$

where we have used the value of $\Psi_{j,k}$ in eq. (2.17) which is specific to the Tonks-Girardeau model and will not generalize to the Lieb-Liniger model. We remind the reader that $(-1)^g$ is the parity of the permutation.

Let us assume that q divides N and choose q constant sized bins

$$\mathcal{S}_i = \left[(i-1) \frac{N}{q} + 1, i \frac{N}{q} \right], \quad i \in [1, q]. \quad (4.5)$$

This restriction on N means that we may miss some states which should exist in the continuum limit, such as the half-integer spin states in the XXX spin chain [10], which arise only at odd values of N [11]. In principle this problem can always be avoided by introducing one additional bin which contains r elements and allowing r to assume each value in the interval $[0, n/q - 1]$. We believe that, even though this extra bin is smaller than the others, the generalization of our construction to this case will be straightforward.

Elements of $H_{\{F_i\}}$ are represented by maps P such that $P(\mathcal{S}_i) = F_i$ for all i . For a given image $\{F_i\}$, choose one such map $P_{\{F_i\}}$ corresponding to any fixed group element $g_{\{F_i\}} \in H_{\{F_i\}}$. Any other element of $H_{\{F_i\}}$ can be obtained by acting on $g_{\{F_i\}}$ from the right by a permutation of each of the subsets $\{F_i\}$. These permutation groups are each $S_{N/q}$ and so a general permutation which fixes all $\{F_i\}$ is an element of the q -fold product group $\times_q S_{N/q}$. The subset $H_{\{F_i\}} \subset S_N$ is then a right coset

$$H_{\{F_i\}} = g_{\{F_i\}}(S_{N/q} \times \cdots \times S_{N/q}). \quad (4.6)$$

The group S_N is the disjoint union of these cosets

$$S_N = \coprod_{\{F_i\}} H_{\{F_i\}} \quad (4.7)$$

which is the reason that the sum over S_N in eq. (2.18) can be decomposed into sums over $H_{\{F_i\}}$ in eq. (4.4). In other words, the decomposition (4.7) justifies step 4. In A we explicitly calculate the decomposition $q = 2$, $N = 4$.

Step 3 on the other hand can be calculated using the decomposition (4.6). It implies that every element of $H_{\{F_i\}}$ can be written as the product of an arbitrary fixed element $g_{\{F_i\}}$ times an element in each of the q copies of the permutation group $S_{N/q}$, one of which acts on each $\mathcal{S}_i$. As a result, the sum over $H_{\{F_i\}}$ in eq. (4.4) can be factored into q sums over the $S_{N/q}$

$$\psi_{\{F_i\}}(\mathbf{x}) = (-1)^{g_{\{F_i\}}} \prod_{i=1}^q \psi_i(\mathbf{x}) \quad (4.8)$$

where

$$\psi_i(\mathbf{x}) = \sum_{g \in S_{N/q}} (-1)^g \exp \left(i \sum_{j \in \mathcal{S}_i} m_j K_{P(j)} \right). \quad (4.9)$$

Eq. (4.8) is the motivation for our binning approach. It states that, up to a phase which we will calculate momentarily, the sum for each F may be calculated explicitly as a product of factors, one for each bin.

We can go further if we approximate the discrete function m to be linear inside each bin

$$m_j = \alpha + \beta j \quad (4.10)$$

corresponding to a constant spacing of the particles in an interval. In general, α and β will depend on the bin i . However to avoid clutter we will simply restrict our attention to a

fixed i . Now

$$\psi_i(\mathbf{x}) = \exp \left[i \left(\alpha + \beta \left((i-1) \frac{N}{q} + 1 \right) \right) \sum_{j \in \mathcal{S}_i} K_{P(j)} \right] \sum_{g \in \mathcal{S}_{N/q}} (-1)^g \exp \left(i \beta \sum_{j=0}^{N/q-1} j K_{P(j)} \right). \quad (4.11)$$

As $P(\mathcal{S}_i) = F_i$, this can be slightly simplified to

$$\psi_i(\mathbf{x}) = \exp \left[i \left(\alpha + \beta \left((i-1) \frac{N}{q} + 1 \right) \right) \sum_{j \in F_i} K_j \right] \sum_{g \in \mathcal{S}_{N/q}} (-1)^g \exp \left(i \beta \sum_{j=0}^{N/q-1} j K_{P(j)} \right). \quad (4.12)$$

The sum is now in the Vandermonde form, and so

$$\begin{aligned} \psi_i(\mathbf{x}) &= (2i)^{\frac{N}{2q} \left(\frac{n}{q} - 1 \right)} \exp \left[i \left(\alpha + \beta \left(\left(i - \frac{1}{2} \right) \frac{N}{q} + \frac{1}{2} \right) \right) \sum_{j \in F_i} K_j \right] \\ &\quad \times \prod_{\{j < k\} \in F_i} \sin \left(\frac{\beta}{2} (K_k - K_j) \right). \end{aligned} \quad (4.13)$$

In the product it is understood that both j and k are elements of F_i . This equation sums the $(N/q)!$ terms in ψ_i , in the case of the matrix element between an arbitrary Hamiltonian eigenstate and a state whose density is approximated to be constant on q intervals, each containing the same number of particles. The full matrix element ψ can be obtained by multiplying q such expressions, following eq. (4.8), thus completing step 3.

4.4 The two bin case

In the rest of the paper we will restrict our attention to the case of two bins. This is because in the general case of an arbitrary number of bins, the phase in eq. (4.13) is a complex number which varies with $\{F_i\}$. While we can evaluate this complex number for each $\{F_i\}$, we still have not devised a systematic expansion which allows us to perform the sum in step 4 to arbitrary precision. This will be the goal of future work. In the two bin case, we will see that there are matrix elements for which each $\{F_i\}$ has the same phase and so this complication is avoided. Therefore in the present paper, we will restrict our attention to the two bin case, for which, at least in the case of some matrix elements, we believe that our techniques produce the exact solution with arbitrary precision.

Let N be even. We will divide $[1, N]$ into two bins

$$\mathcal{S}_A = [1, N/2], \quad \mathcal{S}_B = [N/2 + 1, N]. \quad (4.14)$$

Choose a subset $F \subset [1, N]$ consisting of $N/2$ elements. Recall that each permutation $g \in \mathcal{S}_N$ corresponds to a function $P : [1, N] \rightarrow [1, N]$. Let $H_F \subset \mathcal{S}_N$ be the subgroup such that if $g \in H_F$ then the corresponding function P satisfies

$$P(\mathcal{S}_A) = F. \quad (4.15)$$

We will refer to F as the image, reflecting eq. (4.15).

We have thus completed steps 1 and 2. The $q = 2$ intervals are $\mathcal{S}_A$ and $\mathcal{S}_B$ and the image $P(\mathcal{S}_A)$ is F , whereas $P(\mathcal{S}_B)$ is the complement of F in $[1, N]$. Next for step 3. We need to evaluate

$$\psi_F(\mathbf{x}) = \sum_{g \in H_F} \exp \left(i \sum_{j=1}^N m_j K_{P(j)} + \frac{i}{2} \sum_{j < k}^N \Psi_{P(j), P(k)} \right) \quad (4.16)$$

For example, in the case of the matrix element of the ground state and the constant density state this will simplify to

$$\psi_F(\mathbf{x}) = \sum_{g \in H_F} (-1)^g \chi^{\sum_{j=1}^N j P(j)}. \quad (4.17)$$

As in the q bin case, we may decompose S_N into cosets. Recall that elements of H_F are represented by maps P such that $P(\mathcal{S}_A) = F$. For a given image F , choose one such map P_F corresponding to any fixed group element $g_F \in H_F$. Any other element of H_F can be obtained by acting on g_F from the right by a permutation of the image F and of its complement $[1, N] \setminus F$. These two permutation groups are each $S_{N/2}$ and so H_F is a right coset

$$H_F = g_F(S_{N/2} \times S_{N/2}). \quad (4.18)$$

Note that eq. (4.18) is just the case $q = 2$ of eq. (4.6). Similarly, any element of H_F can be obtained by acting on g_F from the left by a permutation $S_{N/2}$ of $\mathcal{S}_A$ and a permutation $S_{N/2}$ of $\mathcal{S}_B$ and so H_F is also a left coset

$$H_F = (S_{N/2} \times S_{N/2})g_F. \quad (4.19)$$

Using either representation, the key point is that the two copies of $S_{N/2}$ commute with each other. Therefore, just as in the q bin case in eq. (4.8), the ψ_F in eq. (4.17) can be factored into one expression for each copy

$$\psi_F(\mathbf{x}) = (-1)^{g_F} \psi_A \psi_B \quad (4.20)$$

where $(-1)^{g_F}$ is the parity of g_F and ψ_A and ψ_B are respectively sums over the two $S_{N/2}$ factors in eq. (4.18)

$$\begin{aligned} \psi_A &= \sum_{g_A \in S_{N/2}} (-1)^{g_A} \exp \left(i \sum_{j=1}^{N/2} m_j K_{P(j)} \right) \\ \psi_B &= \sum_{g_B \in S_{N/2}} (-1)^{g_B} \exp \left(i \sum_{j=N/2+1}^N m_j K_{P(j)} \right). \end{aligned} \quad (4.21)$$

Thus one may first compute the factors ψ_A and ψ_B and then assemble them into the matrix element.

4.5 The rest of this paper

In the rest of this paper we will systematically apply eq. (4.20) to decompose matrix elements into two bins. In section 5 we consider the case (3.1) of a constant density gas. In this case we already know the exact solution from section 3, and so can easily compare our results. In subsection 5.1 we find that this case enjoys a number of simplifications, and in particular the phases of the sums in step 3 are all equal. This reduces our problem to the sum of a set of positive numbers. In the rest of this section we approximate this sum by an iterative procedure. Our first approximation is simply the largest summand. In the second approximation we sum over a subset which we believe has the largest summands. Eventually we define an iterative scheme for summing larger and larger subsets, yielding better and better approximations of the matrix element. This section is the heart of the paper.

In section 6 we attempt to go beyond the constant density case, while still restricting our attention to two bins. As our basic strategy is to reduce the sum over S_N to that over individual bins via eq. (4.20), first we attempt to try to understand individual bins of constant but arbitrary density. We do this in subsection 6.1 by calculating matrix elements of the ground state with respect to states in which the density is constant not over the whole circle, but only over a subinterval. This calculation is done by approximating the Vandermonde determinant and does not make use of our binning strategy. We also consider constant density not on a subinterval of the circle, but rather of its covering space so that it wraps more than once about the circle. In this case particles from different covers collide at certain densities, yielding a zero matrix element as the particles are impenetrable. These zeroes are not captured by the approximation scheme in section 5, where all summands were positive.

This shortcoming of the approximation in subsection 6.1 is essential in subsection 6.2. Here we consider the case in which there are two regions of distinct densities. We find that if both regions have the same number of particles, then a straightforward generalization of the cancelations in subsection 5.1 implies that the summands in step 3 all have the same complex phase. However, as a result of the zeroes found in subsection 6.1, now there are sign flips and so the contributions have relative signs. Simply ignoring the relative signs, we are able to estimate the logarithms of the matrix elements with errors beneath about 10%. However we have not yet succeeded in turning this into a systematic expansion converging upon the true answer. We feel that this problem must be solved before we can meaningfully attack general matrix elements.

5 The constant density matrix element

5.1 The master formula

In the case of the matrix element of the ground state and the constant density state (3.1), the sum over each bin may be decomposed as in (4.20) into

$$\psi_A = \sum_{g_A \in S_{N/2}} (-1)^{g_A} \chi^{\sum_{j=1}^{N/2} j P_A(j)}, \quad \psi_B = \sum_{g_B \in S_{N/2}} (-1)^{g_B} \chi^{\sum_{j=N/2+1}^N j P_B(j)}$$

where

$$P_A : \mathcal{S}_A \longrightarrow F, \quad P_B : \mathcal{S}_B \longrightarrow [1, N] \setminus F \quad (5.1)$$

represent $g_F g_A$ and $g_F g_B$ respectively.

Simplification 1. Up to a total phase, ψ_A and ψ_B are Vandermonde determinants. This we have seen is true not only for the ground state, but also for all excited states. Each ψ_I is always a Slater determinant and it is equal to a phase times the Vandermonde determinant whenever the x_j with j in the bin $j \in \mathcal{S}_I$ are evenly spaced.²

To bring ψ_A into Vandermonde form, we pull out an overall phase

$$\psi_A = \chi^{\sum_{j \in F} j} \sum_{g_A \in S_{N/2}} (-1)^{g_A} \chi^{\sum_{j=0}^{N/2-1} j P_A(j+1)} = \chi^{\sum_{j \in F} j} \prod_{j < k}^{N/2} \left(\chi^{P_A^F(k)} - \chi^{P_A^F(j)} \right) \quad (5.2)$$

where P_A^F is the map $P_A : [1, N/2] \rightarrow F$ obtained from the action of g_F on $[1, N/2]$. Note that each χ^j appears in the product $N/2 - 1$ times, and so we may pull out the midpoints of the exponents from each factor to obtain

$$\begin{aligned} \psi_A &= \chi^{\left(\frac{N/2-1}{2}+1\right) \sum_{j \in F} j} \prod_{j < k}^{N/2} \left(\chi^{\frac{1}{2}(P_A^F(k) - P_A^F(j))} - \chi^{\frac{1}{2}(P_A^F(j) - P_A^F(k))} \right) \\ &= (2i)^{\frac{N}{4}(\frac{N}{2}-1)} \chi^{\left(\frac{N}{4}+\frac{1}{2}\right) \sum_{j \in F} j} \prod_{j < k}^{N/2} \sin \left(\frac{\pi}{N} (P_A^F(k) - P_A^F(j)) \right). \end{aligned} \quad (5.3)$$

Recall that we are free to choose any g_F such that $P_A^F(\mathcal{S}_A) = F$. Let us fix g_F by demanding, from here on, that the $P_A^F(j)$ are monotonically increasing with j . In this case, the arguments $\frac{\pi}{N} (P_A^F(k) - P_A^F(j))$ of the sine functions on the right hand side of (5.3) are strictly positive.

The factor ψ_B can similarly be put in Vandermonde form

$$\begin{aligned} \psi_B &= \chi^{\left(\frac{N}{2}+1\right) \sum_{j \in [1, N] \setminus F} j} \sum_{g_B \in S_{N/2}} (-1)^{g_B} \chi^{\sum_{j=0}^{N/2-1} j P_B(j+N/2+1)} \\ &= (-1)^{\sum_{j \in F} j} \chi^{\sum_{j \in [1, N] \setminus F} j} \sum_{g_B \in S_{N/2}} (-1)^{g_B} \chi^{\sum_{j=0}^{N/2-1} j P_B(j+N/2+1)}. \end{aligned} \quad (5.4)$$

This is identical to the expression (5.2) for ψ_A except for an overall F -dependent sign. Therefore

$$\psi_B = (-1)^{\sum_{j \in F} j} (2i)^{\frac{N}{4}(\frac{N}{2}-1)} \chi^{\left(\frac{N}{4}+\frac{1}{2}\right) \sum_{j \in [1, N] \setminus F} j} \prod_{j < k}^{N/2} \sin \left(\frac{\pi}{N} (P_B^F(k+N/2) - P_B^F(j+N/2)) \right). \quad (5.5)$$

²In more general models whose states can be put in Bethe Ansatz form, matrix elements are not given by Vandermonde determinants. However, simple cases in which the x are distributed linearly, such as the Néel state in the XXX spin chain, do have known matrix elements [8, 12, 13]. Therefore we hope that in general our binning method will reduce the calculation of piecewise linear x matrix elements to that of linear x matrix elements, which at least in some cases can be found via traditional methods.

Again, we will choose g_F such that P_B^F is monotonically increasing, so the arguments of the sine functions will be positive. By demanding that both P_A^F and P_B^F are monotonically increasing, we have now completely determined the element $g_F \in H_F$. For the rest of this note, g_F will be fixed by this choice.

Multiplying ψ_A by ψ_B , the F -dependence drops out of the χ term, as the j sum is now over the entire range $[1, N]$. Therefore the F -dependence of the phase of ψ_F lies only in the overall signs $(-1)^{g_F}(-1)^{\sum_{j \in F} j}$

$$\begin{aligned} \psi_F(\mathbf{x}) &= (-1)^{g_F}(-1)^{\sum_{j \in F} j} (2i)^{\frac{N}{2}(\frac{N}{2}-1)} \\ &\times \prod_{j < k}^{N/2} \sin\left(\frac{\pi}{N} (P_A^F(k) - P_A^F(j))\right) \sin\left(\frac{\pi}{N} (P_B^F(k + N/2) - P_B^F(j + N/2))\right). \end{aligned} \quad (5.6)$$

Now we will show that this phase is just

$$(-1)^{g_F}(-1)^{\sum_{j \in F} j} = (-1)^{\frac{N}{4}(\frac{N}{2}+1)} \quad (5.7)$$

and so it is also independent of the choice of image F .

First, consider the case $F = [1, N/2]$. In this case g_F is the identity permutation. Therefore

$$(-1)^{g_F} = 1, \quad (-1)^{\sum_{j \in F} j} = (-1)^{\sum_{j=1}^{N/2} j} = (-1)^{\frac{N}{4}(\frac{N}{2}+1)} \quad (5.8)$$

and so (5.7) is satisfied.

Next consider a single permutation which exchanges two elements $j, k \in [1, N]$. Let $j < k$. If both j and k are in F or in $[1, N] \setminus F$ then F is invariant and so the quantity (5.7) cannot change. For concreteness, let $j \in F$ and $k \in [1, N] \setminus F$. Then a permutation which exchanges j and k , acting on g from the right, will yield a new image

$$F' = (F \setminus \{j\}) \cup \{k\}. \quad (5.9)$$

Let j' be the number of elements of F which are less than j and let k' be the number of elements of F which are less than k . Since there are $j-1$ and $k-1$ elements which are lower than j and k in $[1, N]$, then there must be $j-j'-1$ elements of $[1, N] \setminus F$ which are lower than j and $k-k'-1$ which are lower than j .

Now we are ready to calculate the change in $(-1)^{g_F}$ resulting from the exchange of j and k . Note that the resulting permutation $g_{F'}$ is not just the transposition (jk) . This is because once j and k have been exchanged, the elements of F' and $[1, N] \setminus F'$ are no longer in order. k now appears in F in the old position of j , which has j' elements before it. However k has k' elements beneath it in F . One of these elements is j , which is not present in F' , and so k has $k'-1$ elements beneath it in F . The result is that in order to order F' , k needs to be moved to the right $k'-j'-1$ places, requiring $k'-j'-1$ transpositions. Similarly, j now appears after $k-k'-1$ elements of $[1, N] \setminus F'$ but it is only greater than $j-j'-1$ elements, and so to order $[1, N] \setminus F'$ requires $(k-j-k'+j')$ transpositions. Summing the transposition (jk) and the transpositions in F' and its complement we find

$$(-1)^{g_{F'}}(-1)^{g_F} = (-1)^{1+(k'-j'-1)+(k-j-k'+j')} = (-1)^{k-j}. \quad (5.10)$$

The change in the other sign factor is

$$(-1)^{\sum_{j \in F} j} (-1)^{\sum_{j \in F'} j} = (-1)^{2 \sum_{j \in F \cap F'} j} (-1)^{j+k} = (-1)^{j+k} = (-1)^{g_{F'}} (-1)^{g_F} \quad (5.11)$$

and so

$$(-1)^{g_F} (-1)^{\sum_{j \in F} j} = (-1)^{g_{F'}} (-1)^{\sum_{j \in F'} j}. \quad (5.12)$$

Therefore the phase factor is invariant under all interchanges of two elements. But one may obtain any image F' via interchanges of pairs of elements, and so the phase factor is the same for all F , therefore it is always equal to its value when $F = [1, N/2]$ which is given in eq. (5.8). Therefore we have proven (5.7).

Substituting (5.7) into (5.6) we find our master formula for the contribution of one choice of image F to ψ_0

$$\begin{aligned} \psi_F(\mathbf{x}) = & (-1)^{\frac{N}{4}(\frac{N}{2}+1)} (2i)^{\frac{N}{2}(\frac{N}{2}-1)} \\ & \times \prod_{j < k}^{N/2} \sin\left(\frac{\pi}{N} (P_A^F(k) - P_A^F(j))\right) \sin\left(\frac{\pi}{N} (P_B^F(k + N/2) - P_B^F(j + N/2))\right) \end{aligned} \quad (5.13)$$

completing step 3.

Recall that the sine terms are all positive and the phases are manifestly independent of F , and so we have arrived at

Simplification 2. The phase of ψ_F is independent of F .

5.2 An all even image

As we may decompose the symmetric group into cosets

$$S_N = \coprod_F H_F \quad (5.14)$$

we may also decompose the matrix elements into terms of the form (5.13)

$$\psi_0 = \left| \sum_F \psi_F \right|. \quad (5.15)$$

This is step 4 of our method.

As the contributions ψ_F to ψ_0 are all of the same phase, our approach will be to find the largest ψ_F and then to expand about it. The sine functions are maximized when their arguments are closest to $\pi/2$. This occurs when the distances (modulo N) between the elements of F or respectively its complement are largest. Therefore a reasonable guess at the largest ψ_F is the case when F consists of all even or all odd elements of $[1, N]$.

Let us introduce some notation. Define the vector

$$\mathcal{F}_i = \begin{cases} 1 & \text{if } i \in F \\ 0 & \text{otherwise} \end{cases}, \quad i \in [1, N]. \quad (5.16)$$

So $\mathcal{F}$ is a sequence of N 0s and 1s with $N/2$ of each. Let us name two common subsequences

$$\uparrow = \{1, 0\}, \quad \downarrow = \{0, 1\}. \quad (5.17)$$

So the all even subset F corresponds to $\mathcal{F} = \downarrow \cdots \downarrow$ with $N/2$ $\downarrow$'s. Similarly the odd subset F corresponds to $N/2$ $\uparrow$'s. For example, if $\mathcal{F} = \{0, 1, 1, 0\}$ then one could write the shorthand $\mathcal{F} = \downarrow \uparrow$.

In eq. (5.13) one sees that ψ_F only depends upon the differences between the $P_A^F(j)$. These differences are identical for the even and odd subsets, so they yield the same ψ_F . We will now calculate ψ_F for the even subset

$$P_A^{\downarrow \cdots \downarrow}(j) = 2j, \quad P_B^{\downarrow \cdots \downarrow}(j + N/2) = 2j - 1, \quad j \in \mathcal{S}_A = [1, N/2]. \quad (5.18)$$

Now (5.13) becomes

$$\psi_{\downarrow \cdots \downarrow}(\mathbf{x}) = (-1)^{\frac{N}{4}(\frac{N}{2}+1)} (2i)^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{j < k}^{N/2} \sin^2 \left(\frac{2\pi}{N}(k-j) \right). \quad (5.19)$$

The term on the right looks quite similar to the first expression in (3.4).

In that case we saw

$$\psi_0 = \prod_{j < k}^N |\chi^j - \chi^k| = \prod_{j < k}^N |\chi^{(j-k)/2} - \chi^{(k-j)/2}| = (2i)^{N(N-1)/2} \prod_{j < k}^N \sin \left(\frac{\pi}{N}(k-j) \right) = N^{N/2}. \quad (5.20)$$

Replacing N by $N/2$ one finds

$$(2i)^{\frac{N}{4}(\frac{N}{2}-1)} \prod_{j < k}^{N/2} \sin \left(\frac{2\pi}{N}(k-j) \right) = \left(\frac{N}{2} \right)^{N/4}. \quad (5.21)$$

Substituting the square of this identity into (5.19) one finds

$$\psi_{\downarrow \cdots \downarrow}(\mathbf{x}) = (-1)^{\frac{N}{4}(\frac{N}{2}+1)} \left(\frac{N}{2} \right)^{N/2}. \quad (5.22)$$

Remember (3.19) that

$$\psi_0(\mathbf{x}) = \left| \sum_F \psi_F(\mathbf{x}) \right| = N^{N/2} \quad (5.23)$$

as we fixed the wave function to be real and positive. As all ψ_F have the same phase, this phase is inessential. The key point is that $\psi_{\downarrow \cdots \downarrow}$ contributes $(N/2)^{N/2}$ of the total, which is $N^{N/2}$. Stated differently, we may expand

$$\ln(\psi_0(\mathbf{x})) = \alpha N^2 + \beta N \ln(N) + \gamma N + \dots \quad (5.24)$$

The exact result is

$$\alpha = \gamma = 0, \quad \beta = \frac{1}{2}. \quad (5.25)$$

However just the even image $\psi_{\downarrow\ldots\downarrow}$ alone contributes

$$\alpha = 0, \quad \beta = \frac{1}{2}, \quad \gamma = -\frac{\ln(2)}{2}. \quad (5.26)$$

Therefore $\psi_{\downarrow\ldots\downarrow}$ alone exactly yields the leading zero and nonzero terms, although the subleading term is incorrect. This is no surprise, as the other terms in the sum (5.23), corresponding to other images, have not been included.

5.3 Summing over choices of image: arrows

In the rest of this section, we will include more images F in an attempt to improve our estimate of γ . Let us start by replacing the i th $\downarrow$ with an $\uparrow$. Now

$$P_A^{\downarrow\ldots\uparrow\ldots\downarrow}(j) = \begin{cases} 2j & \text{if } j \neq i \\ 2j-1 & \text{if } j = i \end{cases}, \quad P_B^{\downarrow\ldots\uparrow\ldots\downarrow}(j+N/2) = \begin{cases} 2j-1 & \text{if } j \neq i \\ 2j & \text{if } j = i. \end{cases} \quad (5.27)$$

How does the expression (5.13) for $\psi_{\downarrow\ldots\uparrow\ldots\downarrow}$ differ from $\psi_{\downarrow\ldots\downarrow}$? In the latter case, the differences in the P_A^F are all even. Now, when $j = i$, $P_A(j)$ decreases by one unit and so the argument of the first sine decreases by π/N , while that of the second increases. When $k = i$, the first increases and the second decreases. As a result

$$\frac{\psi_{\downarrow\ldots\uparrow\ldots\downarrow}}{\psi_{\downarrow\ldots\downarrow}} = \prod_{j \neq i}^{N/2} \frac{\sin\left(\frac{2\pi}{N}\left(i-j-\frac{1}{2}\right)\right) \sin\left(\frac{2\pi}{N}\left(i-j+\frac{1}{2}\right)\right)}{\sin^2\left(\frac{2\pi}{N}(i-j)\right)}. \quad (5.28)$$

So far our calculation has been exact. Now let us consider an expansion around large N . Choose a natural number $k \ll N$. Consider just those terms of the product such that $|i-j| \leq k$. If the large N limit is taken before the large k limit, then this product converges in k and we may approximate each $\sin(x)$ by x to obtain

$$\frac{\psi_{\downarrow\ldots\uparrow\ldots\downarrow}}{\psi_{\downarrow\ldots\downarrow}} \sim \prod_{j=i-k}^{i+k} \frac{\left(i-j-\frac{1}{2}\right)\left(i-j+\frac{1}{2}\right)}{(i-j)^2} = \left(\prod_{j=1}^k \frac{j^2 - \frac{1}{4}}{j^2}\right)^2 \quad (5.29)$$

where the first product excludes $j = i$. The first approximation, $k = 1$, is

$$\frac{\psi_{\downarrow\ldots\uparrow\ldots\downarrow}}{\psi_{\downarrow\ldots\downarrow}} \sim \frac{9}{16}. \quad (5.30)$$

The approximation $k = 1$ is easy to understand. It means that one only considers adjacent arrows. Each pair corresponds to the following term

$$\downarrow\downarrow=\uparrow\uparrow=\sin^2\left(\frac{2\pi}{N}\right), \quad \downarrow\uparrow=\uparrow\downarrow=\sin\left(\frac{\pi}{N}\right)\sin\left(\frac{3\pi}{N}\right). \quad (5.31)$$

Above we saw sequences of three arrows

$$\downarrow\downarrow\downarrow=\sin^4\left(\frac{2\pi}{N}\right), \quad \downarrow\uparrow\downarrow=\sin^2\left(\frac{\pi}{N}\right)\sin^2\left(\frac{3\pi}{N}\right) \quad (5.32)$$

and so we obtained the ratio

$$\frac{\psi_{\downarrow\cdots\uparrow\cdots\downarrow}}{\psi_{\downarrow\cdots\downarrow}} \sim \frac{\downarrow\uparrow\downarrow}{\downarrow\downarrow\downarrow} \sim \frac{9}{16}. \quad (5.33)$$

More generally, one may use the $k = 1$ approximation (5.31) to find ψ_F for any F that corresponds to a sequence of arrows, in other words any F that includes precisely one member of each pair $\{2j - 1, 2j\}$. The $k = 1$ approximation is reasonable because $k = 1$ yields the contribution to (5.29) which is furthest from unity, and so the most important contribution to the absolute value of the matrix element. This is the case because the corresponding $|i - j| = 1$ term is the most important in (5.28). But there is one exception to this observation. If there is a $|i - j| = N/2 - 1$ term, it will have the same value. Such a term only appears for the arrow at the beginning and at the end of $\mathcal{F}$. Therefore the $k = 1$ approximation means that one considers only adjacent pairs of arrows in the $N/2$ -vector $\mathcal{F}$, where adjacency is understood modulo $N/2$ so that the first and last arrows are also adjacent.

For an arbitrary string of arrows $\mathcal{F}$, let j be the number of occurrences of $\downarrow\uparrow$, with the above cyclicity condition understood. Then j will also be the number of occurrences of $\uparrow\downarrow$ and equivalently the number of strings of consecutive $\uparrow$'s and the number of strings of consecutive $\downarrow$'s. For example $\downarrow\downarrow\uparrow\uparrow\uparrow\uparrow\downarrow$ corresponds to $j = 1$. There are thus $2j$ subsequences $\uparrow\downarrow$ or $\downarrow\uparrow$, each leading to a factor of $3/4$ in the $k = 1$ approximation according to eq. (5.31). Therefore

$$\frac{\psi_F}{\psi_{\downarrow\cdots\downarrow}} = \left(\frac{3}{4}\right)^{2j}. \quad (5.34)$$

How many F 's are there with each j ? One needs to place the $2j$ places where the direction of the arrow changes amongst $N/2$ arrows. At each of these places one inserts $\uparrow\downarrow$ or $\downarrow\uparrow$, which is of length 2. Thus there are a total of $N/2$ slots where one needs to place $2j$ transitions. The number of such choices is $\binom{N/2}{2j}$. Note that one can choose between $\uparrow\downarrow$ and $\downarrow\uparrow$ only once, leading to a single factor of 2 which we will ignore. For example, if the first transition is $\uparrow\downarrow$ then the second will necessarily be $\downarrow\uparrow$ and so on. The corresponding contribution to ψ_0 is ψ_1 where

$$\begin{aligned} \frac{\psi_1}{\psi_{\downarrow\cdots\downarrow}} &= \sum_{j=0}^{N/4} \binom{N/2}{2j} \left(\frac{3}{4}\right)^{2j} = \frac{1}{2} \left(\sum_{j=0}^{N/2} \binom{N/2}{j} \left(\frac{3}{4}\right)^j + \sum_{j=0}^{N/2} \binom{N/2}{j} \left(-\frac{3}{4}\right)^j \right) \\ &= \frac{1}{2} \left(1 + \frac{3}{4}\right)^{N/2} + \frac{1}{2} \left(1 - \frac{3}{4}\right)^{N/2} \sim \left(\frac{7}{4}\right)^{N/2}, \end{aligned} \quad (5.35)$$

and so summing over all such F and ignoring the overall phase we obtain

$$\psi_1 = \left(\frac{7}{4}\right)^{N/2} \left(\frac{N}{2}\right)^{N/2} = \left(\frac{7N}{8}\right)^{N/2} \quad (5.36)$$

or equivalently

$$\gamma = -\frac{\ln(8/7)}{2}. \quad (5.37)$$

5.4 Summing over choices of image: beyond arrows

In general, not every pair $\{2j-1, 2j\}$ will have precisely one element in F . It may have 0 or 2 elements in F . Let us include these cases by adding to our notation for pairs in (5.17)

$$2 = \{1, 1\}, \quad 0 = \{0, 0\}. \quad (5.38)$$

With this short hand, any N -vector $\mathcal{F}$, consisting of 0's and 1's, can be rewritten as an $(N/2)$ -vector consisting of 0's, $\uparrow$'s, $\downarrow$'s and 2's. For example

$$\mathcal{F} = \{0, 1, 1, 1, 1, 0, 0, 0\} \text{ corresponds to } \downarrow 2 \uparrow 0. \quad (5.39)$$

What happens to the calculation in subsection 5.3 if we include a single adjacent pair 20 or 02 starting at the i th position? Recall that in that subsection we adopted the approximation $k = 1$ in which we only considered sine terms in (5.13) corresponding to differences between neighboring symbols, understood modulo $N/2$ so that the first and last symbols are neighbors. Let us refer to the sine terms that we consider as interactions. In that case, there were $N/2$ symbols $\uparrow$ or $\downarrow$, each of which had an interaction with the neighbor on each side leading to a total of $N/2$ interactions in $\mathcal{S}_A$ and also in $\mathcal{S}_B$, for a total of N interactions. Each interaction came with a factor of $1/N$, but these factors cancel in ratios ψ_F 's because there were the same number of interactions, N .

Now it is no longer obvious that the number of interactions is invariant when a 20 replaces two arrows. The two arrows were involved in 3 interactions: one with their neighbor on the left, one with their neighbor on the right and one between the two arrows. When the two arrows are removed, so are these three interactions. Now the 20 has two interactions with the arrow on its left, and one interaction between the two arrows in the 2. So the number of interactions counted is the same when two arrows are replaced by a 20. This would not be the case were the 2 and the 0 separated. Therefore, in the case of 20 and 02 insertions, we can continue to use our $k = 1$ approximation in which only interactions between neighboring symbols are considered, but one must recall that there is now also an interaction inside of the 2.

So what are the interactions of a 20 with its neighbors? The internal interaction gives the difference between the two elements of the 2. These are elements $2i-1$ and $2i$ so their difference is one, giving a factor of $\sin(\frac{\pi}{N})$ for both the $\mathcal{S}_A$ and the $\mathcal{S}_B$ terms in (5.13). The interactions with the neighbors depend on the value of the neighbor. If the left neighbor is $\downarrow$ then $2i-2 \in F$ while $2i-3 \in [1, N] \setminus F$. The former is separated by 1 unit from $2i-1$ and by 2 units from $2i$, both of which are in F , and so one obtains a factor of $\sin(\frac{\pi}{N})\sin(\frac{2\pi}{N})$ from the $\mathcal{S}_A$ terms. If the left neighbor is $\uparrow$, one similarly obtains a factor of $\sin(\frac{2\pi}{N})\sin(\frac{3\pi}{N})$. As neither $2i-1$ nor $2i$ is in $[1, N] \setminus F$, the left neighbor does not contribute to the $\mathcal{S}_B$ terms. The right hand neighbor similarly has two possibilities which lead to these two factors. In all, we have three possible products of the six neighborly

interactions involving the 20

$$\begin{aligned}
 & \sin\left(\frac{\pi}{N}(P_A^F(i) - P_A^F(i-1))\right) \sin\left(\frac{\pi}{N}(P_A^F(i+1) - P_A^F(i-1))\right) \\
 & \times \sin\left(\frac{\pi}{N}(P_A^F(i+1) - P_A^F(i))\right) \sin\left(\frac{\pi}{N}(P_B^F(i+1+N/2) - P_B^F(i+N/2))\right) \\
 & \times \sin\left(\frac{\pi}{N}(P_B^F(i+2+N/2) - P_B^F(i+N/2))\right) \\
 & \times \sin\left(\frac{\pi}{N}(P_B^F(i+2+N/2) - P_B^F(i+1+N/2))\right) \\
 & = \begin{cases} \sin^4\left(\frac{\pi}{N}\right) \sin^2\left(\frac{2\pi}{N}\right) & \text{for } \downarrow 20 \downarrow \\ \sin^3\left(\frac{\pi}{N}\right) \sin^2\left(\frac{2\pi}{N}\right) \sin\left(\frac{3\pi}{N}\right) & \text{for } \downarrow 20 \uparrow \text{ and } \uparrow 20 \downarrow \\ \sin^2\left(\frac{\pi}{N}\right) \sin^2\left(\frac{2\pi}{N}\right) \sin^2\left(\frac{3\pi}{N}\right) & \text{for } \uparrow 20 \uparrow. \end{cases} \quad (5.40)
 \end{aligned}$$

These can be compared with the $\sin^6\left(\frac{2\pi}{N}\right)$ that would be obtained for $\psi_{\uparrow\dots\uparrow}$ or the average $(7/4)^2 \sin^6\left(\frac{2\pi}{N}\right)$ obtained for ψ_1 in subsection 5.3.

Again let us take the large N limit by approximating $\sin(x) \sim x$. Then we obtain the ratios

$$\begin{aligned}
 \frac{\psi_{\downarrow\dots\downarrow 20 \downarrow \dots \downarrow}}{\tilde{\psi}_1} &= \frac{1^4 2^2}{\left(\frac{7}{4}\right)^2 2^6} = \frac{1}{49}, & \frac{\psi_{\downarrow\dots\downarrow 20 \uparrow \dots \downarrow}}{\tilde{\psi}_1} &= \frac{\psi_{\downarrow\dots\downarrow \uparrow 20 \downarrow \dots \downarrow}}{\tilde{\psi}_1} = \frac{1^3 2^2 3^1}{\left(\frac{7}{4}\right)^2 2^6} = \frac{3}{49} \\
 \frac{\psi_{\downarrow\dots\downarrow \uparrow 20 \uparrow \dots \downarrow}}{\tilde{\psi}_1} &= \frac{9}{49}. \quad (5.41)
 \end{aligned}$$

Here we have introduced the notation $\tilde{\psi}_1$ to denote the sum over 8 terms $\psi_{\downarrow\dots\downarrow xxx\downarrow\dots\downarrow}$ where $x = \downarrow$ or $\uparrow$. As a rough approximation, we will simply average these to obtain

$$\frac{\psi_{\dots 20 \dots}}{\psi_1} = \frac{4}{49} \quad (5.42)$$

where $\psi_{\dots 20 \dots}$ is the sum over all ψ with $\mathcal{F}$ consisting entirely of arrows except for a single 20 insertion at a fixed position. The denominator is ψ_1 and not $\tilde{\psi}_1$ as we have now summed over all combinations of arrows even far from the 20 insertion.

What about multiple insertions of 20? This is a length 4 string. In principle it may be inserted at $N/2$ different places in the string of $N/2$ arrows, 0's and 2's. However, many of these positions overlap. To simplify the calculation, let us make the approximation that there are $N/4$ places to insert it, and simply add a degeneracy factor of 2 reflecting the fact that it could be displaced by $1/2$ of a location, ignoring the possible overlaps. Let us add another degeneracy factor of 2 for the fact that 20 may be replaced by 02. And let us ignore the fact that neighboring 20's and 02's will have different interactions than those treated above, where we assumed that the neighbors were arrows.

With these approximations, the number of ways to insert j strings 20 is $\binom{N/4}{j}$ with a weight of $(4/49)^j$ and a degeneracy factor of 4^j . Altogether this yields the sum

$$\frac{\psi_2}{\psi_1} = \sum_{j=0}^{N/4} \binom{N/4}{j} \left(\frac{16}{49}\right)^j = \left(\frac{65}{49}\right)^{N/4} = \left(\frac{\sqrt{65}}{7}\right)^{N/2} \quad (5.43)$$

	$\frac{1}{N} \left(\frac{\psi_{\downarrow \dots \downarrow}}{\psi_0} \right)^{2/N}$	$\frac{1}{N} \left(\frac{\psi_1}{\psi_0} \right)^{2/N}$	$\frac{1}{N} \left(\frac{\psi_2}{\psi_0} \right)^{2/N}$
Estimate Above	0.5	0.875	1.008
$N = 6$	0.5	0.836	1
$N = 8$	0.5	0.824	0.966
$N = 10$	0.5	0.818	0.956

Table 2. Comparison of large N , $k = 1$ estimates above with an explicit Bethe Ansatz sum of the corresponding subsets of images F at $N = 6$, $N = 8$ and $N = 10$. In the last three rows, the three columns sum over (1) $F = \downarrow \dots \downarrow$, (2) over all combinations of arrows and (3) over all combinations of arrows with arbitrary insertions of 20 and 02. Note that the first column agrees for all N as $\psi_{\downarrow \dots \downarrow}$ is calculated exactly. When $N \leq 6$ all possible images F either consist entirely of arrows or have a single 20 or 02 insertion, so the sum over all such images yields the full S_N .

and so

$$\psi_2 = \left(\frac{\sqrt{65}}{8} N \right)^{N/2}, \quad \gamma = \frac{\ln(65)}{4} - \frac{\ln(8)}{2}. \quad (5.44)$$

This is very close to the correct answer. Of course there are further subdominant upwards corrections coming from adding separated pairs of 2 and 0, and further downwards corrections coming from higher k . These corrections each shift γ by a few percent.

In table 2 we compare the $k = 1$ estimates above for the sums over subsets of the images F with the exact sums over those subsets calculated using (2.18) at $N = 6$, $N = 8$ and $N = 10$. The quantity $\psi_{\downarrow \dots \downarrow}$ was calculated exactly above, for any N , and so it is of no surprise that (5.22), reported in the first row of the first column, is equal to the exact results reported in the first column of the later rows. ψ_1 on the other hand is systematically over estimated by (5.36), reported in the second column of the first row, with respect to the exact results in the second column of the later rows. That is as a result of our $k = 1$ approximation. The $k = 2$ correction, for example, multiplies our $(3/4)$ factor by $(15/16)$ yielding $45/64$ and so it reduces the expected $7/8$ to $109/64$. Each k correction in fact is negative. However it is important to remember that they can only be applied to those insertions which are at least a distance k from their neighbors. This correction of course also applies to ψ_2 , reported in the last column of the first row, as it was calculated from ψ_2/ψ_1 in (5.43). With the $k = 2$ correction, ψ_2 will be less than ψ_0 , as it must be since it is a partial sum over positive contributions to ψ_0 . In our examples we also see that of course ψ_2 is less than ψ_0 . With more work, the estimates could become more precise. However in the table we see that $\psi_2^{2/N}$ already provides an estimate of $\psi_0^{2/N}$ with an error of less than 5%.

5.5 Summing over choices of image: a systematic approach

The above system of ever increasing precision by summing over more possible images F can be made systematic. P is a map from $[1, N]$ to $[1, N]$. We always binned the domain into 2 bins, $\mathcal{S}_A$ and $\mathcal{S}_B$. In the first step, in subsection 5.2, we choose only the value of F which gave the largest ψ_F , corresponding to all even or odd elements of $[1, N]$. This gave us $\psi_{\uparrow \dots \uparrow} = (N/2)^{N/2}$.

Next in subsection 5.3 we also binned the range $[1, N]$, into $N/2$ bins $\{2j-1, 2j\}$ and we only considered those images F which contained precisely one value from each bin. If $f_{Ij}(F)$ is the number of elements of $\mathcal{S}_I$ that are mapped into $\{2j-1, 2j\}$ then this corresponds to the case with all values equal to unity

$$f_{Ij}(F) = 1. \quad (5.45)$$

We are trying to sum over all $g \in S_N$. We have partitioned S_N into cosets $H_F \subset S_N$ each labeled by F . Now we have further partitioned F into bins labeled by the $2 \times N/2$ matrix f_{Ij} with entries 0, 1 and 2. We considered only the sine terms, which we called interactions, in (5.13) which connect neighboring bins j and j' . Therefore in subsection 5.3 we have again considered the contributions to ψ_0 arising from the dominant bin, which now is (5.45). Whereas in subsection 5.2 the dominant bin, $\uparrow \cdots \uparrow$, contained $(N/2)!^2$ elements of the $N!$ elements in S_N , now the dominant bin (5.45) includes permutations of the sets $\{2j-1, 2j\}$ and so contains $2^{N/2}(N/2)!^2$ elements. This bigger sum gave us a contribution of $\psi_1 = (7N/8)^{N/2}$.

Actually one may make a similar interpretation of subsection 5.2, where $\mathcal{S}$ is the first row of a $2 \times N$ matrix f_{Ij} which gives the number of elements of $\mathcal{S}_I$ equal to j , which is necessarily 0 or 1 as j is a single element. In this sense the difference between subsections 5.3 and 5.2 is that the first partitioned the image $[1, N]$ into $N/2$ bins while the second partitioned it into N bins, so that each element was its own bin.

The next logical step would be to partition the image into $N/4$ bins and calculate the contribution of the dominant choice

$$f_{Ij} = 2 \quad (5.46)$$

only considering interactions between neighboring bins. We have not systematically done this. However the condition (5.46) corresponds to 20, 02, $\uparrow\uparrow$, $\uparrow\downarrow$, $\downarrow\uparrow$ and $\downarrow\downarrow$ which are the four combinations considered in subsection 5.4. There our rough approximations, including strings 20 and 02, led us to the sum $\psi_1 = (\sqrt{65}N/8)^{N/2}$, ever closer to the exact answer of $N^{N/2}$.

6 Two densities

We expect that the sum over permutations in the Bethe Ansatz (2.18) is simpler when the m_i or equivalently the x_i are distributed linearly in i . The purpose of binning is that while the x_i are generally not linearly distributed for a matrix element of interest, by considering sufficiently small bins, the x_i may be linear in each bin up to any desired precision.

In section 5 the x_i were already distributed linearly, as the density was taken to be constant. The purpose of that section was merely to show how the binning approximation can be implemented systematically. However it did not lead to any simplification with respect to the direct calculation in section 3, quite on the contrary.

In this section we will therefore consider matrix elements of the Tonks-Girardeau ground state with states with two regions with distinct densities. In subsection 6.1 one of these densities will be zero. In subsection 6.2, each region will have the same number

of particles. In these cases, the Vandermonde determinant form for the ground state is of complexity N^2 , and so in principle requires at least time N^2 to evaluate. In each case we will extract the leading order behavior of the matrix elements in a computation whose complexity which is manifestly independent of N , and so will arrive at results which may be a suitable starting place for a large N limit.

We only apply binning to the second case, the role of the first case is to better understand the second. In the first case we bin the Vandermonde product directly, and so bypass the complicated partitioning of the permutation group S_N . However beyond the ground state in this model and also in other models treatable with the Bethe Ansatz, there is no Vandermonde product form, and so this method has no hope of generalizing.

6.1 When one density vanishes

In this case we will consider particle positions of the form

$$x_j = \frac{\alpha L}{N} j \quad (6.1)$$

where α is an arbitrary real number. Then (2.14) yields

$$\eta_j = e^{2\pi i x_j / L} = e^{2\pi i \alpha j / N} = \chi^{\alpha j / N} \quad (6.2)$$

and the Vandermonde determinant of the matrix η_j^k provides arbitrary matrix elements

$$\begin{aligned} \psi_0(\mathbf{x}) &= \prod_{j < k}^N |\eta_k - \eta_j| = 2^{N(N-1)/2} \prod_{j < k}^N \left| \sin \left(\frac{\pi \alpha}{N} (k - j) \right) \right| \\ &= 2^{N(N-1)/2} \prod_{j=1}^{N-1} \left| \sin \left(\frac{\pi \alpha}{N} j \right) \right|^{N-j}. \end{aligned} \quad (6.3)$$

This expression for the wave function is exact, but to calculate it requires a time polynomial in N and so it is not suitable for a large N limit. To simplify it, we will need to choose a regime for α . The case $\alpha = 1$ was the subject of section 3.

6.1.1 The case: $0 < \alpha \ll 1/2$

In this case the argument of the sine function is always much less than $\pi/2$ and so we may use the approximation $\sin(x) = x$ to obtain

$$\begin{aligned} \psi_0(\mathbf{x}) &\sim 2^{N(N-1)/2} \prod_{j=1}^{N-1} \left(\frac{\pi \alpha}{N} j \right)^{N-j} = \left(\frac{2\pi \alpha}{N} \right)^{N(N-1)/2} \prod_{j=1}^{N-1} (j!) \\ &\sim \left(\frac{2\pi \alpha}{N} \right)^{N(N-1)/2} \prod_{j=1}^{N-1} e^{j \ln(j) - j} = \left(\frac{2\pi \alpha}{eN} \right)^{N(N-1)/2} \exp \left(\sum_{j=1}^{N-1} j \ln(j) \right) \\ &\sim \left(\frac{2\pi \alpha}{eN} \right)^{N(N-1)/2} \exp \left(\int_{j=1}^{N-1} dj \, j \ln(j) \right) \\ &= \left(\frac{2\pi \alpha}{eN} \right)^{N(N-1)/2} \exp \left(\frac{j^2}{2} \ln(j) - \frac{j^2}{4} \Big|_1^N \right) \\ &\sim \left(\frac{2\pi \alpha}{e^{3/2}} \right)^{N^2/2}. \end{aligned} \quad (6.4)$$

Figure 1. Comparison of eq. (6.4) (red) versus the exact matrix elements calculated using the Vandermonde formula (6.3) (black) at $N = 20$ (top left), $N = 100$ (top right) and $N = 200$ (bottom).

Note that the leading term in the exponent, which is of order $N^2 \ln(N)$ in the penultimate line of (6.4), cancels in the final expression so long as α is held fixed in the large N limit.

Even so, the norm of the state is of order $e^{N \ln(N)}$ and the volume of the N -dimensional coordinate space is of order L^N , so matrix elements of order e^{-N^2} will have a measure zero contribution to the wave function even if they are all integrated over. Thus states with $\alpha \ll 1/2$ do not contribute to the wave function in the large N limit.

In figure 1 we compare our approximation (6.4) to the exact Vandermonde formula and find good agreement at large N and $\alpha < 0.5$.

6.1.2 The case: $1/2 \ll \alpha < 1$

We will make use of the identities

$$\prod_{j=1}^{M-1} \left(1 - e^{\frac{2\pi i j}{M}}\right) = M \quad (6.5)$$

and

$$\prod_{j=1}^{M-1} \left(1 - e^{\frac{2\pi i j}{M}}\right)^{M-j} = M^{M/2} \quad (6.6)$$

which were proved in subsection 3.1. These allow us to use the fact that α is near 1 to cleanly separate the N^2 terms in the exponent.

At large N , without loss of generality we can take N/α to be an integer. Then

$$\psi_0(\mathbf{x}) = \left| \prod_{j=1}^{N-1} (1 - \chi^{\alpha j})^{N-j} \right| = \left| \prod_{j=1}^{N/\alpha-1} (1 - \chi^{\alpha j})^{N-j} \right| / \left| \prod_{j=N}^{N/\alpha-1} (1 - \chi^{\alpha j})^{N-j} \right| = AB \quad (6.7)$$

where A and B are defined to be the two factors in the penultimate term. These can be evaluated separately

$$\begin{aligned} A &= \left| \prod_{j=1}^{N/\alpha-1} (1 - \chi^{\alpha j})^{N-j} \right| = \left| \prod_{j=1}^{N/\alpha-1} (1 - \chi^{\alpha j})^{N/\alpha-j} \right| \left| \prod_{j=1}^{N/\alpha-1} (1 - \chi^{\alpha j}) \right|^{N-N/\alpha} \\ &= \left(\frac{N}{\alpha} \right)^{N/(2\alpha)} \left(\frac{N}{\alpha} \right)^{N-N/\alpha} = \left(\frac{N}{\alpha} \right)^{N(1-\frac{1}{2\alpha})} \end{aligned} \quad (6.8)$$

where the two factors in the first line were evaluated using the identities (6.6) and (6.5) respectively, with $M = N/\alpha$.

Next

$$\begin{aligned} B &= \left| \prod_{j=N}^{N/\alpha-1} (1 - \chi^{\alpha j})^{j-N} \right| = \left| \prod_{j=N-N/\alpha}^{-1} (1 - \chi^{\alpha(j+N/\alpha)})^{j-N+N/\alpha} \right| \\ &= \left| \prod_{j=1}^{N/\alpha-N} (1 - \chi^{\alpha j})^{N/\alpha-N-j} \right| \sim \left| \prod_{j=1}^{N(1/\alpha-1)} \left(1 - \exp \left(\frac{2\pi i(1-\alpha)j}{N(1/\alpha-1)} \right) \right)^{N(1/\alpha-1)-j} \right|. \end{aligned} \quad (6.9)$$

This is similar to ψ_0 in case 6.1.1 but with

$$N \longrightarrow N' = N(1 - 1/\alpha), \quad \alpha \longrightarrow \alpha' = 1 - \alpha \quad (6.10)$$

so that $0 < \alpha' \ll 1/2$. With these substitutions, eq. (6.4) yields

$$B \sim \left(\frac{2\pi(1-\alpha)}{e^{3/2}} \right)^{\frac{N^2(1-\frac{1}{\alpha})^2}{2}}. \quad (6.11)$$

Substituting A and B into eq. (6.7) we find the final result

$$\psi_0(\mathbf{x}) \sim \left(\frac{2\pi(1-\alpha)}{e^{3/2}} \right)^{\frac{N^2(1-\frac{1}{\alpha})^2}{2}} \left(\frac{N}{\alpha} \right)^{N(1-\frac{1}{2\alpha})}. \quad (6.12)$$

As expected, when $\alpha = 1$ at finite N , this expression equals (3.19). The limit of interest of course is $\alpha \longrightarrow 1$, $N \longrightarrow \infty$. In this case, it differs from (3.19) when the first term differs from unity, which is the case in which $N^2(1-\alpha)^2 \ln(1-\alpha)$ does not tend to 0.

To see this, let us define

$$\epsilon = 1 - \alpha. \quad (6.13)$$

The leading order terms in (6.12) as $\epsilon \rightarrow 0$ are

$$\begin{aligned} \ln(\psi_0(\mathbf{x})) &\sim \frac{N}{2} \ln(N) + \frac{N^2 \epsilon^2}{2} \left(\ln(\epsilon) + \ln\left(\frac{2\pi}{e^{3/2}}\right) \right) + \frac{N\epsilon}{2} (\ln(N) + 1) + \frac{N\epsilon^2}{2} \\ &\sim \frac{N}{2} \ln(N) + \frac{N^2 \epsilon^2}{2} \ln(\epsilon) + \frac{N\epsilon}{2} \ln(N). \end{aligned} \quad (6.14)$$

The leading term is just (3.19). The corrections will tend to 0 in the large N limit if $N\epsilon \rightarrow 0$, in other words if ϵ shrinks faster than $1/N$. If $N\epsilon$ tends to a nonzero constant, then both constants are of the same order. Otherwise, $N\epsilon \rightarrow \infty$, in which case the N^2 term dominates, corresponding to the first term in (6.12). Which case is relevant?

Recall that the normalization of our energy eigenstates is of order $N!$ and so its logarithm is of order $N \ln(N)$. Also, if x is discretized then the norms of the position states will have logarithms of order N . Therefore we expect contributions from of order e^N neighboring states, and terms in the logarithm of the matrix element of order N and $N \ln(N)$ may contribute to finite quantities. For higher order terms however, no amount of normalization and integration can avoid the fact that one finds zero in the large N limit. This argument suggests that the dominant contributions to deformations will come from those that correct $\ln(\psi_0(\mathbf{x}))$ with corrections of order N or $N \ln(N)$, and so correspond to

$$N^2 \epsilon^2 \ln(\epsilon) \sim N \quad (6.15)$$

and so

$$\epsilon \sim N^{-1/2}. \quad (6.16)$$

In this case it is the first term in (6.12) which dominates. This is fortunate, as we will soon see that the second is the most difficult to estimate in general.

In figure 2 we compare our approximation (6.12) to the exact Vandermonde formula and again find good agreement at large N and $0.5 < \alpha < 1$.

6.1.3 The case: $1 < \alpha \ll 3/2$

This is very similar to the previous case. However now some signs change in the evaluation of B and so we choose

$$N \rightarrow N' = N(1/\alpha - 1), \quad \alpha \rightarrow \alpha' = \alpha - 1 \quad (6.17)$$

which lead to

$$\psi_0(\mathbf{x}) \sim \left(\frac{2\pi(\alpha - 1)}{e^{3/2}} \right)^{\frac{N^2(1-\frac{1}{\alpha})^2}{2}} \left(\frac{N}{\alpha} \right)^{N(1-\frac{1}{2\alpha})}. \quad (6.18)$$

Now the families of matrix elements which contribute to the wave function are neighborhoods of the cases in which $\alpha - 1$ shrinks at least as quickly as $N^{-1/2}$.

In figure 3 we compare our approximation (6.18) to the exact Vandermonde formula and again find reasonable agreement at large N and $1 < \alpha < 1.3$. The exact formula produces a series of sharp dips, corresponding to the values of α at which $\alpha j = N$ for some

Figure 2. Comparison of eq. (6.12) (red) versus the exact matrix elements calculated using the Vandermonde formula (6.3) (black) at $N = 20$ (top left), $N = 100$ (top right) and $N = 200$ (bottom).

$j < N$. Physically these are points where two particles coincide, which have zero wave function due to the impenetrability condition (2.3). The Vandermonde formula has zeroes at

$$\alpha = \frac{N}{j}, \quad j < N. \quad (6.19)$$

These zeroes are not captured by our approximation. At large N the zeroes become thin, but their density increases as N . We do not know whether they persist in the large N limit, or whether they lead to a physically relevant drop in the wave function at $\alpha > 1$ in this limit. However we will see below that in the two bin case, this regime does lead us to systematically overestimate matrix elements.

6.2 Equal number of particles at each density

Finally we are ready for a nontrivial application of the binning formalism. We will consider an equal number of particles at density α and $2 - \alpha$ in two contiguous regions

$$x_j = \begin{cases} \frac{L}{N}\alpha j & \text{if } j \leq N/2 \\ \frac{L}{N}(2 - \alpha)j + L(\alpha - 1) & \text{if } j \geq N/2. \end{cases} \quad (6.20)$$

Recall that our Bethe Ansatz (2.18) describes the wave function in terms of a map $P : [1, N] \rightarrow [1, N]$ where the domain corresponds to our x 's and the range to the momenta, which are integers. We will again consider the two bins $\mathcal{S}_A$ and $\mathcal{S}_B$ defined in eq. (4.14). Our binning approach to the constant density case in section 3 suggests that we can systematically improve our estimates of the leading order term in $\ln(\psi_0)$ by considering

Figure 3. Comparison of eq. (6.18) (red) versus the exact matrix elements calculated using the Vandermonde formula (6.3) (black) at $N = 20$ (top left), $N = 100$ (top right) and $N = 200$ (bottom).

more and more images of $\mathcal{S}_A$. Our analysis of the constant density case in subsection 6.1 suggests that now this leading order will consist of the N^2 terms in $\ln(\psi_0)$.

We now face new complications with respect to section 3. First, as the distance between some pairs of x is greater than L/N , the argument of some of the sine functions in (5.13) will be negative. Second, there is no symmetry condition which determines which set F should lead to the dominant contribution to the matrix elements.

As the x_j in each bin have constant separation, the results of subsection 4.4 largely carry over to this case. The symmetric group S_N can again be divided into cosets labeled by F , defined such that $P(\mathcal{S}_A) = F$ for all g in the coset. Then sum of (2.18) over the coset yields a formula similar to eq. (4.17)

$$\begin{aligned} \psi_F(\mathbf{x}) &= \sum_{g \in H_F} \exp \left(i \sum_{j=1}^N m_j K_{P(j)} + \frac{i}{2} \sum_{j < k}^N \Psi_{P(j), P(k)} \right) \\ &= \sum_{g \in H_F} (-1)^g \chi^\alpha \sum_{j=1}^{N/2} j P(j) + \sum_{j=N/2+1}^N ((2-\alpha)j + (\alpha-1)N) P(j). \end{aligned} \quad (6.21)$$

The average value of x in $\mathcal{S}_A$ still differs from that of $\mathcal{S}_B$ by $L/2$. Therefore again, up to an overall phase, ψ_F can be written as a product $\psi_A \psi_B$.

As the x_j are linear in each bin, these can again be written as Vandermonde determinant and so as a product of sines, now with factors of α and $2-\alpha$ respectively. Recombining

them we obtain the analogue of eq. (5.13)

$$|\psi_F(\mathbf{x})| = 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{j < k}^{N/2} \sin\left(\frac{\pi\alpha}{N} (P_A^F(k) - P_A^F(j))\right) \\ \times \sin\left(\frac{\pi(2-\alpha)}{N} (P_B^F(k + N/2) - P_B^F(j + N/2))\right). \quad (6.22)$$

Again the phases are F -independent except for the sines themselves, which are now negative if their arguments are greater than π . Therefore the ψ_F will all contribute to ψ_0 with the same phase up to a sign. In principle we could calculate the distribution of these signs to refine our estimate of ψ_0 , however we will leave this to future work.

The zeros and sign changes of the sine arise identically to those in eq. (6.19) in the case of a constant density. In that case they had a very physical origin, they were the zeroes of the wave function ψ_0 when two particles overlap. Now the particle positions are given by (6.20) and no particles overlap, therefore the total wave function ψ_0 does not vanish.

Instead the zeroes in (6.22) are zeroes in ψ_F , corresponding to the sum over a single image F . These zeroes come from the fact that the wave function would have vanished had the same x spacing α or $2 - \alpha$ persisted over all of $[1, N]$ instead of just the bin $\mathcal{S}_A$ or $\mathcal{S}_B$. More concretely, they arise from the fact that the wave function Slater determinant contains all behaviors from $e^{-i(N-1)x/2L}$ to $e^{i(N-1)x/2L}$ and so if two adjacent x 's differ by $L/(N-1)$, for example, corresponding to $\alpha = N/(N-1)$, then a minor in the determinant which includes $e^{-i(N-1)x/2L}$ for one x_j and $e^{i(N-1)x/2L}$ for its neighbor will vanish. In this sense, the zeroes of (6.22) are zeroes in a minor and not in the determinant so they are spurious and it would be desirable if (2.18) could be resummed in such a way that they are not present from the beginning.

Our goal will now be to make the crudest possible estimate of ψ_0 , as a ψ_F which is in a sense dominant. What do we mean by dominant? Different definitions give different estimates. One choice of image is just $F = \downarrow \cdots \downarrow$ as before. In figure 4 this choice is compared with the exact ψ_0 calculated using (2.14).

6.3 Optimizing the image

How does our estimate of ψ_0 depend upon our choice of F . Recall that F is a subset of $[1, N]$ consisting of $N/2$ elements. We can parameterize this interval by

$$x = \frac{i}{N}, \quad i \in [1, N] \quad (6.23)$$

and in the large N limit we can associate a local density

$$0 \leq \rho(x) \leq 1 \quad (6.24)$$

with the subset $F \in [1, N]$.

Clearly F determines ψ_F . And F determines $\rho(F)$. But to what extent does $\rho(F)$ determine ψ_F ? We claim that $\rho(F)$ fixes the dominant term in $\ln(\psi_0)$ and also perhaps the subleading term. We cannot prove this claim, but we will provide an example.

Figure 4. Comparison of the estimator in subsection 6.3 (red) with a discretization into 25 parts versus $\ln(\psi_{\downarrow\ldots\downarrow})$ (green) and the exact matrix elements $\ln(\psi_0)$ calculated using the Vandermonde formula (2.14) (black) at $N = 100$ (top left), $N = 250$ (top right), $N = 500$ (middle left), $N = 1000$ (middle right), $N = 2500$ (bottom left) and $N = 5000$ (bottom right).

Consider two choices of image F . One is $F_1 = F_{\downarrow\ldots\downarrow}$ which is the subset of even elements. The second, $F_2 = F_{\downarrow\uparrow\downarrow\ldots\uparrow}$ alternates between $\uparrow$ and $\downarrow$, corresponding to $\mathcal{F} = \{011001100\cdots 01\}$. For simplicity we will first consider $\alpha = 1$. In both cases the separations between elements of F and $[1, N] \setminus F$ are equal, and so $\psi_A = \psi_B$.

In the first case

$$\begin{aligned}
 |\psi_{F_1}(\mathbf{x})| &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{j < k}^{N/2} \sin\left(\frac{\pi}{N} (P_A^F(k) - P_A^F(j))\right)^2 \\
 &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{j < k}^{N/2} \sin\left(\frac{2\pi}{N} (k - j)\right)^2 = \left(\frac{N}{2}\right)^{N/2}, \quad (6.25)
 \end{aligned}$$

and in the second

$$\begin{aligned}
 |\psi_{F_2}(\mathbf{x})| &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{k=1}^{N/4-1} \sin^{N/2-2k} \left(\frac{\pi}{N} (4k-3) \right) \sin^{N/2-2k} \left(\frac{\pi}{N} (4k-1) \right) \sin^{N-4k} \left(\frac{\pi}{N} (4k) \right) \\
 &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{k=1}^{N/4-1} \sin^{N/4} \left(\frac{\pi}{N} (4k-3) \right) \sin^{N/4} \left(\frac{\pi}{N} (4k-1) \right) \sin^{N/2} \left(\frac{\pi}{N} (4k) \right) \\
 &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \left(\prod_{k=1}^{N-1} \sin \left(\frac{\pi}{N} k \right) \right)^{N/4} \left(\prod_{k=1}^{N/2-1} \sin \left(\frac{2\pi}{N} k \right) \right)^{-N/4} \left(\prod_{k=1}^{N/4-1} \sin \left(\frac{4\pi}{N} k \right) \right)^{N/2} \\
 &= N^{N/4} \left(\frac{N}{2} \right)^{-N/4} \left(\frac{N}{4} \right)^{N/2} = \left(\frac{N}{2\sqrt{2}} \right)^{N/2}. \tag{6.26}
 \end{aligned}$$

In this case the change in local density only affects the $O(N)$ term in $\ln(\psi)$ not the dominant $O(N \ln(N))$ term.

At an arbitrary α , we need only insert factors of α in the first lines of eqs. (6.25) and (6.26). Let us consider the case $\alpha \ll 1$, so that we may simplify the calculation as in the first case of subsection 6.1. Now the first wave function is

$$\begin{aligned}
 |\psi_{F_1}(\mathbf{x})| &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{j < k}^{N/2} \sin \left(\frac{2\pi\alpha}{N} (k-j) \right)^2 \\
 &\sim \left(\frac{8\pi\alpha}{N} \right)^{\frac{N}{2}(\frac{N}{2}-1)} \left(\prod_{j=1}^{N/2-1} j! \right)^2 \sim \left(\frac{4\pi\alpha}{e} \right)^{N^2/4} \tag{6.27}
 \end{aligned}$$

and the second is

$$\begin{aligned}
 |\psi_{F_2}(\mathbf{x})| &= 2^{\frac{N}{2}(\frac{N}{2}-1)} \prod_{k=1}^{N/4-1} \sin^{N/2-2k} \left(\frac{\pi\alpha}{N} (4k-3) \right) \\
 &\quad \times \sin^{N/2-2k} \left(\frac{\pi\alpha}{N} (4k-1) \right) \sin^{N-4k} \left(\frac{\pi\alpha}{N} (4k) \right) \\
 &\sim 2^{N/2} \left(\frac{8\pi\alpha}{N} \right)^{N(\frac{N}{4}-1)} \left(\prod_{j=1}^{N/4-1} \Gamma^2(j+1) \frac{\Gamma(j+3/4)}{\Gamma(3/4)} \frac{\Gamma(j+1/4)}{\Gamma(1/4)} \right) \\
 &\sim \left(\frac{4\pi\alpha}{e} \right)^{N^2/4} \tag{6.28}
 \end{aligned}$$

where in each evaluation we have used the Sterling approximation to approximate the Gamma function and then replaced the product with an integral as in eq. (6.4). In each case the $O(N^2 \ln(N))$ term in the logarithm vanishes, and the leading $O(N^2)$ terms agree in the two cases. We suspect that at general α the leading order behavior of the wavefunction will be determined by $\rho(F)$. Indeed, were this not the case, one might doubt whether a continuum limit of this model exists.

Thus the leading behavior of $\ln(\psi)$ appears to be entirely determined by $\rho(x)$. How do we optimize $\rho(x)$? Let us take ψ_A and ψ_B to be real and positive. We will now allow

α to be an arbitrary real number. In this case the arguments of the sine terms in (6.25) and (6.26) do not necessarily lie in the interval $[0, \pi]$. However the absolute value on the left hand side of these equations means that each sine term should be $|\sin|$. Then

$$\begin{aligned}\psi_A &= 2^{\frac{N}{4}(\frac{N}{2}-1)} \text{Exp} \left(\sum_{j < k}^{N/2} \ln \left| \sin \left(\frac{\pi\alpha}{N} (P_A^F(k) - P_A^F(j)) \right) \right| \right) = 2^{\frac{N}{4}(\frac{N}{2}-1)} e^{c_A} \\ \psi_B &= 2^{\frac{N}{4}(\frac{N}{2}-1)} \text{Exp} \left(\sum_{j < k}^{N/2} \ln \left| \sin \left(\frac{\pi(2-\alpha)}{N} (P_B^F(k + N/2) - P_B^F(j + N/2)) \right) \right| \right) \\ &= 2^{\frac{N}{4}(\frac{N}{2}-1)} e^{c_B}\end{aligned}\quad (6.29)$$

where c_A and c_B are defined by the preceding expressions. In the continuum limit the sums are replaced by integrals over x and y and the set F is replaced by its density $\rho(x)$

$$\begin{aligned}c_A &\sim \int_0^1 dx \int_{x+\epsilon}^1 dy \rho(x) \rho(y) \ln |\sin(\pi\alpha(y-x))| \\ c_B &\sim \int_0^1 dx \int_{x+\epsilon}^1 dy (1-\rho(x))(1-\rho(y)) \ln |\sin(\pi(2-\alpha)(y-x))|\end{aligned}\quad (6.30)$$

where ϵ is a small number that cuts off the divergence as $x \rightarrow y$, reflecting the fact that $j \neq k$ in the original discrete expression (6.29). The expressions c_A and c_B are those for the energy of a system of density ρ with an interparticle potential given by $\ln(\sin)$.

In this analogy, the potential $V(z)$ for a particle at x is the functional derivative with respect to $\rho(z)$ of $c_A + c_B$. However F has a fixed number of particles, so our extremization condition must be that $c_A + c_B$ is invariant under moving a particle from z_1 to z_2

$$\begin{aligned}0 &= \left(\frac{\delta}{\delta\rho(z_1)} - \frac{\delta}{\delta\rho(z_2)} \right) (c_A + c_B) = V(z_1) - V(z_2) \\ V(z) &= \int_0^1 dx (\rho(x) \ln |\sin(\pi\alpha(x-z))| - (1-\rho(x)) \ln |\sin(\pi(2-\alpha)(x-z))|).\end{aligned}\quad (6.31)$$

Therefore one finds that extremization requires that the potential must be z -independent, or in other words that the force $F(z)$ vanishes

$$\begin{aligned}0 &= F(z) = \frac{\partial}{\partial z} V(z) \\ &= -\pi \int_0^1 dx [\rho(x) (\alpha \cot(\pi\alpha(x-z)) + (2-\alpha) \cot(\pi(2-\alpha)(x-z))) \\ &\quad + (\alpha-2) \cot(\pi(2-\alpha)(x-z))].\end{aligned}\quad (6.32)$$

We solve this equation for ρ by discretizing, and reducing it to a linear matrix equation. The equation is solved by inverting the operator which multiplies ρ and multiplying it by the term on the last line. The dimension of this matrix is independent of N , and so N has disappeared from the problem. We discretized into 25 parts, so that $\rho(x)$ is defined by at 25 positions. Then we solved the equation to obtain our preferred value of $\rho(x)$, displayed in figure 5. At various values of N , we then created a subset $F \in [1, N]$ with local density ρ .

Figure 5. The optimal values of ρ found by solving (6.32) discretized into 25 parts J . The red, green, blue, black and brown curves correspond to α equal to 0.2, 0.4, 0.6, 0.8 and 1 respectively. Note that when $\alpha = 1$ the optimal density is constant as expected.

Figure 6. Comparison of the estimator in subsection 6.3 calculated with discretization into 20 (green) and 25 (red) parts at $N = 100$ (top left), $N = 250$ (top right), $N = 500$ (bottom left) and $N = 1000$ (bottom right). The exact matrix elements are shown in black.

The results are reported in red in figure 4. As a result of this clumsy optimization procedure, our results are quite noisy. In general at high α it systematically underestimates ψ_0 . However at low values of α it is closer to ψ_0 than $\psi_{\downarrow \dots \downarrow}$.

Needless to say, the choice of 25 parts is arbitrary and introduces additional error. To estimate this error, we compare the case of 20 and 25 parts in figure 6. The two curves do not agree. However no systematic offset is observed and so we believe that error introduced by our discretization is not responsible for our systematically high estimates of α .

7 Conclusions

State of the art computations of coordinate matrix elements in integrable models [12, 13] are only able to compute matrix elements with respect to coordinate states that are in some sense homogenous or linear, such as Néel states. Our goal is to generalize these computations to states which are piecewise linear. The strategy that we adopt is the decomposition of the lattice sites or particle positions into piecewise linear bins, which induces a decomposition of sum in the Coordinate Bethe Ansatz.

More precisely, the CBA computes matrix elements as a sum of $N!$ terms, corresponding to maps

$$P : [1, N] \longrightarrow [1, N] \quad (7.1)$$

or equivalently to elements of the symmetric group S_N . Summing over all $N!$ terms in general is difficult. But we found that if the domain $[1, N]$ is divided into bins $\mathcal{S}_i$ and one fixes the images $P(\mathcal{S}_i)$, the corresponding subset of terms can be summed exactly. Each such subset is a coset $H \subset S_N$ and S_N is the disjoint union of these cosets. We provided this sum in eq. (5.13). Thus the task of summing over all $N!$ group elements is reduced to the task of summing over the cosets, of which there are only of order 2^N .

More critically, in the case treated in section 5 in which the density was constant, the sums over all cosets had the same phase. Thus one needed only to sum over the dominant cosets to obtain a good approximation. More generally, in section 6 the sums over cosets had the same phase up to a sign. Our approximations to the matrix elements were generally too high because we did not consider the fluctuation in signs.

However for densities near the mean density, which are those that will contribute to wave functions in the large N limit, the fluctuations are caused only by the terms near the boundaries of the image $[1, N]$. We thus feel that it should be possible to develop an approximation scheme, along the lines of that developed in section 5 for transitions between $\uparrow$ and $\downarrow$, which allows us to estimate the correction caused by these sign flips. This will be done in future work.

Another critical generalization will be to other Hamiltonian eigenstates. Our motivation, understanding mass gaps in quantum field theory, suggests that we only need to investigate low lying states. For these, only a small number of momenta are changed and so we expect that we only need to modify our approximation scheme to keep track of which bin contains these excitations or, more simply, one may separately sum over P^{-1} of the excited momenta. The more challenging step comes next, we will try to generalize these results to the Lieb-Liniger model. As the Tonks-Girardeau model is the strong coupling limit of the Lieb-Liniger model, we will be able to test our results. If this can be achieved, we expect that spin chains such as the Heisenberg XXX model should present no new conceptual difficulties, at least at $s = 1/2$.

A The decomposition of S_4

In this appendix we will describe the two bin decomposition of S_4 . Now $N = 4$ and (4.14) gives the two bins

$$\mathcal{S}_A = \{1, 2\}, \quad \mathcal{S}_B = \{3, 4\}. \quad (\text{A.1})$$

These each contain just two elements. There are six choices of two element subsets $F \subset [1, 4]$. Recall that each permutation $g \in S_4$ corresponds to a function $P : [1, 4] \rightarrow [1, 4]$. $H_F \subset S_4$ is the subgroup of elements $g \in H_F$ such that the corresponding function P satisfies

$$P(\mathcal{S}_A) = F. \quad (\text{A.2})$$

Let us begin with $F = \{1, 2\}$. In the notation of eqs. (5.17) and (5.38) this corresponds to $\mathcal{F} = 20$. The condition (A.2) now becomes

$$P(\{1, 2\}) = \{1, 2\} \quad (\text{A.3})$$

where the sets are not ordered, so that $P(1)$ may be 1 or 2.

This is satisfied by four elements of $g \in S_4$

$$g_1 = e, \quad g_2 = (12), \quad g_3 = (34), \quad g_4 = (12)(34) \quad (\text{A.4})$$

where we use cyclic notation and e corresponds to the identity element. These correspond to the maps

$$\begin{aligned} P_1\{1, 2, 3, 4\} &= \{1, 2, 3, 4\}, & P_2\{1, 2, 3, 4\} &= \{2, 1, 3, 4\} \\ P_3\{1, 2, 3, 4\} &= \{1, 2, 4, 3\}, & P_4\{1, 2, 3, 4\} &= \{2, 1, 4, 3\}. \end{aligned} \quad (\text{A.5})$$

Here these sets are ordered, so for example $P_1(1) = 1$. Note that only P_1 is monotonically increasing, which identifies g_F as

$$g_F = g_{\{1,2\}} = g_1 = e. \quad (\text{A.6})$$

The set of all four g_i is a right coset of g_F

$$H_F = H_{\{1,2\}} = \{g_1, g_2, g_3, g_4\} = g_{\{1,2\}} (\mathbb{Z}_2 \times \mathbb{Z}_2), \quad \mathbb{Z}_2 \times \mathbb{Z}_2 = \{e, (12), (34), (12)(34)\}. \quad (\text{A.7})$$

Let us next consider $F = \{1, 3\}$, corresponding to $\mathcal{F} = \uparrow\uparrow$ in the notation of eqs. (5.17) and (5.38). Now

$$P(\{1, 2\}) = \{1, 3\} \quad (\text{A.8})$$

whose solutions are

$$g_5 = (23), \quad g_6 = (132), \quad g_7 = (234), \quad g_8 = (1342) \quad (\text{A.9})$$

corresponding to

$$\begin{aligned} P_5\{1, 2, 3, 4\} &= \{1, 3, 2, 4\}, & P_6\{1, 2, 3, 4\} &= \{3, 1, 2, 4\} \\ P_7\{1, 2, 3, 4\} &= \{1, 3, 4, 2\}, & P_8\{1, 2, 3, 4\} &= \{3, 1, 4, 2\}. \end{aligned} \quad (\text{A.10})$$

F	$\mathcal{F}$	g_F	H_F
$\{1, 2\}$	20	e	$\{e, (12), (34), (12)(34)\}$
$\{1, 3\}$	$\uparrow\uparrow$	(23)	$\{(23), (132), (234), (1342)\}$
$\{1, 4\}$	$\uparrow\downarrow$	(243)	$\{(243), (1432), (24), (142)\}$
$\{2, 3\}$	$\downarrow\uparrow$	(123)	$\{(123), (13), (1234), (134)\}$
$\{2, 4\}$	$\downarrow\downarrow$	(1243)	$\{(1243), (143), (124), (14)\}$
$\{3, 4\}$	02	$(13)(24)$	$\{(13)(24), (1423), (1324), (14)(23)\}$

Table 3. Decomposition of S_4 into two bins.

None of these is monotonically increasing. However, the definition of g_F is that it is the group element which is monotonically increasing when restricted to both $\mathcal{S}_A$ and to $\mathcal{S}_B$. This identifies

$$g_F = g_{\{1,3\}} = g_5 = (23) \quad (\text{A.11})$$

as

$$P_5\mathcal{S}_A = \{1, 3\}, \quad \text{and } P_5\mathcal{S}_B = \{2, 4\} \quad (\text{A.12})$$

are both monotonically increasing. Again the set of all four g_i is a right coset of g_F

$$H_F = H_{\{1,3\}} = \{g_5, g_6, g_7, g_8\} = g_{\{1,3\}} (\mathbb{Z}_2 \times \mathbb{Z}_2), \quad \mathbb{Z}_2 \times \mathbb{Z}_2 = \{e, (12), (34), (12)(34)\}. \quad (\text{A.13})$$

Note that the subgroup $\mathbb{Z}_2 \times \mathbb{Z}_2$ is independent of F .

The complete decomposition of S_4 is shown in table 3. Note that each of the six values of F corresponds to a coset consisting of four of the 24 elements of S_4 .

Acknowledgments

JE is supported by the CAS Key Research Program of Frontier Sciences grant QYZDY-SSW-SLH006 and the NSFC MianShang grants 11875296 and 11675223. JE also thanks the Recruitment Program of High-end Foreign Experts for support.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] H. Bethe, *On the theory of metals. 1. Eigenvalues and eigenfunctions for the linear atomic chain*, *Z. Phys.* **71** (1931) 205 [[INSPIRE](#)].
- [2] L. Tonks, *The Complete Equation of State of One, Two and Three-Dimensional Gases of Hard Elastic Spheres*, *Phys. Rev.* **50** (1936) 955.
- [3] M. Girardeau, *Relationship between Systems of Impenetrable Bosons and Fermions in One Dimension*, *J. Math. Phys.* **1** (1960) 516.

- [4] E.H. Lieb and W. Liniger, *Exact analysis of an interacting Bose gas. 1. The General solution and the ground state*, *Phys. Rev.* **130** (1963) 1605 [[INSPIRE](#)].
- [5] D.J. Gross, *Meron Configurations in the Two-Dimensional $O(3)$ σ -model*, *Nucl. Phys. B* **132** (1978) 439 [[INSPIRE](#)].
- [6] V.A. Fateev, I.V. Frolov and A.S. Shvarts, *Quantum Fluctuations of Instantons in the Nonlinear σ -model*, *Nucl. Phys. B* **154** (1979) 1 [[INSPIRE](#)].
- [7] I. Affleck, *Mass Generation by Merons in Quantum Spin Chains and the $O(3)$ σ Model*, *Phys. Rev. Lett.* **56** (1986) 408 [[INSPIRE](#)].
- [8] M. Brockmann and J.-M. Stéphan, *Universal terms in the overlap of the ground state of the spin-1/2 XXZ chain with the Néel state*, *J. Phys. A* **50** (2017) 354001 [[arXiv:1705.08505](#)] [[INSPIRE](#)].
- [9] M. Cohen, *The Energy Spectrum of Excitations in Liquid Helium*, Ph.D. Thesis, Caltech (1956) [<https://resolver.caltech.edu/CaltechETD:etd-03192004-153651>].
- [10] L.D. Faddeev and L.A. Takhtajan, *Spectrum and scattering of excitations in the one-dimensional isotropic Heisenberg model*, *J. Sov. Math.* **24** (1984) 241 [[INSPIRE](#)].
- [11] L.D. Faddeev, *How algebraic Bethe ansatz works for integrable model*, in *Relativistic gravitation and gravitational radiation. Proceedings, School of Physics, Les Houches, France, 26 September–6 October 1995*, pp. 149–219 (1996) [[hep-th/9605187](#)] [[INSPIRE](#)].
- [12] J. De Nardis, B. Wouters, M. Brockmann and J.-S. Caux, *Solution for an interaction quench in the Lieb-Liniger Bose gas*, *Phys. Rev. A* **89** (2014) 033601 [[arXiv:1308.4310](#)].
- [13] M. Brockmann, J. De Nardis, B. Wouters and J.-S. Caux, *A Gaudin-like determinant for overlaps of Néel and XXZ Bethe states*, *J. Phys. A* **47** (2014) 145003 [[arXiv:1401.2877](#)].