# 5 Coordinate Bethe Ansatz

Given a 2D model of equilibrium statistical physics, one would like to compute the partition function and certain correlation functions. This is rarely possible on a finite lattice—with dimer coverings being a notable exception—but in many cases exact results can be obtained in the thermodynamical limit. The question arises under which conditions such computations are possible.

Via the transfer matrix formulation, a 2D model can be recast as the time evolution of a 1D quantum spin chain. A first step in the analysis of a given problem is to identify the particles being described by the spin chain as well as their exact dynamics. The free energy and the correlation functions are related in the usual way to the ground state and the excitations of the spin chain Hamiltonian. The goal is then to diagonalise exactly this Hamiltonian, i.e., to identify its eigenvectors and eigenvalues.

If there were only one particle in a spin chain with periodic boundary conditions, its eigenvectors would be plane waves $\mathrm{e}^{ikx}$ of a certain momentum $k$. When more than one particle is present, a product of plane waves would be an appropriate wave function only if the particles moved independently. This is of course not the case in any non-trivial model. But a natural idea, pioneered by Bethe [Be31] in his 1931 paper on the Heisenberg model (also known as the XXX spin chain), is to try an Ansatz of coupled plane waves.

In this chapter we shall illustrate this approach on a more general model, known as the six-vertex model in its 2D incarnation, or equivalently as the XXZ spin chain. This model has important relations to the Potts model and to the Temperley-Lieb lattice algebra, as will be discussed in later chapters.

The six-vertex model was solved by Lieb [Li67]. It is a special case of the eight-vertex model which was later solved by Baxter [Ba72].

The Bethe Ansatz technique comes in several variants. In this chapter we focus on the so-called coordinate Bethe Ansatz, following roughly chapter 8 of Baxter's book [Ba82a]. The procedure is here to construct the eigenvectors explicitly for $n$-particle states, by identifying the relations—known as the Bethe Ansatz equations (BAE)—under which the so-called unwanted terms cancel out. Studying in detail the cases $n = 1$ and $n = 2$ usually gives crucial insight into the general form of these relations, and one proceeds to the general case by reasonable guesswork (which for the simplest models can be justified by detailed arguments).

In later chapters we shall rederive those results in an algebraic framework, culminating in the so-called algebraic Bethe Ansatz. It will gradually emerge

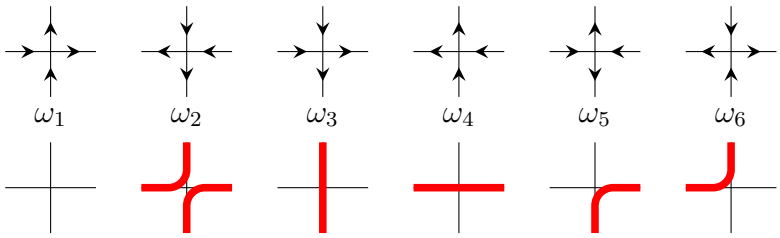



Figure 15: The allowed arrow arrangements (top) around a vertex that define the six-vertex model, with the corresponding particle trajectories (bottom).

that the exact solvability of a model hinges on the possibility to factorise any multi-particle scattering as a product of two-particle scatterings. This solvability condition is encoded in the celebrated Yang-Baxter equation.

## 5.1 Six-vertex model

The six-vertex model is defined by placing arrows on the edges of a square lattice, in such a way that every vertex is adjacent on two incoming and two outgoing arrows. The six possible configurations around a vertex are shown in the first line of Fig. 15 along with their respective Boltzmann weights $\omega_1, \ldots, \omega_6$. The corresponding energies are denoted $\epsilon_i$, and we have $\omega_i = \exp(-\frac{\epsilon_i}{k_B T})$. If there are $n_i$ vertices of type $i$ on the given lattice, the goal is to compute the asymptotic behaviour in the limit of a large lattice of the partition function

$$Z = \sum_{\text{arrows}} \prod_{i=1}^{6} (\omega_i)^{n_i} \,. \tag{5.1}$$

In the transfer matrix formulation, we impose for the moment periodic boundary conditions along the horizontal lattice direction. The row-to-row transfer matrix then conserves the net arrow flux in the time direction. To fully exploit this conservation law we move to an equivalent particle picture. Recall that in the $R$-matrix factorisation of the transfer matrix, time flows in the North-Eastern direction. We therefore define that an edge is occupied by a particle if and only if it sustains a right-pointing or an up-pointing arrow. This is shown in the second line of Fig. 15.

The resulting world-lines of particles are conserved, and moreover they have a very simple dynamics. Following them from the bottom of the system

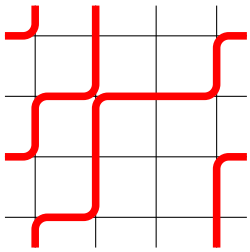


Figure 16: A possible configuration of world-lines on a $4 \times 4$ lattice.

to the top, they can only move up or to the right. This is illustrated in Fig. 16.

The six-vertex model is solvable when the weights are chosen to be invariant under a global reversal of arrows. The weights are traditionally denoted

$$\omega_1 = \omega_2 = a\,, \quad \omega_3 = \omega_4 = b\,, \quad \omega_5 = \omega_6 = c\,. \tag{5.2}$$

Note however that since the arrow arrangement 5 (resp. 6) acts as a sink (resp. source) of horizontal arrows, we must have the same number of each in every line. In particular, $n_5 = n_6$. For any $k \neq 0$ we can therefore take instead $\omega_5 = kc$ and $\omega_6 = k^{-1}c$ without changing $Z$. This "gauge symmetry" will turn out useful in later chapters.

## 5.2 Transfer matrix

The row-to-row transfer matrix $T$ conserves the number $n$ of world-lines when going from one row to the next. The positions $x_i$ of the lines specifies a state $|x_1, x_2, \ldots, x_n\rangle$, where we have assumed $x_1 < x_2 < \cdots < x_n$. The transfer matrix element

$$\langle y_1, y_2, \ldots, y_n | T | x_1, x_2, \ldots, x_n \rangle$$

equals the product of Boltzmann weights along the row, provided that the state of the upper row $\langle y_1, y_2, \ldots, y_n|$ is compatible with the state of the lower row $|x_1, x_2, \ldots, x_n\rangle$. If the two states are not compatible, the matrix element is defined to be zero.

The two states are compatible if and only if the positions in the upper row $y_i$ *interlace* those in the lower row $x_i$. The precise meaning of interlacing is the following:


Figure 17: The two ways in which $\langle y_1, y_2, \ldots, y_n|$ can interlace $|x_1, x_2, \ldots, x_n\rangle$, here shown for $n = 2$.

- If no line goes through the periodic lattice direction we have $x_i \leq y_i \leq x_{i+1}$ for $i = 1, 2, \ldots, n-1$, and $x_n \leq y_n \leq N$.
- If a line goes through the periodic direction, we have $1 \leq y_1 \leq x_1$, and $x_i \leq y_{i+1} \leq x_{i+1}$ for $i = 1, 2, \ldots, n-1$.

By self-avoidance, at most one line can go through the periodic direction. The two possible interlacings are shown for $n = 2$ in Fig. 17.

We now wish to construct $n$-particle states

$$|\Psi_n\rangle = \sum_{1 \leq x_1 < \cdots < x_n \leq N} g(x_1, \ldots, x_n)|x_1, \ldots, x_n\rangle \,, \tag{5.3}$$

which are eigenvectors of $T$:

$$T|\Psi_n\rangle = \Lambda|\Psi_n\rangle \,. \tag{5.4}$$

To this end we try an Ansatz of the form

$$g(x_1, \ldots, x_n) = \sum_{p \in \mathfrak{S}_n} A_p z_{p(1)}^{x_1} z_{p(2)}^{x_2} \cdots z_{p(n)}^{x_n} \,, \tag{5.5}$$

where the sum runs over all permutations $p \in \mathfrak{S}_n$ of the particle labels $\{1, 2, \ldots, n\}$. The complex numbers $z_j$ are related to the so-called quasi-momenta $k_j$ through the relation $z_j = \exp(ik_j)$. For the moment this Ansatz can be considered loosely as "coupled plane waves"; we shall come back to its physical interpretation in due course.

### 5.2.1 Sector with $n = 0$ particles

When $n = 0$, the unique state is completely empty. The horizontal row of edges if either empty (i.e., completely filled with $\omega_1$ vertices) or filled by

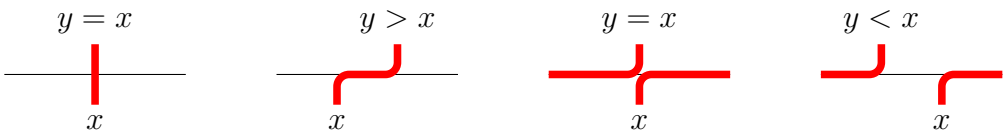



Figure 18: The four possible transitions between row states in the $n = 1$ particle sector.

a horizontal line (i.e., completely filled with $\omega_4$ vertices). Thus $T$ is one-dimensional and takes the value

$$\Lambda = a^N + b^N \,. \tag{5.6}$$

### 5.2.2 Sector with $n = 1$ particle

When $n = 1$, $\dim T = N$. The eigenvectors (5.3) read $|\Psi\rangle = \sum_x g(x)|x\rangle$, and the Ansatz (5.5) is $g(x) = z^x$. The particle at position $x$ can undergo four different processes as shown in Fig. 18. The eigenvalue equation (5.4) projected on a basis state of the eigenvector (5.3) then becomes[13]

$$\begin{aligned} \Lambda z^x &= a^{N-1} b z^x + \sum_{y=x+1}^{N} a^{N-(y-x+1)} b^{y-x-1} c^2 z^y \\ &+ a b^{N-1} z^x + \sum_{y=1}^{x-1} a^{x-y-1} b^{N-(x-y+1)} c^2 z^y \,. \end{aligned} \tag{5.7}$$

Recalling now the geometric series

$$\sum_{n=N_1}^{N_2} \omega^n = \frac{\omega^{N_1} - \omega^{N_2+1}}{1-\omega}$$

the first sum in (5.7) reads

$$a^N \left(\frac{a}{b}\right)^{x+1} \left(\frac{c}{a}\right)^2 \frac{\left(\frac{bz}{a}\right)^{x+1} - \left(\frac{bz}{a}\right)^{N+1}}{1 - \frac{bz}{a}} \,, \tag{5.8}$$

[13] To be precise, (5.7) contains the transition probabilities that $x$ becomes $y$, hence it is the transcription of $\langle \Psi_n | T | x \rangle = \Lambda \langle \Psi_n | x \rangle$.

while the second sum is

$$b^N \left(\frac{a}{b}\right)^{x+1} \left(\frac{c}{a}\right)^2 \frac{\left(\frac{bz}{a}\right)^1 - \left(\frac{bz}{a}\right)^x}{1-\frac{bz}{a}} \,. \tag{5.9}$$

Collecting everything, we first have the "wanted terms" multiplying $a^N z^x$. They come from the first term in (5.7) and the first term in (5.8), and their coefficient is

$$\frac{b}{a} + \frac{c^2 z}{a(a-bz)} = \frac{ab + (c^2 - b^2)z}{a(a-bz)} \equiv L(z) \,. \tag{5.10}$$

Another class of wanted terms, multiplying $b^N z^x$, comes from the third term in (5.7) and the second term in (5.9). The coefficient of these terms is

$$\frac{a}{b} - \frac{c^2}{b(a-bz)} = \frac{a^2 - c^2 - abz}{b(a-bz)} \equiv M(z) \,. \tag{5.11}$$

The remaining terms, namely the second term in (5.8) and the first term in (5.9), are "unwanted boundary terms" that read

$$\frac{a^{x-1} b^{N-x} c^2 z}{a-bz} (1 - z^N) \,. \tag{5.12}$$

The unwanted terms cancel out provided we impose the following condition on the allowed quasi-momenta:

$$z^N = 1 \,. \tag{5.13}$$

Note that this has precisely $N$ solutions, and we have thus determined $N$ eigenvectors for the $N$-dimensional matrix $T$. Its eigenvalue is given by the wanted terms and reads simply

$$\Lambda = a^N L(z) + b^N M(z) \,. \tag{5.14}$$

### 5.2.3 Sector with $n = 2$ particles

When $n = 2$, $\dim T = \binom{N}{2} = \frac{N(N-1)}{2}$. The two possibilities that $y_1, y_2$ can interlace $x_1, x_2$ are shown in Fig. 17. Let us define a function $E(x, y)$ that contains the weight of a world-line entering at $x$ and exiting at $y \geq x$

$$E(x,y) = \begin{cases} \frac{b}{c} & \text{if } y = x \\ c b^{y-x-1} & \text{if } y > x \end{cases} \tag{5.15}$$

and another function $D(y,x)$ giving the weight of the empty segment between two world-lines

$$D(y,x) = \begin{cases} \frac{a}{c} & \text{if } x = y \\ ca^{x-y-1} & \text{if } x > y \end{cases} . \tag{5.16}$$

The eigenvalue equation (5.4) projected on one basis state $|x_1, x_2\rangle$ of the eigenvector (5.3) then reads

$$\begin{aligned} \Lambda g(x_1,x_2) &= \sum_{y_1=x_1}^{x_2} \sum_{y_2=x_2}^{N} a^{x_1-1} E(x_1,y_1) D(y_1,x_2) E(x_2,y_2) c a^{N-y_2} g(y_1,y_2) \\ &+ \sum_{y_1=1}^{x_1} \sum_{y_2=x_1}^{x_2} b^{y_1-1} D(y_1,x_1) E(x_1,y_2) D(y_2,x_2) c b^{N-x_2} g(y_1,y_2) , \end{aligned}$$

where the two terms correspond to the situations shown in Fig. 17. Note that the special cases when one of the $x$ coincides with one of the $y$ are already provided for in the definitions of $E(x,y)$ and $D(y,x)$. However, the double sums must be constrained to exclude terms with $y_1 = y_2$. This is best done by first computing the sums without the constraint, then subtracting off the disallowed contribution $y_1 = x_2 = y_2$ to the first double sum, and $y_1 = x_1 = y_2$ to the second.

Despite of (5.5) we first insert the simper Ansatz

$$g(x_1,x_2) = A_{12} z_1^{x_1} z_2^{x_2} .$$

It is convenient to introduce the short-hand notations

$$L_j \equiv L(z_j) , \quad M_j \equiv M(z_j) , \quad \rho_j \equiv \rho(z_j) = \frac{c^2 z_j}{a(a - b z_j)}$$

and to define the function

$$R_j(x_1,x_2) = L_j a^{x_2-x_1} z_j^{x_1} + M_j b^{x_2-x_1} z_j^{x_2} .$$

As before the terms coming from the constrained double summations are of several types:

**Wanted terms**

These wanted terms are those proportional to $z_1^{x_1} z_2^{x_2}$:

$$A_{12} \left( a^N L_1 L_2 + b^N M_1 M_2 \right) z_1^{x_1} z_2^{x_2}$$

and they determine the eigenvalue

$$\Lambda = a^N L_1 L_2 + b^N M_1 M_2 \,. \tag{5.17}$$

**Unwanted internal terms**

There are of the form $(z_1 z_2)^{x_2}$ or $(z_1 z_2)^{x_1}$. One can verify that both such terms are proportional to

$$M_1 L_2 - 1 = -\frac{c^2 s_{12}}{(a - b z_1)(a - b z_2)} \,, \tag{5.18}$$

where

$$s_{12} = 1 - 2\Delta z_2 + z_1 z_2 \,, \tag{5.19}$$

$$\Delta = \frac{a^2 + b^2 - c^2}{2ab} \,. \tag{5.20}$$

The quantities $s_{12}$ (scattering phase) and $\Delta$ (anisotropy parameter) play a very important role in the solution of the six-vertex model, and in the physical interpretation of the scattering theory described by the Bethe Ansatz. We shall come back to this later.

**Unwanted boundary terms**

These come from the $y_2 = N$ or the $y_1 = 1$ summation limits and their sum is

$$A_{12} a^{x_1} b^{N - x_2} \left( R_2(x_1, x_2) \rho_1 - R_1(x_1, x_2) \rho_2 z_2^N \right) \,. \tag{5.21}$$

**Elimination of the unwanted terms**

The justification of the complete Ansatz (5.5) is precisely that it permits us to eliminate the unwanted terms. We therefore set

$$g(x_1, x_2) = A_{12} z_1^{x_1} z_2^{x_2} + A_{21} z_2^{x_1} z_1^{x_2} \,.$$

The unwanted internal terms cancel under the condition

$$s_{12}A_{12} + s_{21}A_{21} = 0\,. \tag{5.22}$$

The sum of the unwanted boundary terms is

$$a^{x_1}b^{N-x_2}\left\{\rho_2 R_1(x_1,x_2)(A_{21} - z_2^N A_{12}) + \rho_1 R_2(x_1,x_2)(A_{12} - z_1^N A_{21})\right\}$$

and this will vanish under the conditions

$$\begin{aligned} z_1^N &= \frac{A_{12}}{A_{21}} = -\frac{s_{21}}{s_{12}}\,, \\ z_2^N &= \frac{A_{21}}{A_{12}} = -\frac{s_{12}}{s_{21}}\,. \end{aligned} \tag{5.23}$$

## 5.3 Bethe Ansatz equations

The structure of the solution for the case of general $n$ is very much visible in the above detailed treatment for $n = 2$. By generalising the argument (or proceding by educated guesswork) it emerges that the eigenvalue is

$$\Lambda = a^N L_1 L_2 \cdots L_n + b^N M_1 M_2 \cdots M_n\,. \tag{5.24}$$

The condition for the vanishing of the unwanted internal terms becomes

$$s_{p_j,p_{j+1}} A_{p_1,\ldots,p_j,p_{j+1},\ldots,p_n} + s_{p_{j+1},p_j} A_{p_1,\ldots,p_{j+1},p_j,\ldots,p_n} = 0 \tag{5.25}$$

for each $j = 1, 2, \ldots, n-1$ and all permutations $p \in \mathfrak{S}_n$. Finally, the condition for the vanishing of the unwanted boundary terms reads

$$z_{p_1}^N = \frac{A_{p_1,p_2,\ldots,p_{n-1},p_n}}{A_{p_2,p_3,\ldots,p_n,p_1}} \tag{5.26}$$

for all $p \in \mathfrak{S}_n$.

There is a nice alternative way of deriving (5.26) using the consideration of translational invariance. Indeed the eigenstate must be unchanged upon taking any of the particles through the periodic boundary condition and back to its original position. In particular

$$g(x_1, x_2, \ldots, x_{n-1}, x_n) = g(x_2, x_3, \ldots, x_n, x_1 + N)\,. \tag{5.27}$$

Note that this respects our conventions that the arguments of $g$ must be written in increasing order. Using this, the form of the Ansatz (5.5) implies (5.26).

To see this, consider for simplicity the case of $n = 3$ particles. We have then

$$\begin{aligned} g(x_1, x_2, x_3) &= A_{123} z_1^{x_1} z_2^{x_2} z_3^{x_3} + A_{132} z_1^{x_1} z_3^{x_2} z_2^{x_3} + \dots , \\ g(x_2, x_3, x_1 + N) &= A_{231} z_2^{x_2} z_3^{x_3} z_1^{x_1+N} + A_{321} z_3^{x_2} z_2^{x_3} z_1^{x_1+N} + \dots . \end{aligned}$$

Since this must be valid for all $x_j$ we can identify terms

$$\begin{aligned} A_{231} z_1^N &= A_{123} , \\ A_{321} z_1^N &= A_{132} . \end{aligned} \tag{5.28}$$

This proves (5.26) for the case $n = 3$.

Obviously (5.25)–(5.26) provides many more equations that the $n$ unknown quasi-momenta $z_1, z_2, \dots, z_n$. Generalising (5.19) we define

$$s_{ij}(z_i, z_j) = 1 - 2\Delta z_j + z_i z_j . \tag{5.29}$$

One then easily verifies that (5.25) is solved by

$$A_{p_1,p_2,\dots,p_n} = \epsilon_p \prod_{1 \le i < j \le n} s_{p_j,p_i} , \tag{5.30}$$

where $\epsilon_p$ is the signature of the permutation $p \in \mathfrak{S}_n$. Inserting this into (5.26) gives

$$z_{p_1}^N = (-1)^{n-1} \prod_{l=2}^{n} \frac{s_{p_l,p_1}}{s_{p_1,p_l}}$$

for all $p \in \mathfrak{S}_n$. But since the right-hand side is symmetric in $p_2, p_3, \dots, p_n$ there are actually only $n$ distinct equations:

$$z_j^N = (-1)^{n-1} \prod_{\substack{l=1 \\ l \neq j}}^{n} \frac{s_{l,j}}{s_{j,l}} \quad \text{for } j = 1, 2, \dots, n . \tag{5.31}$$

These are the Bethe Ansatz equations (BAE) for the six-vertex model.

The progress obtained by now is considerable. Rather than diagonalising a transfer matrix of dimension $2^N$ we have to solve only a set of $n$ coupled (but non-linear) equations for each $n = 1, 2, \dots, N$.

### 5.3.1 Scattering phases and the Yang-Baxter equation

It is useful to define the modified scattering phases

$$\hat{S}_{ij}(z_i, z_j) = -\frac{s_{ij}}{s_{ji}} = -\frac{1 - 2\Delta z_j + z_i z_j}{1 - 2\Delta z_i + z_i z_j}\,. \tag{5.32}$$

The Bethe Ansatz equations can the be written in the suggestive form

$$z_j^N = \prod_{\substack{l=1 \\ l \neq j}}^{n} \hat{S}_{lj}(z_l, z_j) \quad \text{for } j = 1, 2, \ldots, n\,. \tag{5.33}$$

This can be interpreted physically as follows. When the particle $j$ is taken around the periodic direction and back to its original position, it picks up a scattering phase $\hat{S}_{lj}$ each time it crosses another particle $l$. These phases are also known as the $S$-matrix elements of the scattering theory.

We now derive some important physical properties of the $S$-matrix. Let us again focus on the case of $n = 3$ particles. Eliminating $z_1^N$ from (5.28) we obtain $\frac{A_{123}}{A_{231}} = \frac{A_{132}}{A_{321}}$. Doing the same for $z_2^N$ and $z_3^N$, and making some rearrangements, we arrive at

$$\frac{A_{213}}{A_{123}} = \frac{A_{321}}{A_{312}}\,, \quad \frac{A_{312}}{A_{132}} = \frac{A_{231}}{A_{213}}\,, \quad \frac{A_{321}}{A_{231}} = \frac{A_{132}}{A_{123}}\,. \tag{5.34}$$

This tells us that the interchange of two particles (e.g., 1 and 2 in the first relation) is independent of the position of the third particle (which on the left-hand side of the relations is to the right of the two particles being interchanges, and vice versa).

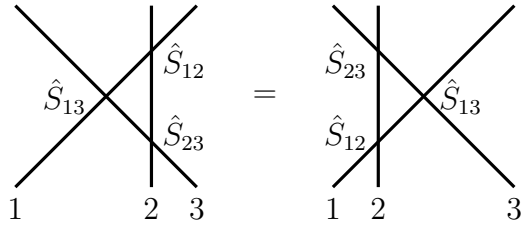
It thus emerges that the $S$-matrix possesses a locality property, according to which the scattering amplitude of $n$ quasi-particles factorises into a product of $\binom{n}{2}$ two-particle $S$-matrices. To make this more precise, consider the following relations which follow from (5.30):

$$A_{321} = \begin{cases} \hat{S}_{12} A_{312} = \hat{S}_{12}\hat{S}_{13} A_{132} = \hat{S}_{12}\hat{S}_{13}\hat{S}_{23} A_{123} \\ \hat{S}_{23} A_{231} = \hat{S}_{23}\hat{S}_{13} A_{213} = \hat{S}_{23}\hat{S}_{13}\hat{S}_{12} A_{123} \end{cases}$$

Eliminating $A_{123}$ yields the so-called Yang-Baxter relation

$$\hat{S}_{12}\hat{S}_{13}\hat{S}_{23} = \hat{S}_{23}\hat{S}_{13}\hat{S}_{12}\,, \tag{5.35}$$

which can be represented diagramatically in terms of the world-lines of the particles:




(5.36)

The graphical reading of this diagram is that any world-line can be moved across the intersection of two other world-lines.

Obviously the above argument is at most suggestive, since after all the factors in (5.35) are just scalars, and as such the identity is trivial. We shall however see later that the same relation holds for matrix-valued quantities (operators), such as $R$-matrices and monodromy ($\simeq$ transfer) matrices.

The Yang-Baxter equation (5.35) is at the heart of the algebraic approach to the Bethe Ansatz and we shall return to it extensively in later chapters.

## 5.4 Phase diagram

The six-vertex model has a non-trivial phase diagram, and it is hardly surprising that its thermodynamic limit depends on the parameter $\Delta$. In fact the Bethe Ansatz equations (5.31) show that the limit depends *only* on $\Delta$. Although the phase diagram can be derived in details, let us first discuss a few qualitative arguments.

If either $a$ or $b$ is large compared to the other weights, the system will freeze into a unique state in which all vertical arrows and all horizontal arrows point in the same direction. It turns out that this freezing occurs whenever $\Delta > 1$. The largest eigenvalue is that of the $n = 0$ particle sector, whence trivially $\Lambda_{\max} = a^N + b^N$. The free energy per vertex is therefore $f = \min(\epsilon_1, \epsilon_3)$.

If $c$ is very large compared to the other weights, the predominant configuration is the one where all vertices on the even (resp. odd) sublattice are of the type $\omega_5$ (resp. $\omega_6$). The system is however not frozen, meaning that it exhibits fluctuations around the predominant configuration.[14] One would

[14]We shall discuss the nature of these fluctuations more carefully when dealing with the Coulomb gas. Suffice it to say here that the least possible change of a configuration is to reverse a path of consistently oriented arrows. Any such path has infinite length in the

expect non-critical behaviour as long as $c$ remains reasonably large. This is indeed the case: when $\Delta < -1$ the system is in a non-critical phase.

The most interesting phase occurs for $-1 < \Delta < 1$. In that range the six-vertex model is critical, and it turns out that the critical exponents depend continuously on $\Delta$. This is an interesting counter-example to naive ideas of universality.

We shall concentrate most of the subsequent discussion on the critical case $-1 < \Delta < 1$, where the free energy can be expressed in terms of Fourier integrals. (The non-critical case $\Delta < -1$ can also be worked out in details and calls instead for the use of Fourier series.)

## 5.5 Thermodynamic limit for $\Delta < 1$

It is not known how to solve the Bethe Ansatz equations (5.31) for finite $n$ and $N$. This situation is quite common in the study of integrable systems. By contrast, we have seen that the partition function of dimer coverings can be exactly computed on a finite lattice—a highly unusual situation.

Nevertheless, it turns out that the six-vertex model is exactly solvable in the thermodynamic limit. By this we mean precisely that the free energy $f = -\frac{1}{\beta N} \log \Lambda_{\max}$, or equivalently the ground state energy in the spin chain, can be determined analytically for $N \to \infty$. The same is true for the low-lying excitations, but for the moment we concentrate on the ground state.

### 5.5.1 Location of the quasi-momenta

The BAE (5.31) possess many solutions for the quasi-momenta $z_j$. It is not a priori clear which one corresponds to the ground state. In what follows we shall admit the following fact:

- The solution of the BAE (5.31) that maximises the eigenvalue $\Lambda$ is such that $z_1, z_2, \ldots, z_n$ are distinct, lie on the unit circle, are distributed symmetrically about unity, and are packed as closely as possible.

This can actually be proved quite rigorously, using some lengthy analysis [YY66]. An easier method, that usually works quite well for more general

frozen phase, whereas the $\Delta < -1$ has an exponential number of short paths (of length four).

integrable models, is to study numerically the solutions for low values of $N$—confronting the results with exact diagonalisations of the transfer matrix—until the pattern has become clear.

### 5.5.2 Transformation to a set of real equations

We introduce the momenta $k_j \in \mathbb{R}$ and the function $\Theta(p,q)$ so that

$$z_j = \exp(ik_j)\,, \tag{5.37}$$

$$\frac{s_{i,j}}{s_{j,i}} = \exp(-i\Theta(k_j,k_i))\,. \tag{5.38}$$

By (5.32) we have then

$$\mathrm{e}^{-i\Theta(p,q)} = \frac{1-2\Delta\mathrm{e}^{ip}+\mathrm{e}^{i(p+q)}}{1-2\Delta\mathrm{e}^{iq}+\mathrm{e}^{i(p+q)}}\,. \tag{5.39}$$

To see that $\Theta(p,q)$ is a real function, it suffices to notice that

$$\tan\left(\frac{1}{2}\Theta(p,q)\right) = -i\,\frac{1-\mathrm{e}^{-i\Theta(p,q)}}{1+\mathrm{e}^{-i\Theta(p,q)}} = \frac{\Delta\sin\left(\frac{p-q}{2}\right)}{\cos\left(\frac{p+q}{2}\right)-\Delta\cos\left(\frac{p-q}{2}\right)}\,,$$

where the right-hand side is manifestly real.

Including the term $l=j$ obviously leaves the right-hand side of (5.31) unchanged, so we can rewrite it as

$$\exp(iNk_j) = (-1)^{n-1}\prod_{l=1}^{n}\exp(-i\Theta(k_j,k_l))\,,$$

where now both sides of the equation are unimodular. Taking logarithms we have

$$Nk_j = 2\pi I_j - \sum_{l=1}^{n}\Theta(k_j,k_l)\,, \tag{5.40}$$

where $I_j$ ranges between $\pm\left(\frac{n-1}{2}\right)$, hence is an integer (resp. half an odd integer) if $n$ is odd (resp. even). Note that both sides of this equation are real.

The hypothesis that $k_1,k_2,\ldots,k_n$ be distinct, symetrically distributed about the origin, and packed as closely as possible implies that the ground state is obtained by choosing

$$I_j = j - \frac{1}{2}(n+1)\,, \quad \text{for } j=1,2,\ldots,n\,. \tag{5.41}$$

### 5.5.3 Continuum limit

The thermodynamic limit is obtained by sending $n, N \to \infty$, while keeping the ratio $n/N$ fixed and finite. This ratio describes the (fixed) ratio of up-pointing arrows in each row of the lattice. The distribution function $\rho(k)$ of Bethe roots is defined so that $N\rho(k)\,\mathrm{d}k$ is the number of $k_j$ lying between $k$ and $k + \mathrm{d}k$. By assumption $\rho(k)$ has support on a symmetric interval $[-Q, Q]$, where $Q$ will be determined later. Thus

$$\int_{-Q}^{Q} \rho(k)\,\mathrm{d}k = \frac{n}{N}\,. \tag{5.42}$$

For a given value $k_j$ of $k$, the quantity $I_j + \frac{1}{2}(n+1) = N\int_{-Q}^{k}\rho(k')\,\mathrm{d}k'$ is the number of momenta $k_l$ with $l < j$. Passing from sums to integrals in (5.40)—and denoting $k_j$ simply as $k$—then produces

$$Nk = -\pi(n+1) + 2\pi N\int_{-Q}^{k}\rho(k')\,\mathrm{d}k' - N\int_{-Q}^{Q}\Theta(k,k')\rho(k')\,\mathrm{d}k'\,.$$

Taking derivatives with respect to $k$, and dividing by $N$, then leads to a linear integral equation for $\rho(k)$

$$2\pi\rho(k) = 1 + \int_{-Q}^{Q}\frac{\partial\Theta(k,k')}{\partial k}\rho(k')\,\mathrm{d}k'\,. \tag{5.43}$$

The free energy is then given by (5.24) as

$$f = -\frac{1}{\beta}\max\left\{\log a + \frac{1}{N}\sum_{j=1}^{n}\log L(z_j)\,, \log b + \frac{1}{N}\sum_{j=1}^{n}\log M(z_j)\right\}\,.$$

In the thermodynamic limit this becomes

$$\begin{aligned} f \;=\; -\frac{1}{\beta}\max\Big\{&\log a + \int_{-Q}^{Q}[\log L(\mathrm{e}^{ik})]\rho(k)\,\mathrm{d}k\,,\\ &\log b + \int_{-Q}^{Q}[\log M(\mathrm{e}^{ik})]\rho(k)\,\mathrm{d}k\Big\}\,. \end{aligned} \tag{5.44}$$

## 5.6 Free energy for $-1 < \Delta < 1$

A natural strategy for solving the linear integral equation (5.43) would be to use Fourier transformation. This is however only possible if we can find a transformation to a difference kernel. Fortunately this is possible.

### 5.6.1 Difference kernel transformation

For $-1 < \Delta < 1$ we parameterise

$$\Delta = -\cos\mu\,, \quad \text{with } 0 < \mu < \pi \tag{5.45}$$

and we trade $k$ for a new variable $\alpha$ defined by

$$\mathrm{e}^{ik} = \frac{\mathrm{e}^{i\mu} - \mathrm{e}^{\alpha}}{\mathrm{e}^{i\mu+\alpha} - 1}\,. \tag{5.46}$$

Note that $\alpha \in \mathbb{R}$. Indeed, supposing this, it is easily seen that $|\mathrm{e}^{ik}|^2 = 1$ from the right-hand side of (5.46), as is consistent with the hypothesis that $k \in \mathbb{R}$.

Differentiating logarithmically—i.e., using $\frac{\mathrm{d}}{\mathrm{d}\alpha}\mathrm{e}^{ik} = i\mathrm{e}^{ik}\frac{\mathrm{d}k}{\mathrm{d}\alpha}$ to isolate $\frac{\mathrm{d}k}{\mathrm{d}\alpha}$—we find

$$\frac{\mathrm{d}k}{\mathrm{d}\alpha} = \frac{\sin\mu}{\cosh\alpha - \cos\mu}\,. \tag{5.47}$$

This proves in particular that $k(\alpha) \in \mathbb{R}$ is a monotonically increasing function (since $0 < \mu < \pi$), and by (5.46) it maps the interval $(-\infty, \infty)$ onto $(\mu - \pi, \pi - \mu)$. It follows directly from (5.46) that $k(-\alpha) = -k(\alpha)$, i.e., the function is odd.

Let us define $p = k(\alpha)$ and $q = k(\beta)$, so that

$$\mathrm{e}^{ip} = \frac{\mathrm{e}^{i\mu} - \mathrm{e}^{\alpha}}{\mathrm{e}^{i\mu+\alpha} - 1}\,, \qquad \mathrm{e}^{iq} = \frac{\mathrm{e}^{i\mu} - \mathrm{e}^{\beta}}{\mathrm{e}^{i\mu+\beta} - 1}\,.$$

Inserting this into (5.39) the scattering phase becomes

$$\mathrm{e}^{-i\Theta(p,q)} = \frac{\mathrm{e}^{\alpha-\beta} - \mathrm{e}^{2i\mu}}{\mathrm{e}^{\beta-\alpha} - \mathrm{e}^{2i\mu}}\,. \tag{5.48}$$

Crucially, this depends only on the difference $\alpha - \beta$ (and on the constant $\mu$).

We shall need the root density function $R(\alpha)$ transformed to the $\alpha$ variable (and renormalised by $\frac{1}{2\pi}$ for later convenience)

$$R(\alpha)\,\mathrm{d}\alpha = 2\pi\rho(k)\,\mathrm{d}k\,. \tag{5.49}$$

Plugging this into (5.43) leads to

$$R(\alpha) = \frac{\mathrm{d}k}{\mathrm{d}\alpha} + \frac{1}{2\pi}\int_{-Q_1}^{Q_1} \frac{\partial\Theta(\alpha,\beta)}{\partial\beta}\,R(\beta)\,\mathrm{d}\beta\,,$$

where we note that there is a new integration range $(-Q_1, Q_1)$ corresponding to the $\alpha$ variable. Computing the derivatives from (5.47) and (5.48) finally leads to

$$R(\alpha) = \frac{\sin\mu}{\cosh\alpha - \cos\mu} - \frac{1}{2\pi}\int_{-Q_1}^{Q_1} \frac{\sin(2\mu)}{\cosh(\alpha-\beta) - \cos(2\mu)}\, R(\beta)\,\mathrm{d}\beta\,, \quad (5.50)$$

and the normalisation condition (5.42) for the root density function now reads

$$\frac{1}{2\pi}\int_{-Q_1}^{Q_1} R(\alpha)\,\mathrm{d}\alpha = \frac{n}{N}\,. \quad (5.51)$$

### 5.6.2 Parameterisation

We have already parameterised $\Delta = -\cos\mu$ in (5.45). The Bethe Ansatz equations (5.31)—and hence the universality class of the six-vertex model—depend only on the initial vertex weights $a, b, c$ through $\Delta = \frac{a^2+b^2-c^2}{2ab}$. We must therefore choose a parameterisation of the two independent ratios $a : b : c$ that respects this latter constraint (we "uniformise the spectral curve"). This can be done in this case using trigonometric functions:

$$a : b : c = \sin\left(\frac{\mu - w}{2}\right) : \sin\left(\frac{\mu + w}{2}\right) : \sin\mu\,, \qquad -\mu < w < \mu \quad (5.52)$$

defining another parameter $w$.

The eigenvalues of the transfer matrix are determined by (5.24) through the functions $L(z)$ and $M(z)$ given by (5.10)–(5.11). Recalling that $z = \mathrm{e}^{ik}$ is parameterised by (5.46), the parametric form of these functions now becomes

$$L(\mathrm{e}^{ik}) = \frac{\mathrm{e}^{i(w+\mu)} - \mathrm{e}^{\alpha - i\mu}}{\mathrm{e}^{\alpha} - \mathrm{e}^{iw}}\,,$$
$$M(\mathrm{e}^{ik}) = \frac{\mathrm{e}^{i(w-\mu)} - \mathrm{e}^{\alpha + i\mu}}{\mathrm{e}^{\alpha} - \mathrm{e}^{iw}}\,. \quad (5.53)$$

### 5.6.3 Solution by Fourier integrals

The integral equation (5.50) that determines the root density function now has a difference kernel. One can therefore solve it by Fourier transformation,

provided that $Q_1 = \infty$. Let us suppose that this is so, and justify the assumption below. The Fourier transformed root density function then reads

$$\widetilde{R}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\alpha) \mathrm{e}^{ix\alpha} \, \mathrm{d}\alpha \, . \tag{5.54}$$

Let us define the function

$$\phi_\mu(\alpha) = \frac{\sin \mu}{\cosh \alpha - \cos \mu} \, , \tag{5.55}$$

which is often referred to as the *source term* of the Bethe Ansatz equations. The difference kernel equation (5.50) then decouples upon Fourier transformation, since the Fourier transform of a convolution is the product of Fourier transforms. Explicitly, multiplying both sides of (5.50) by $\frac{1}{2\pi}\mathrm{e}^{ix\alpha}$ and integrating over $\alpha$ leads to

$$\widetilde{R}(x) = \widetilde{\phi_\mu}(x) - \widetilde{\phi_{2\mu}}(x) \cdot \widetilde{R}(x) \, . \tag{5.56}$$

Exercise: Show that the Fourier transform of the source term is

$$\widetilde{\phi_\mu}(x) = \frac{\sinh\left((\pi - \mu)x\right)}{\sinh(\pi x)} \, . \tag{5.57}$$

Inserting this yields

$$\widetilde{R}(x) = \frac{\sinh((\pi - \mu)x)}{\sinh(\pi x)} - \frac{\sinh((\pi - 2\mu)x)}{\sinh(\pi x)} \widetilde{R}(x) \tag{5.58}$$

and we can then finally isolate

$$\widetilde{R}(x) = \frac{1}{2\cosh(\mu x)} \, . \tag{5.59}$$

The normalisation condition (5.51) is such that $\widetilde{R}(0) = \frac{n}{N}$, and evaluating (5.59) we arrive at

$$\frac{n}{N} = \frac{1}{2} \, . \tag{5.60}$$

This simple result justifies the assumption $Q_1 = \infty$ a posteriori. Indeed, the largest sector of the transfer matrix precisely corresponds to the case where

there are as many up-pointing as down-pointing arrows. By a simple entropic reasoning, this is also the ground state sector.[15]

From (5.53) one obtains

$$\begin{aligned} |L(\mathrm{e}^{ik})|^2 &= \frac{\cos(w+2\mu)-\cosh\alpha}{\cos w-\cosh\alpha}\,, \\ |M(\mathrm{e}^{ik})|^2 &= \frac{\cos(w-2\mu)-\cosh\alpha}{\cos w-\cosh\alpha}\,. \end{aligned}$$

This implies that $|L|>|M|$ for $w<0$, and $|L|<|M|$ for $w>0$.

Suppose in the sequel that $w<0$; a similar calculation for $w>0$ can be shown to lead to exactly the same end result. The free energy is then given by the first term in (5.44):

$$f=-\frac{1}{\beta}\left(\log a+\frac{1}{2\pi}\int_{-\infty}^{\infty}[\log|L(\mathrm{e}^{ik})|]R(\alpha)\,\mathrm{d}\alpha\,.\right) \tag{5.61}$$

Using (5.59) and the fact that parity is conserved by Fourier transformation, we see that $R(\alpha)$ is an even function. Under the integral we can therefore replace the other factor $\log|L(\mathrm{e}^{ik})|$ by its even part, which is also its real part. From (5.53) we get

$$\mathrm{Re}\,L(\mathrm{e}^{ik})=-\cos\mu+\frac{\sin\mu\,\sin w}{\cos w-\cosh\alpha}\,, \tag{5.62}$$

and the Fourier transform of $\log|L(\mathrm{e}^{ik})|$ becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{ix\alpha}\log|L(\mathrm{e}^{ik})|\,\mathrm{d}\alpha=\frac{\sinh((\mu+w)x)\sinh((\pi-\mu)x)}{x\sinh(\pi x)}\,. \tag{5.63}$$

Exercise: Detail this computation!

To compute (5.61) we can use that the Fourier transform of a product is the convolution of Fourier transforms. The end result follows by combining (5.59) and (5.63):

$$f=-\frac{1}{\beta}\left(\log a+\int_{-\infty}^{\infty}\frac{\sinh((\mu+w)x)\sinh((\pi-\mu)x)}{2x\cosh(\mu x)\sinh(\pi x)}\,\mathrm{d}x\right)\,. \tag{5.64}$$

As already stated, exactly the same result is found for $w>0$. We have therefore found, for any $w\in(-\mu,\mu)$, the free energy of the six-vertex model in the critical region $\Delta\in(-1,1)$.

[15] A variant argument is obtained by examining (5.24).

### 5.6.4 Ice model

The equal-weighted case $a = b = c = 1$ is of special interest, both combinatorially and historically (it was solved by Lieb [Li67] before the general case). Physically it can be interpreted as a two-dimensional model of ice: The arrows on the edges represent to which side the electron cloud is pushed by the hydrogen bonding, and the six-vertex constraint corresponds to local charge neutrality.

In this case $w = 0$ and $\mu = \frac{2\pi}{3}$ from (5.52). The integral (5.64) can then be performed explicitly by contour integration; this is true more generally whenever $\mu$ is a rational fraction of $\pi$.

We have

$$-\beta f = \int_{-\infty}^{\infty} \frac{\sinh\left(\frac{\pi x}{3}\right) \tanh\left(\frac{2\pi x}{3}\right)}{2x \sinh(\pi x)} \, \mathrm{d}x \,. \tag{5.65}$$

The integrand is regular for $x \to 0$, even, and decays like $\frac{1}{2x} \mathrm{e}^{-2\pi x/3}$ for $\mathrm{Re}\, x \gg 1$. We can therefore close the contour in the upper-half plane and use the Cauchy integration theorem.

The residues are all on the imaginary axis. They read for $p \in \mathbb{N}$:

$$\begin{aligned}
&\frac{-3i}{4(3p+\frac{3}{4})\pi} \quad \text{at} \quad x = (3p+\frac{3}{4})i \,, \\
&\frac{3i}{4(3p+1)\pi} \quad \text{at} \quad x = (3p+1)i \,, \\
&\frac{3i}{4(3p+2)\pi} \quad \text{at} \quad x = (3p+2)i \,, \\
&\frac{-3i}{4(3p+\frac{9}{4})\pi} \quad \text{at} \quad x = (3p+\frac{9}{4})i \,.
\end{aligned}$$

Thus

$$\begin{aligned}
-\beta f &= 2\pi i \sum_{\mathrm{Im}(x)>0} \mathrm{Res}\, f(x) \\
&= -\frac{3}{2} \sum_{p=0}^{\infty} \left( -\frac{1}{3p+\frac{3}{4}} + \frac{1}{3p+1} + \frac{1}{3p+2} - \frac{1}{3p+\frac{9}{4}} \right) \\
&= \frac{5}{2} \sum_{p=0}^{\infty} \frac{(1+2p)}{(1+3p)(2+3p)(1+4p)(3+4p)} \\
&= \frac{3}{2} \log\left(\frac{4}{3}\right) \,.
\end{aligned} \tag{5.66}$$

The effective number of configurations per vertex

$$Z^{1/MN} = \exp(-\beta f) = \left(\frac{4}{3}\right)^{\frac{3}{2}} \simeq 1.539\,600\cdots \tag{5.67}$$

is known as Lieb's constant [Li67].