ICTP – Eurasian Centre for Advanced Research

II. International Summer School on
Exact and Numerical Methods for
Low-Dimensional Quantum Structures
August 23 – 31, 2014

LECTURE NOTES
INTRODUCTION TO BETHE ANSATZ TECHNIQUES AND SOME OF THEIR APPLICATIONS TO MODELS OF QUANTUM MATTER

Hans-Peter Eckle

Remarks:

These lecture notes consist of chapters and sections of an evolving monograph. Some of the cross references in the text therefore refer to chapters, sections, or equations which are not included in these lecture notes. Some of the references to the literature at the end of the lecture notes are, for the same reason, not mentioned in the text. Please let me know of errors of any kind and any comments you may wish to make (hanspetereckle@googlemail.com).

14
Ice model

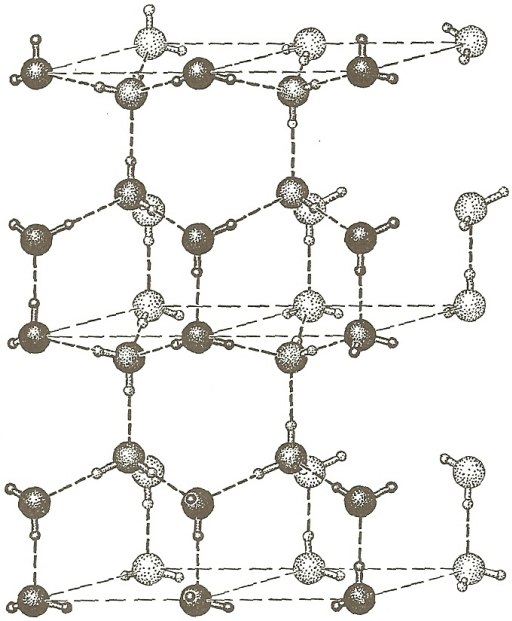

FIG. 12-6.—The arrangement of molecules in the ice crystal. The orientation of the water molecules, as represented in the drawing, is arbitrary; there is one proton along each oxygen-oxygen axis, closer to one or the other of the two oxygen atoms.

Fig. 14.1 The crystal structure of three–dimensional ice. Source: Linus Pauling: The nature of the chemical bond.

The crystal structure of ice as depicted in figure 14.1 shows a peculiarity. In the words of Linus Pauling (1960): "there is one proton along each oxygen–oxygen axis, closer to one or the other of the two oxygen atoms." This peculiarity, arising from a double well potential along the oxygen–oxygen axis, so that the hydrogen atom can sit in one or the other potential minimum, depends only on the configuration and is, hence, independent of temperature. It gives rise to a finite entropy of ice, even at zero temperature, the *residual entropy*.

We shall use this observation as a physical motivation to study a certain type of *two–dimensional* statistical mechanics models, the so-called *vertex* models. This class of models is

of physical interest in its own right. However, we shall study them for a particular reason: We shall study the vertex models as a convenient path to introduce one–dimensional *integrable* quantum mechanical models and the Bethe ansatz. Quite apart from the questions of integrability, there is a correspondence between statistical mechanical lattice models, which are really classical models, and quantum mechanical models. This correspondence relates the statistical mechanical models in $(d+1)$ dimensions to quantum mechanical models in d dimensions, see e.g. section 1 "Introduction" of Shenker (1982) for an elementary discussion. We shall make this correspondence quite explicit, starting with the six–vertex model, and ending up with the Heisenberg XXZ quantum spin chain, having established, along the way, integrability, in particular the heart of integrability, the Yang–Baxter relations.

This is the constructive program of the *algebraic* Bethe ansatz.

14.1 Physical Motivation for square lattice ice model

Careful experiments by Giauque and Stout (1936) on the heat capacity of ice down to low temperatures lead to the extrapolation that there will remain a residual entropy of ice crystals at zero temperature. Experimentally they obtained the value

$$\frac{S}{kN} \approx 0.41 \approx \ln(1.5)$$

for the entropy per oxygen atom.

Concurrent with these experiments, Pauling (1935) suggested an explanation for the residual entropy of ice, based on the observation discussed above (cf. figure 14.1).

Many different microscopic configurations of hydrogen atoms with respect to the oxygen atoms are possible (see figure 14.1) which lead to the same macroscopic state. Then already the simplest assumption leads to a finite value for the residual entropy of ice: Assume a lattice of N oxygen atoms as vertices with one proton per bond (i.e. oxygen–oxygen axis), hence $2N$ hydrogen atoms in total. The proton can be in one of two positions: near to or far from one of the two oxygen atoms, and, consequently, in the opposite position with respect to the other oxygen atom. Therefore we have, ignoring boundaries and, more importantly, possible restrictions on them, 2^{2N} possible configurations. The number of possible microscopic configurations is therefore

$$Z = 2^{2N} = 4^N, \tag{14.1}$$

which leads to the residual entropy

$$S = k \ln Z = kN \ln(4) \quad \Rightarrow \quad \frac{S}{kN} = \ln(4).$$

This value is much too large compared to the experimental result. A simple consideration, going back to Bernal and Fowler (1933) and applied to the problem of the residual entropy of ice by Pauling (Pauling, 1935; Pauling, 1960), improves this result considerably.

Since the oxygen atoms in ice are four–fold coordinated (cf. figure 14.2), each oxygen atom can be surrounded by either zero, one, two, three or four hydrogen atoms in the adjacent position. This accounts for a total of 16 configurations. Pauling argued that only six out of

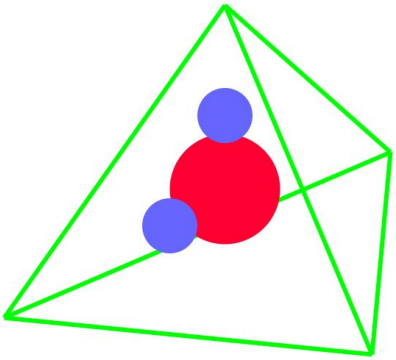

Fig. 14.2 Coordination of water in an ice crystal.

these 16 configurations are, in fact, allowed, namely the ones where there are two hydrogen atoms surrounding the oxygen atom. Therefore we get, instead of (14.1),

$$Z = 2^{2N} \left(\frac{6}{16}\right)^N = \left(\frac{3}{2}\right)^N,$$

and, thus,

$$S = k \ln Z = kN \ln\left(\frac{3}{2}\right) \quad \Rightarrow \quad \frac{S}{kN} = \ln\left(\frac{3}{2}\right).$$

This value is in good agreement with the experimental result of Giauque and Stout ((Giauque and Stout, 1936)).

Introducing the number W such that

$$S = kN \ln(W),$$

numerical calculations by Nagle (1966) confirmed the value of

$$W \approx 1.5.$$

Moreover, numerical calculations, also by Nagle (1966) for a caricature model of ice in two spatial dimensions also yield

$$W \approx 1.540 \pm 0.001.$$

The latter result prompted mathematical physicist Elliott Lieb to investigate whether an exact solution of the problem of two–dimensional square ice was within reach (Lieb, 1967*a*; Lieb, 1967*b*): "It seemed worthwhile to try and find the value of W exactly for the following reasons: (a) It will serve as a check of Nagle's calculations. (b) It is an interesting graph theoretic problem. (c) It is the first step toward the solution of much more interesting problems having phase transitions ..." (Lieb, 1967*b*).

Let's give a definition of the ice model before we return to Lieb's result for W.

Square ice, also called the *ice model*, is defined on a two–dimensional square lattice. Put N "oxygen" atoms at the vertices of the two–dimensional square lattice and "decorate" the

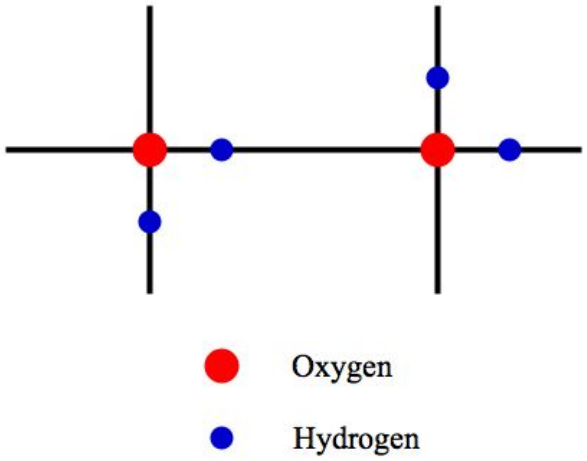

Fig. 14.3 Square ice.

edges with one of the $2N$ "hydrogen" atoms. Each of the $2N$ "hydrogen" atoms can be in one of the two positions, close or far with respect to a given oxygen atom, as depicted in figure 14.3.

Lieb (Lieb, 1967*a*; Lieb, 1967*b*) solved the ice model using transfer matrices. He obtained the exact value

$$W = \left(\frac{4}{3}\right)^{\frac{3}{2}} = 1.5396007\ldots \tag{14.2}$$

Although we will, in the following, present the transfer matrix method in some detail, our main interest is not in two–dimensional statistical mechanics models such as the ice model. These we rather use as a convenient entry point to study one–dimensional quantum mechanical models. In particular, the approach via two–dimensional statistical mechanics models will enable us to understand the integrability of the corresponding one-dimensional quantum models. For the same reason – our primary interest in one–dimensional quantum models – we will also not present the details of Lieb's calculation of W for the ice model.

Let's reiterate the definition of the square lattice ice model and introduce some useful alternative graphical representation in the following section.

14.2 Definition of the ice model

The ice model is a two–dimensional model for square ice, i.e. it is a statistical mechanical model, as just described. However we also need the (figure 14.4)

- *Ice rule*: there are exactly two hydrogen atoms near each oxygen atom.

We translate the "ice picture" into the *vertex picture*: a hydrogen atom close to the oxygen atom, the vertex, is represented by an arrow pointing toward the oxygen atom, a hydrogen atom far from the oxygen atom by an arrow pointing outward. Another way of decorating the lattice consists of using thick lines for arrows pointing downward or to the left.

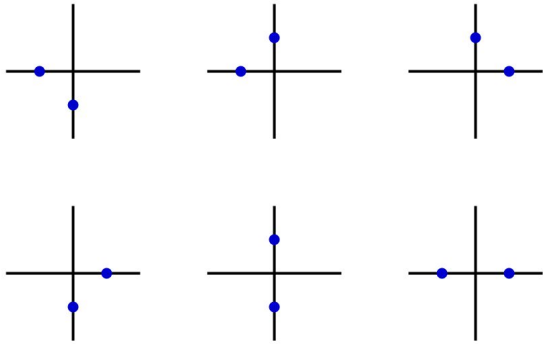

Fig. 14.4 The six local configuration for square ice allowed by the ice rule. The oxygen atoms are not shown.

While we will not be using thick and thin lines to describe the ice model, we shall use the arrow picture, and call the ice model from now on mostly the six–vertex model, the oxygen atoms forming the vertices to or from which the arrows on the bonds or edges point.

The following exercise encourages you to play with a small lattice to become familiar with the notions just introduced. Playing with small lattices is always a good idea to understand what is going on.

Exercise 14.1 Various ways of decorating a small square lattice
Draw a small square lattice, e.g. a 3×3 lattice. The vertices of the lattice correspond to oxygen atoms as described in the text. Decorate this lattice with hydrogen atoms such that the ice rule is satisfied. Now draw your lattice again and decorating it with the arrow configuration corresponding to your first lattice. Finally draw your lattice a third time, now decorating it with thick and thin lines corresponding to your first two configurations.

- How did you treat the horizontal and vertical boundaries of your small finite lattice?
- What is the advantage of the last picture of thick/thin lines?

15
General vertex models

This chapter and the following are partly follow the excellent review article by Leon Takhtajan (1985) and a special lecture series by Tuong Truong (1987) (Ferenc Woynarovich and the author were the only students).

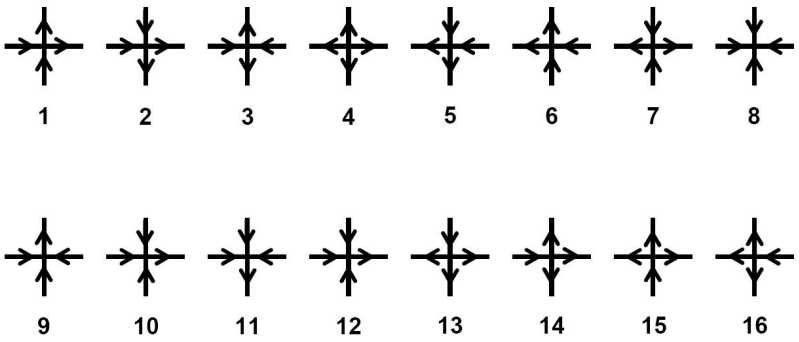

Fig. 15.1 The sixteen possible vertex configurations of the general square lattice vertex model (16 vertex model). The energies and vertex weights (cf. main text) are ϵ_j and v_j, $j = 1, \ldots, 16$, respectively.

Relaxing the ice rule again, we can, of course, define vertex models with more than 6 possible vertices. The general two–dimensional vertex model on a square lattice has $n = 16$ different types of vertices, eight of which, shown in the upper row of figure 15.1, have an even number of outgoing and incoming arrows. Especially vertex type 4 and 8 have all arrows outgoing and incoming, respectively. The other eight vertex types, in the lower row of figure 15.1, have either three incoming and one outgoing arrow, or vice versa.

15.1 General vertex models in two dimensions

Again, there is almost no end of variation. If one allows also for bonds with no arrows, and/or other lattices, e.g. triangular or honeycomb lattices, other numbers of allowed vertices, (e.g. seven and nineteen) can be considered. Such models are of interest, e.g. to study surface

and interface phenomena since bonds with no arrows can be interpreted as bonds with no interaction between the sites connected. Hence, surfaces and interfaces can be described by such models.

Examples can be found in Batchelor, Nienhuis and Warnaaar (Batchelor, Nienhuis and Warnaar, 1989) who discuss a nineteen vertex model on a square lattice, Batchelor and Blöte (Batchelor and Blöte, 1989) who consider a seven vertex model on a honeycomb lattice, or Blum and Shapir (Blum and Shapir, 1990) with a seven vertex model on a square lattice.

15.2 Sixteen– and eight–vertex model

Since, for the time being, the number of allowed vertices will make no difference, we can keep the discussion general for some time before we restrict ourselves again to the six–vertex case. We shall, however, restrict ourselves to the two–dimensional square lattice, and we shall also not consider any of the more purpose built models similar to the ones briefly mentioned in the last section. Our focus will, for some time, be on the sixteen–vertex model of figure 15.1, which we shall soon abandon for the eight–vertex model, since the former seems to be neither integrable nor physically interesting.

15.3 Vertex Boltzmann weights and the partition function

To have some variability, let us call s the number of allowed vertices. We assign an energy

$$\epsilon_j, \qquad j = 1, \ldots s$$

to each of the s vertices. Equivalently, a vertex of type j can be assigned a Boltzmann or statistical weight, called vertex weight in this connection, by

$$v_j = e^{-\beta\epsilon_j} = e^{-\frac{\epsilon_j}{kT}}.$$

Then a given arrangement or configuration of vertices has an energy

$$\mathcal{E} = n_1\epsilon_1 + n_2\epsilon_2 + \ldots + n_s\epsilon_s = \sum_{j=1}^{s} n_j\epsilon_j,$$

where n_j is the number of vertices of type j in this configuration of vertices. The partition function becomes

$$Z = \sum_{\text{allowed configurations}} \exp\left(-\frac{\mathcal{E}}{kT}\right),$$

i.e. a vertex appearing n_j times in the configuration contributes with a statistical weight of $\exp(-\beta n_j\epsilon_j)$ to the partition function. Arrows are a good way to draw and visualize configurations of vertices, but they are a bit tedious for algebraic purposes. Therefore we assign the numbers $\sigma = \pm 1$ to the arrows according to the following correspondences

$$\begin{aligned} \uparrow \text{and} \rightarrow \qquad &\Leftrightarrow \qquad \sigma = +1 \\ \downarrow \text{and} \leftarrow \qquad &\Leftrightarrow \qquad \sigma = -1. \end{aligned}$$

The numbers σ are called spins or spin variables in the sense of the classical spin variables of, e.g., the Ising model. They are *not* quantum mechanical spins. We shall use either notation depending on which is more convenient for the situation.

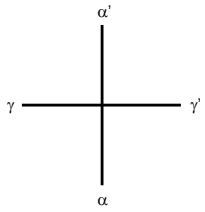

Fig. 15.2 Ice rule $\alpha' + \gamma' = \alpha + \gamma$.

The ice rule for the six–vertex model, written in spin variables is simply (cf. figure 15.2)

$$\alpha' + \gamma' = \alpha + \gamma$$

The s energies ϵ_j or, equivalently, vertex weights v_j characterize the vertex model under consideration. They must be chosen according to the physics we want to describe with the vertex model. However, these s values, energies or vertex weights, are fine to describe an individual vertex, but they are not very suitable to describe a whole two–dimensional lattice of vertices, or even just a single row of vertices. They have yet to be arranged in a proper way for that purpose.

This proper arrangement can be achieved by an object, called the R–matrix, which we shall now discuss.

15.4 R–matrix: matrix of Boltzmann weights of a vertex

Each appearance of one of the $s = 16$ vertices contributes with a Boltzmann weight

$$v_j = \mathrm{e}^{-\beta\epsilon_j} \qquad j = 1, \ldots, 16 \tag{15.1}$$

to the partition function. It will soon prove important to arrange these 16 vertex weights in a 4×4 matrix, the so–called R–matrix. More precisely, a vertex configuration depends on four spin or arrow variables which we denote by $\alpha = \pm 1, \alpha' = \pm 1$ and $\gamma = \pm 1, \gamma' = \pm 1$, such that the Boltzmann weights of a single vertex can be represented by a 4×4 matrix, the R–matrix. The matrix elements are

$$R_{\alpha}^{\alpha'}(\gamma, \gamma') \tag{15.2}$$

with entries

$$v_j = \exp(-\beta\epsilon_j)$$

As the matrix elements in (15.2) may not look like proper matrix elements at first sight, let us spent a moment to clarify this point. Group together the primed and unprimed indices in (15.2) (cf. also figure 15.3) and make the following assignation

$$\{(\alpha', \gamma')\} = \{(+,+), (+,-), (-,+), (-,-)\} \equiv \{1, 2, 3, 4\}$$

such that, with a similar assignation for $\{(\alpha, \gamma)\}$, we now have row and column indices for the R–matrix elements in a more familiar form.

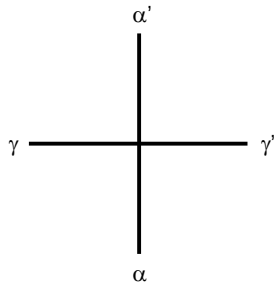

Fig. 15.3 R–matrix element $R_{\alpha}^{\alpha'}(\gamma, \gamma')$.

From the vertices of figure 15.1 and the correspondences between arrows and spin variables, the arrangement of the weights v_j in the R–matrix, and hence the R–matrix itself, is now determined (once we have assigned the weights v_j).

The R–matrix is the fundamental building block of the vertex model which we will now use to build up the two–dimensional lattice step by step, beginning with a single row, then attaching further rows.

If we put two vertices together, described by the R–matrix elements $R_{\alpha_1}^{\alpha_1'}(\gamma_1, \gamma_1')$ and $R_{\alpha_2}^{\alpha_2'}(\gamma_2, \gamma_2')$, respectively, the "inner" spin variables γ_1' and γ_2 have to be equal or the arrows have to point in the same direction (i.e. both left or both right), i.e. we will have to equate the "inner" spin indices $\gamma_1' = \gamma_2$.

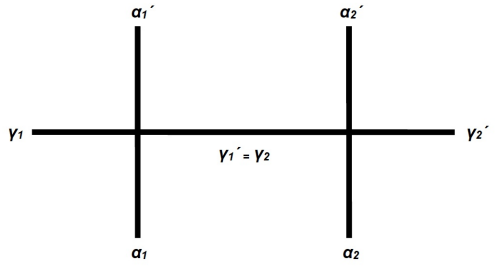

Fig. 15.4 $R_{\alpha_1}^{\alpha_1'}(\gamma_1, \gamma_2) R_{\alpha_2}^{\alpha_2'}(\gamma_2, \gamma_2')$: The R–matrices of two vertices multiplied together.

Moreover we can sum over this inner variable. We obtain the following object

$$\sum_{\gamma_2 = \pm} R_{\alpha_1}^{\alpha_1'}(\gamma_1, \gamma_2) R_{\alpha_2}^{\alpha_2'}(\gamma_2, \gamma_2')$$

which represents the statistical weight of two vertices, i.e. it is the R–matrix element of two vertices.

In the same way we can build a whole row of, say, N vertices and we obtain the matrix elements of a large $2^{N+1} \times 2^{N+1}$ matrix

$$T_{\{\alpha\},\{\alpha'\}}(\gamma_1, \gamma'_N) = \sum_{\gamma_2} \dots \sum_{\gamma_N} R^{\alpha'_1}_{\alpha_1}(\gamma_1, \gamma_2) R^{\alpha'_2}_{\alpha_2}(\gamma_2, \gamma_3) \dots R^{\alpha'_N}_{\alpha_N}(\gamma_N, \gamma'_N) \tag{15.3}$$

The first and last horizontal spin variable are not yet specified. We will specify them through boundary conditions, the simplest of which and the ones we will adopt are cyclic or periodic ones, i.e. $\gamma_1 = \gamma'_N$ [1]. Therefore we can also sum over $\gamma_1 = \pm 1$ in (15.3) to obtain the $2^N \times 2^N$ matrix

$$T_{\{\alpha\},\{\alpha'\}} = \sum_{\gamma_1} \dots \sum_{\gamma_N} R^{\alpha'_1}_{\alpha_1}(\gamma_1, \gamma_2) R^{\alpha'_2}_{\alpha_2}(\gamma_2, \gamma_3) \dots R^{\alpha'_N}_{\alpha_N}(\gamma_N, \gamma_1).$$

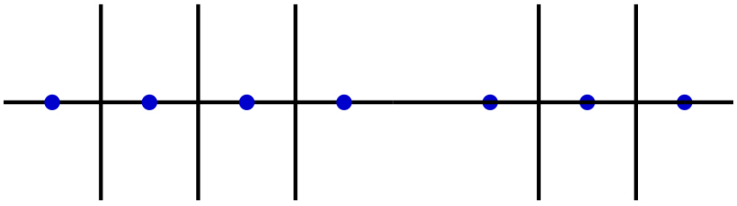

Fig. 15.5 Transfer matrix $T_{\{\alpha\},\{\alpha'\}}$. The vertical spins, $\{\alpha\}$ and $\{\alpha'\}$ are free variables, while the horizontal spins (indicated by blue dots (•)) are summed over. To satisfy periodic boundary conditions the first and the last horizontal spin are identified $\gamma_{N+1} = \gamma_1$.

This matrix will prove to be another central object, called the transfer matrix. It has many "dangling" indices: $\{\alpha\}, \{\alpha'\}$, cf. figure 15.5.

The transfer matrix, in a sense a one–dimensional object since it describes a row of vertices, can now be used to build up the two–dimensional square lattice row by row. For example the statistical weight of a lattice consisting of only two rows can now be written as the matrix product of two transfer matrices (cp. figure 15.6)

$$T^{(2)}_{\{\alpha\},\{\alpha''\}} = \sum_{\{\alpha'\}} T_{\{\alpha\},\{\alpha'\}} T_{\{\alpha'\},\{\alpha''\}}.$$

We repeat this procedure M times. The resulting matrix elements will depend on the lower and upper row of outer horizontal spin variables $\{\alpha\} = \alpha_1, \dots, \alpha_N$ and $\{\alpha^{(M)}\} = \alpha^{(M)}_1, \dots, \alpha^{(M)}_N$, and, hence, will still be the matrix elements of a $2^N \times 2^N$ matrix. This matrix is the Mth power T^M of the transfer matrix T.

Imposing cyclic boundary conditions also in the vertical direction, hence using the topology of a torus (toroidal boundary conditions), we can finally equate and then sum over the last remaining free spin variables $\{\alpha\} = \{\alpha^{(M)}\}$. With this last summation we have performed a trace in the space of the $2^N \times 2^N$ matrices. On the other hand, in total we have summed the

[1] It should be emphasized at this point, that other boundary conditions *are* possible, the use of which may lead to quite interesting physics.

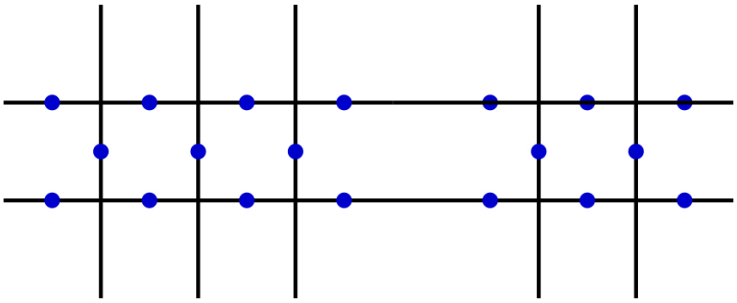

Fig. 15.6 Product of two transfer matrices $T_{\{\alpha\},\{\alpha''\}} = \sum_{\{\alpha'\}} T_{\{\alpha\},\{\alpha'\}} T_{\{\alpha'\},\{\alpha''\}}$. All spin variables indicated by a blue dots (•) are summed over.

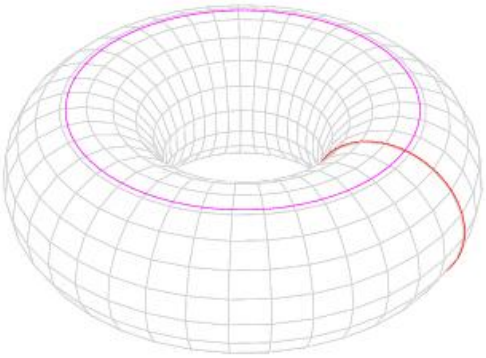

Fig. 15.7 Toroidal boundary conditions.

statistical weights of all possible configurations of the two–dimensional square lattice with toroidal boundary conditions, which is nothing else than the partition function of the lattice model. We are therefore allowed to write

$$Z = \mathrm{trace}(T^M).$$

The computation of the partition function Z has now been reduced to the task of calculating the eigenvalues of the transfer matrix T.

The transfer matrix, by construction from the non–negative vertex weights, is a matrix with non–negative matrix elements. The Perron–Frobenius theorem of matrix theory then tells us that, for such a matrix, there is a unique positive real eigenvalue λ_1 such that for all other eigenvalues λ_j

$$\lambda_1 > |\lambda_j| \quad j \neq 1$$

holds.

Supposing we had achieved this task of diagonalizing the transfer matrix T, then, according to the Perron–Frobenius theorem, the partition function becomes

$$Z = \sum_{j=1}^{2^N} \lambda_j^M = \lambda_1^M \left(1 + \left(\frac{\lambda_2}{\lambda_1}\right)^M + \left(\frac{\lambda_3}{\lambda_1}\right)^M + \ldots\right),$$

which, in the limit $M \to \infty$ reduces to just

$$Z = \lambda_1^M.$$

This result looks marvelously simple, but we are, of course, not there yet. We first have to find a way to achieve the diagonalization of the transfer matrix.

15.5 Integrability and the transfer matrix

The transfer matrix allows us to address the question of integrability. We pose the following question: Can we diagonalize a whole family of transfer matrices simultaneously? In other words, can we find arbitrary sets of Boltzmann weights v_j and v_j' and the corresponding R–matrices such that the corresponding transfer matrices $T = T(v_j)$ and $T' = T(v_j')$ commute

$$TT' - T'T \equiv [T, T'] = 0. \tag{15.4}$$

Before we search for an answer to this question, we shall pause for a moment and investigate the implications of (15.4).

15.6 Commuting transfer matrices

Recall the expression we found for the partition function of vertex models for the case $N = 8$, i.e. the eight–vertex model

$$Z = \sum_{\text{allowed configurations}} \exp\left(-\frac{\mathcal{E}}{kT}\right),$$

where

$$\mathcal{E} = n_1\epsilon_1 + n_2\epsilon_2 + \ldots + n_8\epsilon_8 = \sum_{j=1}^{8} n_j\epsilon_j.$$

We can also write the partition function in the following form

$$Z = \sum_{\text{allowed configurations}} a^{n_a} b^{n_b} c^{n_d} d^{n_d} \tag{15.5}$$

where we introduced the notation for the vertex weights

$$a = v_1 = v_5, \quad b = v_2 = v_6, \quad c = v_3 = v_7 \quad d = v_4 = v_8,$$

and the numbers of times the vertices appear in a configuration are

$$n_a = n_1 + n_5, \quad n_b = n_2 + n_6, \quad n_c = n_3 + n_7 \quad n_d = n_4 + n_8.$$

In the rectangular lattice of $N \times M$ lattice sites, we are considering, these numbers sum up $n_a + n_b + n_c + n_d = NM$. We assume that there is a spin or arrow inversion symmetry, i.e. flipping all spin variables, $\sigma \to -\sigma$ or, equivalently, reversing all arrows, of a vertex does not change the energy. This symmetry holds quite obviously as long as there are no external fields which might break the symmetry.

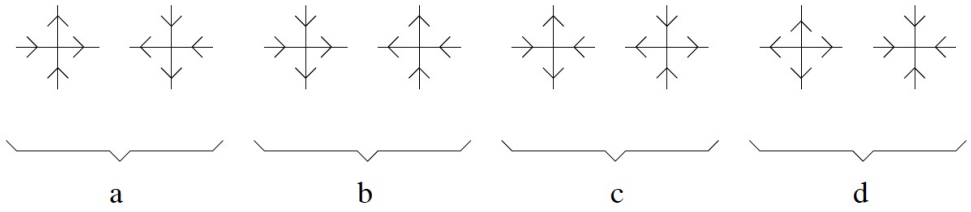

Fig. 15.8 The vertices of the eight–vertex model in the $abcd$ notation. The arrow reversal symmetry is quite obvious. Note that the sequence of the vertices is different from the one in 15.1. The vertices which are arrow–reversal symmetric to each other have ben grouped in pairs.

Writing the partition function in the form of (15.5) is quite instructive, as we shall see. First we note that it can be written, e.g., as

$$Z = d^{NM} \sum_{\text{allowed configurations}} \left(\frac{a}{d}\right)^{n_a} \left(\frac{b}{d}\right)^{n_b} \left(\frac{c}{d}\right)^{n_c}.$$

The partition function thus depends only on the three ratios

$$\frac{a}{d}, \quad \frac{b}{d}, \quad \text{and} \quad \frac{c}{d}.$$

We can, thus, describe the eight–vertex model partition function in the three–dimensional space space formed by these three vertex weight ratios.

The same conclusions hold also for the transfer matrix T. Generically, it will turn out, transfer matrices for different sets of vertex weights will not commute. However, we shall also see that we can find families of curves, parametrized by a parameter u, in the space of vertex weights $\mathbf{w} \equiv \left(\frac{a}{d}, \frac{b}{d}, \frac{c}{d}\right)^T$

$$\mathbf{w} = \mathbf{w}(u) \equiv \begin{pmatrix} \frac{a}{d}(u) \\ \frac{b}{d}(u) \\ \frac{c}{d}(u) \end{pmatrix},$$

such that the transfer matrices for different positions, i.e. different values of u, on those curves do commute. Different members of the families of curves are distinguished by a parameter λ, the crossing parameter. The parameter u is called the spectral parameter for reasons that will become clear a little later.

In figure 15.9, this is illustrated for the six–vertex model where there are only two parameters $\left(\frac{a}{c}, \frac{b}{c}\right)$ spanning the parameter space, since the d–vertex in the six–vertex model is not allowed by the ice rule.

Now let's get back to how we can actually perform this programme of finding commuting transfer matrices.

15.7 Monodromy matrix

To make progress toward an answer of the question we posed in section 15.5, we recall an object which we encountered above in (15.3)

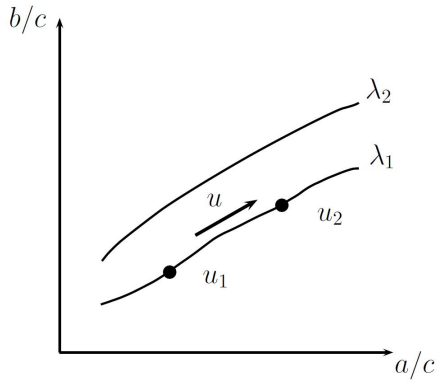

Fig. 15.9 Families of curves $\mathbf{w}(u)$ in the parameter space of the six–vertex model. The different families are specified by a parameter λ, the crossing parameter. For a fixed value of λ, pairs of transfer matrices for different positions along the curve, e.g. u_1 and u_2, commute.

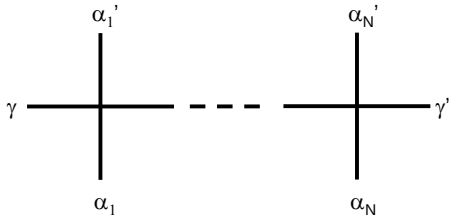

Fig. 15.10 The monodromy matrix. All the inner spin variables, indicated by the dashed line, are summed over, while the two outer spin variables γ and γ' are free.

$$\mathcal{T} = T_{\{\alpha\},\{\alpha'\}}(\gamma,\gamma') = \sum_{\gamma_2} \dots \sum_{\gamma_N} R^{\alpha_1'}_{\alpha_1}(\gamma,\gamma_2') R^{\alpha_2'}_{\alpha_2}(\gamma_2,\gamma_3') \dots R^{\alpha_N'}_{\alpha_N}(\gamma_N,\gamma')$$

namely the transfer matrix where no boundary conditions have been specified, i.e. where the first and last horizontal spin variables are free. This object bears the name monodromy matrix. If we focus on the dependence of the monodromy matrix of these first and last horizontal spin variables, it can be regarded as a 2×2 matrix

$$\mathcal{T} = \mathcal{T}(\gamma,\gamma')$$

whose four matrix elements are, however, themselves $2^N \times 2^N$ matrices. The transfer matrix can then be written as a trace in the two–dimensional space of these $2^N \times 2^N$ matrices

$$T = \mathrm{tr}\,\mathcal{T} = \mathcal{T}(+,+) + \mathcal{T}(-,-). \tag{15.6}$$

We use the notation tr for this 2×2 trace, reserving the notation trace for a trace of $2^N \times 2^N$ matrices.

The monodromy matrix $\mathcal{T}$, as well as the transfer matrix T, obviously are objects which describe an entire row of vertices. The derivation of $\mathcal{T}$ or T, we have given above, is built on a local object, the R–matrix. Next, we want to construct the monodromy matrix $\mathcal{T}$ from a different set of elementary matrices, which are, on the one hand also local objects in the sense that they describe a particular vertex. On the other hand, they are connected with a whole row as we shall see in a moment. The whole point is, similar to when we moved from vertex weights v_j to their neat arrangement as R–matrix, that we now write the R–matrix in such a way that the vertex to which it refers becomes easily localized in the row of vertices we consider. This is what the L–matrix does. The L–matrix bears a position label and is written more precisely as L_n, where n refers to the nth vertex of the row. The R–matrix does not bear such a position label.

First, let's write the R–matrix in a more explicit form as

$$R_{\alpha}^{\alpha'}(\gamma, \gamma') = \sum_{i,j=1}^{4} w_{ij} \sigma^i_{\gamma',\gamma} \sigma^j_{\alpha,\alpha'}$$

where the matrices σ^j, $j = 1, 2, 3, 4$ are the Pauli matrices and the 2×2 unit matrix

$$\sigma^1 = \sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma^4 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We postpone a discussion of the detailed form of the Boltzmann weights w_{ij}, especially their connection to the previously introduced weights v_j, since they are not relevant yet for the still formal development of our arguments. We just note now that any complex 4×4 matrix can be written in such a way.

Exercise 15.1 First, prove that every 2×2 matrix A can be written as a linear combination of the 2×2 unit matrix and the Pauli matrices.

Then, second, also prove that every 4×4 matrix A can be written as a linear combination of the tensor product matrices $\sigma^i \otimes \sigma^j$ where the σ^i are the 2×2 unit matrix and the Pauli matrices.

For the definition and properties of tensor product matrices, see the following section 15.7.1.

As we shall see, the drawback of the R–matrix is that, although it clearly is an object connected to one vertex, it cannot be easily located within the lattice, it does not bear a spatial coordinate. Motivated by this form of the R–matrix, however, we can easily write down a matrix which does depend on the position within the lattice, more specifically on the position of the vertex we are looking at in a row of vertices. This object is therefore a $2^{N+1} \times 2^{N+1}$ matrix, called the local L–matrix or, more commonly, the L–operator.

$$L_n = \sum_{i,j=1}^{4} w_{ij} \sigma^i \otimes \sigma^j_n \tag{15.7}$$

where

$$\sigma_n^j = \overbrace{I \otimes I \otimes \ldots \underbrace{\sigma^j}_{n} \otimes \ldots I}^{N} \tag{15.8}$$

is a $2^N \times 2^N$ matrix which acts non–trivially only on the position n along the row of vertices.

15.7.1 Interlude: direct or tensor product of matrices

In the following, the direct or tensor product of two matrices will be important. Moreover we have just used it in writing down (15.7) and (15.8). Therefore, let's define it explicitly:

For the 2×2 (complex) matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

the tensor product $A \otimes B$ is defined by[2]

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \left(\begin{array}{cc|cc} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ \hline a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{array}\right).$$

There is an easy way to remember this

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix}.$$

The 2×2 matrices act on a two–dimensional complex vector space

$$V = \mathcal{C}^2,$$

whose elements are two–component vectors

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

with $x_i \in \mathcal{C}$.

The tensor product $A \otimes B$ acts on the tensor product space $V \otimes V$ whose elements are of the form

$$x \otimes y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \left(\begin{array}{c} x_1y_1 \\ \hline x_1y_2 \\ \hline x_2y_1 \\ x_2y_2 \end{array}\right). \tag{15.9}$$

It is useful to recall the equation

$$(A \otimes B)(x \otimes y) = Ax \otimes By.$$

[2]The horizontal and vertical lines help to organize the matrix elements; they have no other significance, and can be omitted. This practice follows Rafael Nepomechie (1999).

which generalizes to

$$(A_1 \otimes A_2 \otimes \dots A_n)(x_1 \otimes x_2 \otimes \dots x_n) = A_1 x_1 \otimes A_2 x_2 \otimes \dots A_n x_n.$$

In words, a 2×2 matrix in the ℓ's position ($1 \leq \ell \leq n$) of an n–fold tensor product acts non–trivially only on the two–dimensional vector space in the ℓ's position of the n–fold tensor product space

$$\overset{1}{\underset{\downarrow}{V}} \otimes \cdots \otimes \overset{\ell}{\underset{\downarrow}{V}} \otimes \cdots \otimes \overset{n}{\underset{\downarrow}{V}},$$

i.e. on the corresponding vector in the ℓ's position of the n–fold tensor product of vectors.

Let us make a few simple observations with the tensor product we just defined.

15.7.1.1 **Simple applications of the tensor product**.
Many of the tensor products we will be interested in are related to the Pauli matrices and the corresponding spinors

$$|+\rangle = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle = |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Consider a lattice of two sites only to each of which we attach a quantum spin–$\frac{1}{2}$ object. Then the Pauli matrices acting on the first object only will be

$$\sigma_1^x = \sigma^x \otimes \mathbb{I}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right),$$

while for the second object, we have

$$\sigma_2^x = \mathbb{I}_2 \otimes \sigma^x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right),$$

where $\mathbb{I}_2$ is the 2×2 unit matrix. Similarly, we can obtain the 4×4 matrices σ_i^y and σ_i^z ($i = 1, 2$).

The following little exercise is suggested to gain practice.

Exercise 15.2 Two spins
Show that for two spins we have

$$\begin{aligned} \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 &= \sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y + \sigma_1^z \sigma_2^z \\ &= \sigma^x \otimes \sigma^x + \sigma^y \otimes \sigma^y + \sigma^z \otimes \sigma^z \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

The permutation operator (or rather permutation matrix)

$$\mathbb{P} = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix}$$

acts on the tensor product of two spinors of the tensor product space $V \otimes V$ as

$$\mathbb{P}x \otimes y = \begin{pmatrix} x_1 y_1 \\ x_2 y_1 \\ x_1 y_2 \\ x_2 y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \otimes \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = y \otimes x.$$

The permutation operator squares to the 4×4 identity[3]

$$\mathbb{P}^2 = \mathbb{I}_2 \otimes \mathbb{I}_2 = \mathbb{I}_4.$$

Furthermore, combining this with the exercise 15.2 we immediately observe

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 + \mathbb{I}_4 = 2\mathbb{P}. \tag{15.10}$$

This is an interesting result to which we will come back later. Let us, for the moment, just note that the two–spin operator

$$h_{i,i+1} = \frac{1}{2}\left(\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} + \mathbb{I}_4\right), \tag{15.11}$$

acting on spinors at neighbouring sites, is just the permutation operator $\mathbb{P}$.

15.8 Further to the L–operator

Let us now return to the operator L_n, (15.7). Using the terminology of tensor products introduced in section 15.7.1, the L_n–operator can now be formally decomposed and written as a 2×2 matrix with elements $\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n$ and $\hat{\delta}_n$ which are themselves matrices or operators of dimension $2^N \times 2^N$

$$L_n = \begin{pmatrix} \hat{\alpha}_n & \hat{\beta}_n \\ \hat{\gamma}_n & \hat{\delta}_n \end{pmatrix}. \tag{15.12}$$

Now we can write the monodromy matrix in terms of the local L–operators

$$\mathcal{T} = L_1 L_2 \ldots L_N = \prod_{n=1}^{N} L_n$$

or, because of (15.12)

$$\mathcal{T} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{15.13}$$

Now we can formally perform the trace of (15.6) to obtain

$$T = A + D. \tag{15.14}$$

So far, of course, this was all purely formal and we still have not diagonalized T. This will only be possible once we specify the model, i.e. once we define the energies ϵ_j and therewith

[3] Often we, as many other people, shall be sloppy about unit matrices and identity operators, and not indicate their proper dimensions; we even go as far as just writing "1", as if it were the number 1.

the Boltzmann weights v_j. However, we can now give a *sufficient* condition (not a *necessary* one, though!) for the transfer matrices to commute.

15.9 Yang–Baxter relations

This sufficient condition for commuting transfer matrices are the *Yang–Baxter relations* which we can write formally

$$R''(\mathcal{T} \otimes \mathcal{T}') = (\mathcal{T}' \otimes \mathcal{T})R'' \tag{15.15}$$

where the dyadic product is given by

$$\mathcal{T} \otimes \mathcal{T}' = \begin{pmatrix} A\mathcal{T}' & B\mathcal{T}' \\ C\mathcal{T}' & D\mathcal{T}' \end{pmatrix}.$$

The R–matrix and the two monodromy matrices are to be taken for different sets of vertex weights v_j or w_{ij} which is indicated by no prime, prime and double prime, i.e.

$$\mathcal{T} = \mathcal{T}(v_j) \qquad \mathcal{T}' = \mathcal{T}(v'_j) \qquad R'' = R(v''_j).$$

The Yang–Baxter relations equation (15.15) implies that the transfer matrices for different sets of vertex weights commute, equation (15.4). To see this rewrite the Yang–Baxter relations (15.15) as

$$(\mathcal{T} \otimes \mathcal{T}') = {R''}^{-1}(\mathcal{T}' \otimes \mathcal{T})R'' \tag{15.16}$$

and take the appropriate trace in the 4 dimensional space.

Exercise 15.3 Convince yourself that taking the trace in an appropriate four–dimensional space of the previous equation, (15.16), is sufficient for the transfer matrices for different sets of vertex weights to commute

$$[T, T'] = 0. \tag{15.17}$$

Using the local L–operators we can also write down local Yang–Baxter relations

$$R''(L_n \otimes L'_n) = (L'_n \otimes L_n)R''$$

or, more explicitly, using the weights w_{ij} this time

$$R(w''_{ij})(L_n(w_{ij}) \otimes L_n(w'_{ij})) = (L_n(w'_{ij}) \otimes L_n(w_{ij}))R(w''_{ij}). \tag{15.18}$$

We pause for a moment to clarify what our results so far mean. We found the Yang–Baxter relations as a sufficient condition for the integrability of the transfer matrix of the vertex model. Integrability here means that we can diagonalize the transfer matrix of a given vertex model and hence calculate the partition function.

15.10 More on Yang–Baxter relations

The Yang–Baxter relations appear in many different versions. We have already encountered two versions, the first involving, besides the R–matrix, the whole monodromy matrix $\mathcal{T}$. The second has the *local* L_n matrices in place of the monodromy matrix $\mathcal{T}$, suggesting that there is a version of the Yang–Baxter relations which employs only R–matrices.

While the versions of the Yang–Baxter relations involving the monodromy matrix $\mathcal{T}$ or the L_n matrices are connected to the square lattice of the vertex model, the version involving only R matrices consists of just three vertices attached to each other to form a triangle (cf. figure 15.11).

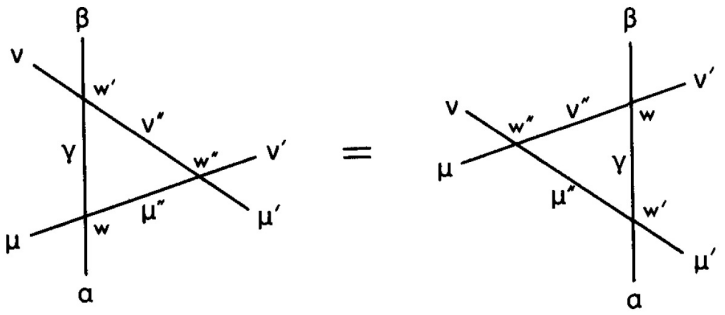

Fig. 15.11 Yang–Baxter relations for the R matrices. From Rodney Baxter's book.

Exercise 15.4 Write down the explicit Yang–Baxter relations corresponding to figure 15.11.

15.11 Exploiting integrability

We will now exploit the local Yang–Baxter relations, (15.18), i.e. the integrability condition. To do so, let's finally establish an explicit relation between the vertex weights v_j and the elements of the R–matrix:

$$R_{\alpha}^{\alpha'}(\gamma,\gamma') = \sum_{i,j=1}^{4} w_{ij}\sigma^{i}_{\gamma',\gamma}\sigma^{j}_{\alpha,\alpha'}. \tag{15.19}$$

The logic is the following: We are searching for vertex weights w_{ij} such that the local Yang–Baxter relations are satisfied.

So far, we have not imposed any restrictions on the model. Thus the R–matrix (15.19) still corresponds to the most general case of a sixteen–vertex model. Unfortunately, there are no vertex weights known for a sixteen–vertex model and, hence, there is no R–matrix either such that the local Yang–Baxter relations (15.18) could be satisfied.

We will accordingly restrict our attention to the eight–vertex model and eventually even the six–vertex model where (15.19) simplifies to

$$R_{\alpha}^{\alpha'}(\gamma,\gamma') = \sum_{j=1}^{4} w_j \sigma^j_{\gamma',\gamma}\sigma^j_{\alpha,\alpha'}. \tag{15.20}$$

and the L–operators are

$$L_n = \sum_{j=1}^{4} w_j \sigma^j \otimes \sigma_n^j = \begin{pmatrix} w_4 I_n + w_3\sigma_n^z & w_1\sigma_n^x - iw_2\sigma_n^y \\ w_1\sigma_n^x + iw_2\sigma_n^y & w_4 I_n - w_3\sigma_n^z \end{pmatrix}$$

$$R_{+}^{+}(+,+) \equiv v_1 = w_3 + w_4$$

Furthermore

$$v_2 = w_4 - w_3 \qquad v_3 = w_1 + w_2 \qquad v_4 = w_1 - w_2$$

For the six–vertex model, the ice rule requires

$$v_4 = 0 \qquad \Leftrightarrow \qquad w_1 = w_2$$

The local Yang–Baxter relations, equation (15.18) are equations relating the different sets of vertex weights w_j'', w_j' and w_j. For the eight–vertex model, these equations describe an elliptic curve which can be parametrized by elliptic functions $\mathrm{sn}(u,k)$ (elliptic sine of modulus k). These remarks connect the present discussion with the remarks in section 15.6, especially figure 15.9.

At this point, we at last specialize to the six–vertex model, pointing the reader to the literature (see Baxter (1982)) for the further treatment of the eight–vertex model.

16
Six–vertex model

We restrict ourselves in the following to the six–vertex model, where the corresponding curves obtained from (15.18) can be parametrized by ordinary trigonometric sine functions [1]. These would also emerge from appropriate limits (i.e. $k \to 1$) of the elliptic functions.

16.1 Parameterization of the six–vertex model

The local Yang–Baxter relations, (15.18), of the six–vertex model are satisfied for the parameterization

$$\begin{aligned} v_1 &= w_3 + w_4 = \sin(\lambda + \eta) \\ v_2 &= w_4 - w_3 = \sin(\lambda - \eta) \\ v_3 &= 2w_1 = \sin(2\eta) \end{aligned} \tag{16.1}$$

i.e. we have two parameters whose significance will become clear in the following.

In fact, we shall regard λ, the so–called spectral parameter, as the main parameter, which will play a major role in diagonalizing the transfer matrix. The parameter η will turn out to describe different regimes of the model.

The replacements in the parameterizations, (16.1),

$$\lambda \to \mu$$

and

$$\lambda \to \lambda - \mu + \eta$$

give sets w'_j and w''_j, respectively, such that the local Yang–Baxter relations (15.18) are satisfied. The local Yang–Baxter relations thus take the form

$$R(\lambda - \mu)(L_n(\lambda) \otimes L_n(\mu)) = (L_n(\mu) \otimes L_n(\lambda))R(\lambda - \mu)$$

The L– and R–matrices can now be given explicitly

$$L_n(\lambda) = \begin{pmatrix} w_4(\lambda) I_n + w_3(\lambda)\sigma_n^z & \frac{\sin 2\eta}{2}\sigma_n^- \\ \frac{\sin 2\eta}{2}\sigma_n^+ & w_4(\lambda) I_n - w_3(\lambda)\sigma_n^z \end{pmatrix}$$

where

$$w_4(\lambda) \pm w_3(\lambda) = \sin(\lambda \pm \eta)$$

[1] This is only one possible parameterization which corresponds to the anisotropic XXZ Heisenberg quantum spin chain with anisotropy parameter $|\Delta| \leq 1$.

and

$$R(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & b(\lambda) & c(\lambda) & 0 \\ 0 & c(\lambda) & b(\lambda) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where

$$b(\lambda) = \frac{\sin 2\eta}{\sin(\lambda + 2\eta)}$$

$$c(\lambda) = \frac{\sin \lambda}{\sin(\lambda + 2\eta)}$$

Now we are in a position to start the diagonalization of the transfer matrix. Introduce the local vacuum state

$$\omega_n = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n,$$

i.e. the state which is defined by

$$\sigma_n^z \omega_n = \sigma_n^z \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = + \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = +\omega_n,$$

$$\sigma_n^+ \omega_n = \sigma_n^+ \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = 0,$$

$$\sigma_n^- \omega_n = \sigma_n^- \begin{pmatrix} 1 \\ 0 \end{pmatrix}_n = + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_n,$$

and

$$\sigma_n^z \begin{pmatrix} 0 \\ 1 \end{pmatrix}_n = - \begin{pmatrix} 0 \\ 1 \end{pmatrix}_n.$$

The L–operator acts on ω_n to produce

$$L_n(\lambda)\omega_n = \begin{pmatrix} \alpha(\lambda)\omega_n & * \\ 0 & \delta(\lambda)\omega_n \end{pmatrix} \tag{16.2}$$

which is an upper triangular matrix, where

$$\alpha(\lambda) = \sin(\lambda + \eta)$$

and

$$\delta(\lambda) = \sin(\lambda - \eta)$$

Because this is a triangular matrix, the matrix element indicated by an $*$ will not play any further role in the analysis and we don't bother to write it down explicitly.

The reference state

$$\Omega = \prod_{n=1}^{N} \otimes \omega_n$$

assumes the role of the local vacuum state for the monodromy matrix. Using the triangular form of (16.2), we immediately obtain

$$\mathcal{T}\,\Omega = \begin{pmatrix} \alpha^N(\lambda)\Omega & * \\ 0 & \delta^N(\lambda)\Omega \end{pmatrix}$$

Now, the formal developments we spent quite some space on before introducing a number of formal quantities, are starting to pay off. The transfer matrix was

$$T = tr\,(\mathcal{T}) = A(\lambda) + D(\lambda).$$

We can now immediately read off an eigenvalue of the transfer matrix and also the eigenvalues of the matrices A and D

$$A(\lambda)\Omega = \alpha^N(\lambda)\Omega$$

and

$$D(\lambda)\Omega = \delta^N(\lambda)\Omega$$

and, finally,

$$T(\lambda)\Omega = (\alpha^N(\lambda) + \delta^N(\lambda))\Omega.$$

We note that Ω is an eigenstate of the transfer matrix with eigenvalue

$$\alpha^N(\lambda) + \delta^N(\lambda).$$

We now want to construct other eigenstates of the transfer matrix $T(\lambda)$. To make progress with this we use the fundamental relation, the Yang–Baxter relation

$$R(\lambda - \mu)(\mathcal{T}(\lambda) \otimes \mathcal{T}(\mu)) = (\mathcal{T}(\mu) \otimes \mathcal{T}(\lambda))R(\lambda - \mu) \tag{16.3}$$

It is this constructive, algebraic approach which gives the whole approach its name: algebraic Bethe ansatz.

Writing the Yang–Baxter relations in components, we obtain 16 commutation relations between the operators A, B, C and D. Fortunately many of them are rather trivial or do not give any further information. In fact, we only need 3 of them for our purpose. The operator B commutes for different values of the spectral parameter

$$[B(\lambda), B(\mu)] = 0. \tag{16.4}$$

The other two commutation relations are more complicated.

$$A(\lambda)B(\mu) = \frac{1}{c(\mu - \lambda)}B(\mu)A(\lambda) - \frac{b(\mu - \lambda)}{c(\mu - \lambda)}B(\lambda)A(\mu) \tag{16.5}$$

and

$$D(\lambda)B(\mu) = \frac{1}{c(\lambda - \mu)}B(\mu)D(\lambda) - \frac{b(\lambda - \mu)}{c(\lambda - \mu)}B(\lambda)D(\mu) \tag{16.6}$$

We note, that the operator $C(\lambda)$ does not play a further role in our analysis. It can be shown, that $C(\lambda)$ acts as an annihilation operator in the same sense as $B(\lambda)$ acts as a creation operator (see below).

Exercise 16.1 Explicitly derive (16.4–16.6) from the Yang–Baxter relation, (16.3) of the six–vertex model. You may want to derive further commutation relations.

These commutation relations suggest an interpretation of $B(\lambda)$ as a kind of creation operator. Acting with the B–operator for different values λ_j of the spectral parameter on the reference state Ω, we construct the state

$$\Phi(\lambda_j) = \prod_{j=1}^{\ell} B(\lambda_j)\Omega,$$

which apparently depends on the string of mutually different spectral parameters $\lambda_j,\ j = 1, \ldots, \ell$.

We expect that this state will be an eigenstate of the transfer matrix. We shall see, however, that this is not immediately true. Rather we find that

$$T(\lambda)\Phi = \Lambda(\lambda)\Phi + \ \textit{"junk"–terms.}$$

Provided we can make the "junk"–terms disappear, this would indeed be an eigenvalue equation.

We find, that the "junk" or unwanted terms indeed do disappear, provided the distinct values of the spectral parameters $\{\lambda_j\}$ satisfy

$$\frac{\alpha^N(\lambda_j)}{\delta^N(\lambda_j)} = \prod_{k=1,k\neq j}^{\ell} \frac{c(\lambda_k - \lambda_j)}{c(\lambda_j - \lambda_k)} \qquad j = 1, \ldots, \ell. \tag{16.7}$$

With this provision, we find the eigenvalue

$$\Lambda(\lambda; \lambda_1, \ldots, \lambda_\ell) = \alpha^N(\lambda) \prod_{j=1}^{\ell} \frac{1}{c(\lambda_j - \lambda)} + \delta^N(\lambda) \prod_{j=1}^{\ell} \frac{1}{c(\lambda - \lambda_j)}. \tag{16.8}$$

The proof of these statements is rather lengthy but elementary. It only uses the commutation relations (16.4-16.6) to carry the operators $A(\lambda)$ and $D(\lambda)$ of the transfer matrix $T(\lambda) = A(\lambda) + D(\lambda)$ through the strings of $B(\lambda_j)$–operators. The equations (16.7) stem from the terms in (16.5) and (16.6) where the arguments of A or D, respectively, and B are exchanged, i.e. the second terms on the respective right–hand sides of the commutation relations (16.5) and (16.6). These unwanted terms cancel if (16.7) hold.

Exercise 16.2 Derive the above results explicitly for the cases $\ell = 1$ and $\ell = 2$, respectively.

The equations (16.7) are the celebrated Bethe ansatz equations, now derived within the algebraic framework.

We are now only left with the task to show the connection of this approach with the Heisenberg quantum spin–$1/2$ chain. To do this, we have to consider the transfer matrix $T(\lambda)$ in some detail.

19.5 Excitations of the XXZ spin chain

For any given set of quantum numbers $\{\lambda_i\}$ the Bethe ansatz equations (19.10)

$$k_i N = 2\pi\lambda_i + \sum_{j=1, j\neq i}^{M} \Theta(k_i, k_j) \quad \text{where} \quad \lambda_i = 0, 1, \ldots, N-1; \quad i = 1, \ldots, M$$

describe a state. In general, we do not know beforehand which state this will be. This is one of the main problems to be solved within the framework of the coordinate Bethe ansatz. Generically, by prescribing a set of quantum numbers $\{\lambda_i\}$, we will have hit on some excited state. At this stage, we would have to invest a lot of work to determine the low–lying excitations of the spin chain.

We shall, instead, restrict ourselves to a few simple examples which we examine in the thermodynamic limit $N \to \infty$ thereby simplifying the analysis considerably. The complexity of the Bethe ansatz equations will become apparent even for the simplest examples.

19.5.1 One magnon state in the ferromagnetic chain

The simplest case is the isotropic ferromagnetic spin chain, i.e. $\Delta = -1$, whose ground state is the reference state $|0\rangle$ introduced in section 18.3, (18.6). In section 19.3, we have seen that the quasi–momenta for the isotropic ferromagnetic chain can be parametrized by

$$k_i = 2\tan^{-1} 2v_i$$

where, for real $0 \leq k_i \leq 2\pi$, the variables v_i are distributed over the whole real line. The phases, which, in the new variables, depend only on the difference, are given by

$$\Theta(k_i, k_j) = 2\cot^{-1}(v_i - v_j)$$

and the energy of the state by

$$E_M = \frac{1}{2}\sum_{j=1}^{M} \frac{1}{v_j + \frac{1}{4}} = \frac{1}{2}\sum_{j=1}^{M} \frac{dk_j}{dv_j}.$$

The Bethe ansatz equations in product form are then for even N

$$\left(\frac{v_i - \frac{i}{2}}{v_i + \frac{i}{2}}\right)^N = \prod_{j=1, j\neq i}^{M} \frac{v_i - v_j - i}{v_i - v_j + i}, \quad i = 1, \ldots, M. \tag{19.22}$$

For $M = 1$ flipped spin, this reduces to

$$\left(\frac{v - \frac{i}{2}}{v + \frac{i}{2}}\right)^N = 1$$

or, equivalently,

$$\frac{v - \frac{i}{2}}{v + \frac{i}{2}} = e^{ik}$$

with e^{ik} one of the Nth roots of unity, i.e.

$$k = \frac{2\pi\lambda}{N}, \quad \lambda = 0, 1, 2, \ldots, N-1.$$

In the thermodynamic limit $N \to \infty$ the quasi–momentum k gets indeed distributed over the whole interval $0 \leq k \leq 2\pi$ and, hence, the rapidity variable v over the whole real line.

An excitation of this type is called a *magnon* or spin wave. The energy of a magnon of the ferromagnetic spin chain is given by

$$E_1(k) = 1 + \cos k.$$

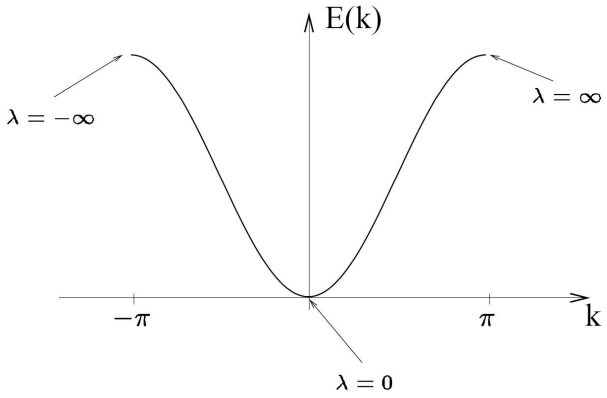

Fig. 19.5 Energy spectrum of one magnon excitations of the isotropic ferromagnetic Heisenberg model. Indicated are also the corresponding values of the rapidity variable v (denoted by λ in the figure). Note that, compared to the calculation in the text, a different Brillouin zone has been chosen: $-\pi \leq k \leq \pi$ instead of $0 \leq k \leq 2\pi$.

Since we choose the Hamiltonian in such a way, that the ferromagnetic reference state $|0\rangle$ has energy $E_0 = 0$, we obviously have the result, that for the magnon energy for all quasi–momenta $0 \leq k \leq 2\pi$

$$E_1(k) \geq E_0.$$

This demonstrates that the energy of the reference state is indeed a good candidate for the absolute ground state of the spin chain.

More ambitious is the case of two flipped spins to which we now turn.

19.5.2 Bound states or spin complexes

In the case of two flipped spins, $M = 2$, we have the two coupled Bethe ansatz equations

$$\left(\frac{v_1-\frac{i}{2}}{v_1+\frac{i}{2}}\right)^N = \frac{v_1-v_2-i}{v_1-v_2+i}, \tag{19.23}$$

$$\left(\frac{v_2-\frac{i}{2}}{v_2+\frac{i}{2}}\right)^N = \frac{v_2-v_1-i}{v_2-v_1+i}. \tag{19.24}$$

19.5.2.1 Real solutions: two magnon state of the ferromagnetic spin chain. The first question to answer is whether these equations admit real solutions v_1 and v_2.

Using again the identity

$$\exp\left(2i\cot^{-1}x\right) = \frac{x+i}{x-i}$$

for real x, we can rewrite the two coupled equations

$$\frac{v_1-\frac{i}{2}}{v_1+\frac{i}{2}} = \left(\frac{v_1-v_2-i}{v_1-v_2+i}\right)^{\frac{1}{N}} = e^{-i\phi(v_1-v_2)} = e^{ik_1}$$

$$\frac{v_2-\frac{i}{2}}{v_2+\frac{i}{2}} = \left(\frac{v_2-v_1-i}{v_2-v_1+i}\right)^{\frac{1}{N}} = e^{i\phi(v_1-v_2)} = e^{ik_2}$$

where

$$k_i = \frac{2\pi\lambda_i}{N}, \quad \lambda_i = 0,1,\ldots,N-1, \quad i=1,2$$

and

$$\phi(v_1-v_2) = \frac{1}{N}\Theta(v_1-v_2) = \frac{2}{N}\cot^{-1}(v_1-v_2)$$

using the principal value of the $\cot^{-1}$ function. Since the principal value of the $\cot^{-1}$ function is a bounded function, $\phi(v_1-v_2)$ vanishes as $N\to\infty$. Therefore, in this limit,

$$\frac{v_i-\frac{i}{2}}{v_i+\frac{i}{2}} = e^{ik_i}, \quad i=1,2.$$

In other words, for real v_1 and v_2, the Bethe ansatz equations for two flipped spins decouple and reduce to two independent equations of the type we know from the discussion of the one magnon state. Thus, we have the two magnon state of the ferromagnetic spin chain with the quasi–momenta k_1 and k_2. The energy of this state is given by

$$E_2 = 1+\cos k_1 + 1 + \cos k_2 = 2\left(1+\cos\frac{k_1+k_2}{2}\cos\frac{k_1-k_2}{2}\right).$$

Again we see, that $E_2 \geq E_0$.

However, in the case of two flipped spins, the two magnon state does not exhaust the possible states of the spin chain. We have to take into account the possibility of complex solutions of the Bethe ansatz equations[6]. While the two magnon solution discussed so far can be regarded as a state of two independent single magnon excitations, the solutions with complex rapidities will clearly show interaction effects.

[6]This was where Bethe's analysis extended that of Bloch.

19.5.2.2 Complex solutions: bound state or spin complex. We now investigate the Bethe ansatz equations (19.23) and (19.24) when the rapidities can take complex values

$$v_i = x_i + iy_i, \quad i = 1, 2.$$

Then we have for the moduli of the Bethe ansatz equations (19.23) and (19.24)

$$\left[\frac{x_1^2 + (y_1 - \frac{1}{2})^2}{x_1^2 + (y_1 + \frac{1}{2})^2}\right]^N = \frac{(x_1 - x_2)^2 + (y_1 - y_2 - 1)^2}{(x_1 - x_2)^2 + (y_1 - y_2 + 1)^2} \tag{19.25}$$

$$\left[\frac{x_2^2 + (y_2 - \frac{1}{2})^2}{x_2^2 + (y_2 + \frac{1}{2})^2}\right]^N = \frac{(x_1 - x_2)^2 + (y_1 - y_2 + 1)^2}{(x_1 - x_2)^2 + (y_1 - y_2 - 1)^2} \tag{19.26}$$

which are symmetric with respect to exchange of the indices 1 and 2.

We examine (19.25) and (19.26) in the limit $N \to \infty$. Assuming that $y_1 > 0$, the LHS of (19.25) vanishes for $N \to \infty$ and moreover it follows that

$$x_1 = x_2 \quad \text{and} \quad y_1 - y_2 = 1.$$

The same result follows from (19.26) assuming $y_2 < 0$. Using the exchange symmetry of the indices, in summary we have either

$$x_1 = x_2 \quad \text{and} \quad y_1 - y_2 = 1 \quad \text{for} \quad y_1 > 0, y_2 < 0 \tag{19.27}$$

or

$$x_1 = x_2 \quad \text{and} \quad y_2 - y_1 = 1 \quad \text{for} \quad y_1 < 0, y_2 > 0. \tag{19.28}$$

Next, we multiply the Bethe ansatz equations (19.23) and (19.24) together. Using (19.27) yields

$$\left[\frac{x_1 + i(y_1 - \frac{3}{2})}{x_1 + i(y_1 + \frac{1}{2})}\right]^N = 1.$$

If we use (19.28) instead, we have to replace y_1 by y_2 in this expression. It follows, that the complex number in the brackets of this expression must have unit modulus. For this to be true, we must either have $y_1 = \frac{1}{2}$ and, hence, $y_2 = -\frac{1}{2}$, or else $y_2 = \frac{1}{2}$ and $y_1 = -\frac{1}{2}$, depending on whether we use (19.27) or (19.28). In both cases, however, we end up with

$$\left[\frac{x_1 - i}{x_1 + i}\right]^N = 1$$

or

$$\exp\left(2i \cot^{-1} x_1\right) = e^{ik}$$

where, again, we have to use the principal value of the $\cot^{-1}$ function and e^{ik} is one of the Nth roots of unity. In the limit $N \to \infty$ both k and $2 \cot^{-1} x_1$ are defined on the same interval $[0, 2\pi]$. Hence, we do not find any restriction for the value of the real part x_1 which can thus assume any real value.

The complex solution of the Bethe ansatz equations (19.23) and (19.24) is thus found to be of the form (dropping the index 1 on the variable x)

$$v_1 = x \pm \frac{i}{2}, \quad v_2 = x \mp \frac{i}{2}, \quad -\infty < x < \infty, \tag{19.29}$$

i.e. they form a pair of conjugate complex values (of course, (19.29) only describes two, not four rapidities).

The important question which remains to be answered is, what kind of state is it that is described by the complex solution (19.29). To answer this question, let's go back from rapidities to quasi–momenta. Recall the relation between the quasi–momenta k_i and the rapidities v_i for the isotropic ferromagnetic spin chain with $\Delta = -1$

$$k_i = 2 \tan^{-1} 2v_i.$$

For the two flipped spins we are considering, we have

$$k_1 \pm k_2 = 2(\tan^{-1} 2v_1 \pm \tan^{-1} v_2)$$

which, with the help of the identity,

$$\tan^{-1} z = \frac{1}{2i} \ln \frac{1+iz}{1-iz}$$

after some algebra, yields

$$k_1 + k_2 = -2 \cot^{-1} x \tag{19.30}$$

and

$$k_1 - k_2 = -i \ln \left[\frac{x^2}{1+x^2} \right]. \tag{19.31}$$

Written differently, this means that k_1 and k_2 are conjugate complex numbers

$$k_1 = \alpha + i\beta, \quad k_2 = \alpha - i\beta$$

where

$$\alpha = -\cot^{-1} x, \quad \beta = -\frac{1}{2} \ln \left[\frac{x^2}{1+x^2} \right].$$

For pairs of conjugate complex quasi–momenta, it has already been shown by Bethe (1931) that the probability $|a(n_1, n_2)|^2$ calculated from the Bethe ansatz wave function falls off roughly exponentially for increasing distance $n_2 - n_1$ of the two flipped spins. The two flipped spins are thus in a bound state, or, as Bethe himself called it, they form a spin complex. We leave the details of this calculation as an exercise.

Exercise 19.7 Bound state of the $M = 2$ sector of the isotropic ferromagnetic spin chain
Calculate the probability $|a(n_1, n_2)|^2$ from the Bethe ansatz wave function for increasing distance $n_2 - n_1$ of the two flipped spins. This exercise is rather an enticement to read Bethe's original paper, either in the German original (Bethe, 1931) or in the English translation, available in Bethe (1997) or Mattis (1993).

It is also instructive to calculate the energy E_2^b of the bound state which is what we shall do now. Recalling the results of section 19.3, especially of exercise 19.5, the energy for two flipped spins of the isotropic ferromagnetic chain is

$$E_2 = \frac{1}{2}\sum_{i=1}^{2}\frac{1}{v_i^2+\frac{1}{4}} = \frac{1}{2}\sum_{i=1}^{2}\frac{dk_i}{dv_i}.$$

For the complex rapidities (19.29) this yields

$$E_2^b = \frac{1}{x^2+1} = \frac{1}{2}\left(1-\cos(k_1+k_2)\right),$$

the last expression following from equation (19.30).

The bound state energy and the two–magnon energy are plotted in figures 19.6 and 19.7.

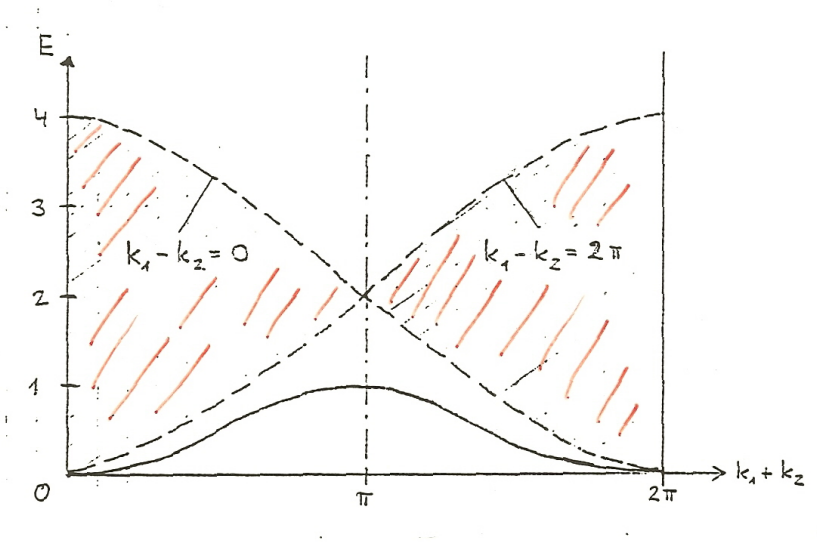

Fig. 19.6 Excitation spectrum of the isotropic ferromagnetic ($\Delta = -1$) spin chain for $M = 2$. Red shaded area: continuum of two–magnon excitations. Full line: bound state energy.

It is apparent from the figure and the equations, too, that E_2^b is always above the ground state energy E_0: $E_2^b(k_1+k_2) \geq E_0$, but also always below the two–magnon excitation energy: $E_2^b(k_1+k_2) \leq E_2^S(k_1+k_2)$.

It seems plausible to assume that the above results can be generalized to three and more flipped spins. However, the algebra will also become more tedious. Since, moreover, we do not expect to gain any further physical insight, just more complicated excitation spectra, we will not pursue this further. Let's close this section with a description what we would be finding in the case $M = 3$ of three flipped spins from the totally ferromagnetic state.

In the case $M = 3$, the following types of solution occur:

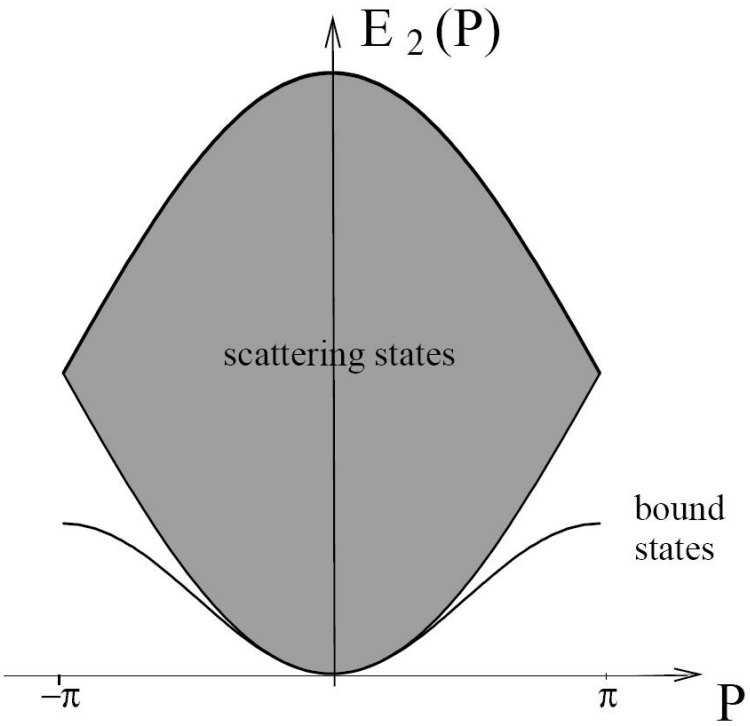

Fig. 19.7 Spectrum of the two magnon excitations of the isotropic ferromagnet. It consists of a continuum of scattering states described by real solutions of the Bethe ansatz equations and a branch of bound states corresponding to two–string solutions. Note, that compared to the calculations in the text, a different Brillouin zone has been chosen: $-\pi \leq P \leq \pi$ instead of $0 \leq (k_1 + k_2) \leq 2\pi$. Otherwise, this is the same spectrum as shown in figure 19.6.

1. all three v_i are real and independent; this solution characterizes the three–magnon state, three independent magnons with quasi-momenta k_1, k_2 and k_3;

2. one of the v_i, v_1 say, is real, the other two form a conjugate complex pair; this corresponds to a one–magnon state with real quasi–momentum k_1 and a bound state or two–spin complex with real quasi–momentum $k = k_2 + k_3$;

3. the three rapidities v_i have the form $v_1 = x + i$, $v_2 = x$, and $v_3 = x - i$ with a common real part x; this is a three-spin complex, i.e. a bound state of the three flipped spins, with the real quasi–momentum $k = k_1 + k_2 + k_3$.

Despite the slightly disparaging remark above about the physical interest in doing the corresponding calculations, it is nevertheless important to generalize the emerging picture of how the states are characterized by rapidities and quasi–momenta. One reason for this is, that otherwise we cannot answer the question which started Bethe himself in his investigation, the question of correctly enumerating *all* states of the spin chain. A second reason is connected

with the extension of the Bethe ansatz to finite temperature: the thermodynamic Bethe ansatz[7], where knowledge of the structure of the Bethe ansatz solutions is the starting point of the analysis; see part VII.

Let's therefore continue with generalizing from what we learned in the cases $M = 1, 2, 3$ for the structure of the solutions of the Bethe ansatz equations.

The solutions of the Bethe ansatz equations are sometimes also called Bethe ansatz roots. From here on, we will be using both terminologies.

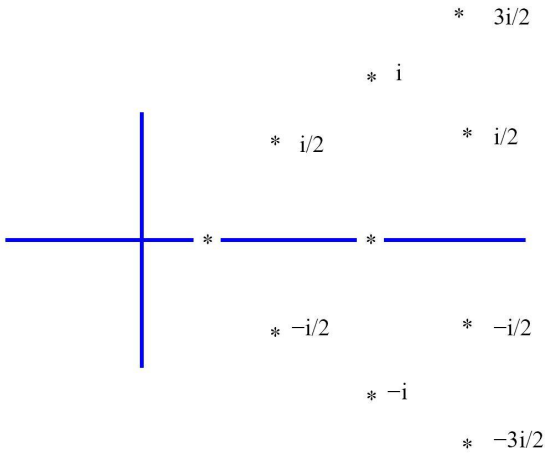

Fig. 19.8 Strings: a one–, two–, three–, and four–string.

19.5.3 String hypothesis

We would now like to generalize from what we have seen in the previous sections. The generalization of the picture of Bethe ansatz solutions which emerged for the case $M = 3$ suggests the following hypothesis. The solutions of the Bethe ansatz equations in product form

$$\left[\frac{v_i - \frac{i}{2}}{v_i + \frac{i}{2}}\right]^N = \prod_{j=1, J\neq i}^{M} \frac{v_i - v_j - i}{v_i - v_j + i}, \quad i = 1, \ldots, M, \quad M \leq \frac{N}{2} \tag{19.32}$$

for arbitrary M contains always spin complexes of various lengths n. An n–spin complex consists, as the examples we discussed in detail suggest, of an aggregate of n complex rapidities v_i, which all have a common real part and whose imaginary parts differ by unity. The spin complex is thus described by n complex numbers

$$v^{n,j} = v^n + \frac{i}{2}(n + 1 - 2j), \quad j = 1, 2, \ldots, n \tag{19.33}$$

[7]Not to be confused with the limit of large N (formally $N \to \infty$), the thermodynamic limit, which as we have seen is not concerned with finite temperature.

where v^n is real. This aggregate is called a *string*[8] of length n. The hypothesis states that this is the general structure of the solutions of the Bethe ansatz equations[9]. We shall see below how the hypothesis simplifies the analysis of the Bethe ansatz equations.

This hypothesis goes basically back to Bethe himself. It was extended to the anisotropic case $\Delta \neq \pm 1$. However, it is not correct as has been shown, e.g. in Woynarovich (1982*a*, Woynarovich (1982*b*). There are indeed excitations of the Heisenberg spin chain which do not obey the string hypothesis. However, the standard argument runs, these non–string excitations are not relevant in the thermodynamic limit because their number is small (i.e. remain microscopic of order unity) compared to the number of string excitations (which become macroscopic of order $\propto N$). Hence, they also do not, as we shall use in part VII, contribute to the thermodynamics of the spin chain. We are not going to discuss the issue of string and non–string solutions further, but adopt the string hypothesis for the most part.

The examples we discussed in detail above do indeed fit the string hypothesis. A one–magnon state is described by a string of length $n = 1$, which is of course a rather fancy way of saying that it is described by a real rapidity. A two–spin complex, a bound state of two flipped spins, is described by a string of length $n = 2$, i.e. a pair of conjugate complex rapidities.

The string structure (19.33) is not yet the most general. We can have several, ν_n say, strings of the same length n which are distinguished by their real parts, i.e.

$$v_\alpha^{n,j} = v_\alpha^n + \frac{i}{2}(n+1-2j), \quad j = 1, 2, \ldots, n \quad \alpha = 1, 2, \ldots, \nu_n \tag{19.34}$$

where ν_n is the number of strings of length n. The index i for a rapidity in equation (19.32), thus, has been replaced by a combination of two indices $j = 1, \ldots, n$, enumerating the string, and $\alpha = 1, 2, \ldots, \nu_n$ of strings of length n with distinct real parts.

Therefore, as it is easy to convince oneself, the numbers ν_n are restricted by the condition

$$M = \sum_{n=1}^{M} n\nu_n.$$

In particular, this can easily be verified for the case $M = 3$ which we discussed in the previous section. Indeed, for $M = 3$ we have the following possibilities

1. $\nu_1 = 3, \nu_2 = \nu_3 = 0$
2. $\nu_1 = \nu_2 = 1, \nu_3 = 0$
3. $\nu_1 = \nu_2 = 0, \nu_3 = 1$

which covers the cases we have discussed.

19.5.3.1 Bethe ansatz equations for the real parts of strings. We now examine the Bethe ansatz equations in product form

$$\left[\frac{v_i - \frac{i}{2}}{v_i + \frac{i}{2}}\right]^N = \prod_{j=1, j\neq i}^{M} \frac{v_i - v_j - i}{v_i - v_j + i}$$

[8] No connection to string theory!

[9] As we have seen in the examples above, the strings take the form (19.33), strictly speaking, only in the thermodynamic limit $N \to \infty$. For finite N, there are small corrections which vanish rapidly as the chain length N increases; see e.g. Faddeev and Takhtajan (1984) and Gaudin (1983) and references therein for details. We ignore this fine detail here.

inserting the string form for the rapidities v_i and v_j

$$v_i \to v_\alpha^{n,k} = v_\alpha^n + \frac{i}{2}(n+1-2k), \quad k = 1, 2, \ldots, n, \quad \alpha = 1, 2, \ldots, \nu_n, \tag{19.35}$$

$$v_j \to v_\beta^{m,l} = v_\beta^m + \frac{i}{2}(m+1-2l), \quad l = 1, 2, \ldots, m, \quad \beta = 1, 2, \ldots, \nu_m, \tag{19.36}$$

which yields the equations in the form

$$\begin{aligned}\left[V_1(v_\alpha^{n,k})\right]^N &= \prod_{\substack{m,\beta \\ \{m,\beta\}\neq\{n,\alpha\}}} \prod_{l=1}^{m} \frac{v_\alpha^{n,k} - v_\beta^{m,l} - i}{v_\alpha^{n,k} - v_\beta^{m,l} + i} \\ &= \prod_{\substack{m,\beta \\ \{m,\beta\}\neq\{n,\alpha\}}} \prod_{l=1}^{m} V_2(v_\alpha^{n,k} - v_\beta^{m,l})\end{aligned}$$

where

$$V_1(a) = \frac{a - \frac{i}{2}}{a + \frac{i}{2}},$$

and

$$V_2(a) = \frac{a-i}{a+i}.$$

In the double product above: $m = 1, 2, \ldots, M$ and $\beta = 1, 2, \ldots, \nu_m$. The notation for the double product is used to indicate that, on the LHS, all terms are to be excluded where $\{m, \beta\} \neq \{n, \alpha\}$ would be fulfilled simultaneously. However, one must allow terms with $m = n$ as long as $\beta \neq \alpha$ and terms with $\beta = \alpha$ as long as $m \neq n$. Since $\{m, \beta\} \neq \{n, \alpha\}$ guarantees that we are either dealing with different real parts or strings of different length (i. e. different strings), the third product does not need a restriction.

For later use, we define also the functions for general index n

$$V_n(a) = \frac{a - \frac{n}{2}i}{a + \frac{n}{2}i}.$$

The inner product over l can be easily evaluated. First, using (19.36), we obtain

$$\prod_{l=1}^{m} \frac{v_\alpha^{n,k} - v_\beta^{m,l} - i}{v_\alpha^{n,k} - v_\beta^{m,l} + i} = \prod_{l=1}^{m} \frac{v + i(l-1)}{v + i(l+1)}$$

where we set

$$v \equiv v_\alpha^{n,k} - v_\beta^m - i\frac{m+1}{2}$$

as abbreviation. In this form the product is easily seen to be "telescopic", i.e. most factors in the numerator and denominator of the product cancel and what remains is

$$\prod_{l=1}^{m} \frac{v + i(l-1)}{v + i(l+1)} = \frac{v}{v + i(m+1)} \frac{v+i}{v+im}.$$

The remaining two quotients on the RHS can be written with the help of the functions V_1 or V_n we introduced above

$$\begin{aligned}\frac{v}{v+i(m+1)}\frac{v+i}{v+im} &= V_1\left(\frac{v_\alpha^{n,k}-v_\beta^m}{m+1}\right)V_1\left(\frac{v_\alpha^{n,k}-v_\beta^m}{m-1}\right)\\ &= V_{m+1}(v_\alpha^{n,k}-v_\beta^m)V_{m-1}(v_\alpha^{n,k}-v_\beta^m),\end{aligned}$$

so that we get

$$\begin{aligned}\left[V_1(v_\alpha^{n,k})\right]^N &= \prod_{m=1,m\neq n}^{M}\prod_{\beta=1,\beta\neq\alpha}^{\nu_n} V_1\left(\frac{v_\alpha^{n,k}-v_\beta^m}{m+1}\right)V_1\left(\frac{v_\alpha^{n,k}-v_\beta^m}{m-1}\right)\\ &= \prod_{m=1,m\neq n}^{M}\prod_{\beta=1,\beta\neq\alpha}^{\nu_n} V_{m+1}(v_\alpha^{n,k}-v_\beta^m)V_{m-1}(v_\alpha^{n,k}-v_\beta^m).\end{aligned}$$

Let's now perform the product over $k \quad (k=1,\ldots,n)$ on both sides of this equation. The LHS contains again a telescopic product, so that the LHS reduces to

$$\prod_{k=1}^{n}\left[V_1(v_\alpha^{n,k})\right]^N = \left[V_1\left(\frac{v_\alpha^n}{n}\right)\right]^N = \left[V_n\left(v_\alpha^n\right)\right]^N,$$

and, thus, we are left with only the real parts of the string on the LHS. This yields the compact form of the Bethe ansatz equations

$$\left[V_n\left(v_\alpha^n\right)\right]^N = \prod_{m=1,m\neq n}^{M}\prod_{\beta=1,\beta\neq\alpha}^{\nu_n} V_{nm}(v_\alpha^n-v_\beta^m)$$

where

$$\begin{aligned}V_{nm}(v_\alpha^n-v_\beta^m) &= \prod_{k=1}^{n} V_1\left(\frac{v_\alpha^{n,k}-v_\beta^m}{m+1}\right)V_1\left(\frac{v_\alpha^{n,k}-v_\beta^m}{m-1}\right)\\ &= \prod_{k=1}^{n} V_{m+1}(v_\alpha^{n,k}-v_\beta^m)V_{m-1}(v_\alpha^{n,k}-v_\beta^m). \qquad (19.37)\end{aligned}$$

Still, we have not completely reduced the Bethe ansatz equations such that they contain only the real parts of the strings, which is, after all, our objective. The RHS still contains the full complex strings, equation (19.35). These can also be removed, with somewhat more effort, by inserting (19.35) into (19.37). It can then be shown that the Bethe ansatz equations can finally be recast in a form containing only the real parts of the (assumed) strings

$$\left[V_n(v_\alpha^n)\right]^N = \prod_{m=1,m\neq n}^{M}\prod_{\beta=1,\beta\neq\alpha}^{\nu_n} V_{nm}(v_\alpha^n-v_\beta^m), \qquad (19.38)$$

where the functions V_{nm} now contain only the real centers of the strings, and are explicitly given by

$$V_{nm}(x) = V_{|n-m|}(x)V^2_{|n-m|+2}(x)\dots V^2_{n+m-2}V_{n+m}(x),$$

and

$$x = v^n_\alpha - v^m_\beta.$$

Now we have arrived at a set of equations for the real parts of the strings only which has been the goal of this maybe slightly tedious section.

However, we can obtain a further important result with almost no further work: the Bethe ansatz equations for the real parts of the strings are valid, not only for the isotropic ferromagnetic case $\Delta = -1$, for which we just derived them, but also for the isotropic antiferromagnetic case $\Delta = 1$. The Bethe ansatz equations (19.15) for the isotropic antiferomagnetic spin chain transform into the corresponding equations (19.22) for the ferromagnetic chain upon complex–conjugating, however for the rapidities v_i^* and v_j^* instead of v_i and v_j. This does not matter, though, since a look at the string structure equations (19.35) and (19.36) reveals that $\left(v^{n,j}_\alpha\right)^*$ and $(v^{m,l}_\beta)^*$ are the same set of complex numbers as (19.35) and (19.36) themselves. Thus, the Bethe ansatz equations (19.38) are generally valid for the isotropic spin chain $\Delta = \pm 1$.

In a first attempt to use the work invested, let us calculate physical quantities: crystal quasi–momentum and energy of a string solution.

In part VII we shall also see, that the tedious work to derive the Bethe ansatz equations for the real parts of the strings was necessary to correctly investigate the finite temperature thermodynamics of the spin chain.

19.5.3.2 Quasi–momentum and energy of a string solution. Starting with the ferromagnetic case $\Delta = -1$, we get for the crystal quasi–momentum from (19.11)

$$K = \sum_{j=1}^{M} k_j$$

and a result from exercise 19.5

$$k_j = 2\tan^{-1} 2v_j = \frac{1}{i}\ln\frac{\frac{i}{2} - v_j}{\frac{i}{2} + v_j}$$

the intermediate result

$$K = \sum_{j=1}^{M} k_j = \frac{1}{i}\prod_{j=1}^{M}\frac{\frac{i}{2} - v_j}{\frac{i}{2} + v_j}.$$

To make progress, we shall assume that all rapidities $v_j \quad (j = 1, \dots, M)$ belong to a single string of length M, i.e. to an M–string complex, thus

$$v_j = v + \frac{i}{2}(M + 1 - 2j), \quad j = 1, \dots, M.$$

Again the product is telescopic and we find

$$\prod_{j=1}^{M}\frac{\frac{i}{2} - v_j}{\frac{i}{2} + v_j} = (-1)^M V_1\left(\frac{v}{M}\right),$$

and thus

$$K = 2\tan^{-1}\frac{2v}{M} - \frac{\pi}{2}[1 + (-1)^{M+1}].$$

The energy is then from

$$E_M(K) = \frac{1}{2}\sum_{i=1}^{M}\frac{dk_i}{dv_i} = \frac{1}{2}\frac{dK}{dv}$$

given by

$$E_M(K) = \frac{1}{M}[1 - (-1)^M \cos K],$$

which coincides with the special cases for $M = 1$ and $M = 2$ which we calculated above.

Using the corresponding equations for the isotropic antiferromagnetic spin chain $\Delta = 1$ (cf. 19.3.1), a similar calculation yields the results

$$K = 2\cot^{-1}\frac{2v}{M}$$

for the total crystal quasi–momentum, and

$$E_M(K) = -\frac{1}{M}(1 - \cos K)$$

for the energy.

Similar calculations can be performed for the case when there are several string solutions. We leave the example of a two string solution as an exercise to the reader.

Exercise 19.8 Quasi–momentum and energy of a two string solution
Repeat the calculations of section 19.5.3.2 leading to quasi–momentum and energy of a one string solution of the Bethe ansatz equations for a two string solution,i.e. for two strings of length $M - n$ and n, respectively. In this case, the total quasi–momentum splits into two parts

$$K = P + Q$$

one part, P say, corresponding to the string of length $M - n$, the other, Q, to the string of length n.

Exercise 19.9 Ferromagnetic versus antiferromagnetic spin chain Show that for fixed M and fixed K the M–spin complex (one string of length M) the excitation energy of the ferromagnetic case is lowest, whereas the excitation energy of the antiferromagnetic case is highest.

29

Thermodynamics of the isotropic spin chain

29.1 Summary of Bethe ansatz for the isotropic spin chain

Let us first summarize some of the results we obtained in the previous part IV on the coordinate Bethe ansatz for the isotropic antiferromagnetic spin chain. To be able to calculate quantities such as the magnetic susceptibility, we consider in this part on the thermodynamic Bethe ansatz the slightly more general isotropic antiferromagnetic spin chain in an external magnetic field $h \geq 0$ with Hamiltonian

$$\mathcal{H} = \sum_{n=1}^{N} \left(\mathbf{S}_n \mathbf{S}_{n+1} - \frac{1}{4} \right) - h \sum_{n=1}^{N} S_n^z. \tag{29.1}$$

Exercise 29.1 Energy eigenvalue of the isotropic spin chain in a magnetic field
Convince yourself that the earlier result for the energy eigenvalue of the isotropic antiferromagnetic chain, equation (19.16), in section 19.3.1, can easily be generalized to the finite magnetic field case of Hamiltonian (29.1) to give

$$E_M = -\frac{1}{2} \sum_{i=1}^{M} \frac{1}{v_i^2 + \frac{1}{4}} - h \left(\frac{N}{2} - M \right). \tag{29.2}$$

The rapidities $\{v_i\}$ satisfy the Bethe ansatz equations

$$\left(\frac{v_i + \frac{i}{2}}{v_i - \frac{i}{2}} \right)^N = \prod_{j=i, j \neq i}^{M} \frac{v_i - v_j + i}{v_i - v_j - i}, \quad i = 1, 2, \ldots, M, \quad M \leq \frac{N}{2}. \tag{29.3}$$

These equations written in a more compact notation are

$$[V_1(v_i)]^N = \prod_{j=i, j \neq i}^{M} V_2(v_i - v_j),$$

where the functions

$$V_n(v) \equiv \frac{v + \frac{in}{2}}{v - \frac{in}{2}}$$

have been introduced.

We have discussed in section 19.5.3 the string hypothesis for the solution of these equations. Let us recall the first few cases of string solutions:

- a string of length one:
$$v^{(1)} = x$$
with x an arbitrary real number, called the *center* of the string.
- a string of length two:
$$\begin{cases} v^{(1)} = x + \frac{i}{2} \\ v^{(2)} = x - \frac{i}{2} \end{cases},$$
- a string of length three:
$$\begin{cases} v^{(1)} = x + i \\ v^{(2)} = x \\ v^{(3)} = x - i \end{cases},$$
- and so on. Cf. figure 29.1

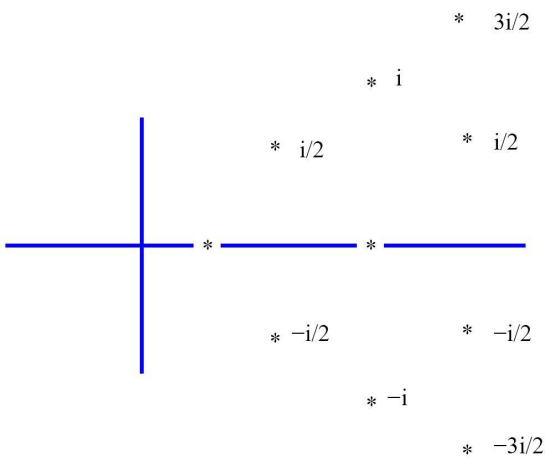

Fig. 29.1 Strings: a one–, two–, three–, and four–string.

Recall the *string hypothesis* which states that all solutions $\{v_i, i = 1, 2, \ldots, M\}$ of the Bethe ansatz equations (29.3) are organized in strings. This hypothesis, despite the fact that it is not correct (cf. the references given in section 19.5.3), is adopted in most treatments of the thermodynamics of Bethe ansatz solvable models, and we shall adopt it here too. Non–string type solutions are assumed to be negligible in the thermodynamic limit.

The general form of the strings is as follows: There are in general M_n strings of length n of the form
$$v_\alpha^{(n,j)} = v_\alpha^n + i\left(\frac{n+1}{2} - j\right),$$
where
$$j = 1, \ldots, n, \quad \alpha = 0, 1, \ldots, M_n \quad n = 1, \ldots, \infty$$
with real centers v_α^n.

We have discussed in section 19.5.3 how to rewrite the Bethe ansatz equations for string solutions $\{v_\alpha^{(n,j)}\}$ in such a way that the resulting Bethe ansatz equations only contain the real centers, $\{v_\alpha^n\}$, of the strings. The result there has been

$$[V_n\left(v_\alpha^n\right)]^N = \prod_{m=1;m\neq n}^{M} \prod_{\beta=0;\beta\neq\alpha}^{\nu_m} V_{nm}\left(v_\alpha^n - v_\beta^m\right) \tag{29.4}$$

$$V_{nm}(x) = V_{|n-m|}(x)V_{|n-m|+2}^2(x)\ldots V_{n+m-2}^2(x)V_{n+m}(x).$$

Since we have the thermodynamic limit $N \to \infty$ in mind, we can let the upper limit of the first product over m grow beyond bounds $M \to \infty$.

Since the products in equation (29.4) only contain phases, taking the logarithm simplifies matters

$$Nk_n(v_\alpha^n) = 2\pi J_\alpha^n + \sum_{m=1,m\neq n}^{\infty} \sum_{\beta=0,\beta\neq\alpha}^{\nu_m} \Theta_{nm}\left(v_\alpha^n - v_\beta^m\right) \tag{29.5}$$

where

$$k_n(v) \equiv i\ln V_n(v)$$

and

$$\Theta_{nm}(v) \equiv i\ln V_{nm}(v).$$

Note, that we have already used the upper limit ∞ in the summation over m.

The numbers J_α^n play the role of quantum numbers and are integers or half–integers which are constrained by

$$-J_{\max}^n \leq J_\alpha^n \leq J_{\max}^n.$$

Faddeev and Takhtajan (1984) have given an explicit description for extracting an expression for $J_{\max}^n$ from equations (29.5). However, we shall not need such an expression for our purposes.

We only need that such a range in fact does exist and call any integer or half–integer within the range *admissible*. The reason for this naming will become clear soon. For every admissible set of mutually distinct quantum numbers $\{J_\alpha^n\}$ there is a unique solution of, again mutually distinct, values $\{v_\alpha^n\}$, of equation (29.5). Such solutions describe "particles" in the Bethe ansatz framework, and are, hence, called particle rapidities.

Yang and Yang (1969), in their treatment of the thermodynamics of the one–dimensional Bose gas with repulsive delta–function interaction, introduced the corresponding concept of "holes". If we prescribe a set of particle quantum numbers $\{J_\alpha^n\}$, there can be, and in a generic case will be, gaps which correspond to numbers $\{\tilde{J}_\alpha^n\}$, which, though perfectly admissible, are *not* in the set of particle quantum numbers $\{J_\alpha^n\}$.

Using these admissible quantum numbers, which we call hole quantum numbers, we define via

$$Nk_n(\tilde{v}_\alpha^n) = 2\pi\tilde{J}_\alpha^n + \sum_{m=1,m\neq n}^{\infty} \sum_{\beta=0,\beta\neq\alpha}^{\nu_m} \Theta_{nm}\left(\tilde{v}_\alpha^n - v_\beta^m\right) \tag{29.6}$$

quantities $\{\tilde{v}_\alpha^n\}$, called hole rapidities.

Observe, that the hole rapidities are determined by the particle rapidities $\{v_\beta^m\}$, and, thus, equations (29.5) and (29.6) are coupled.

34
Euler–Maclaurin formula

Read Euler: he is our master in everything.

Pierre–Simon Laplace, as quoted in George Simmons (1992)

The Bethe ansatz equations contain sums over the finitely many discrete quasi–momenta or rapidities. While these sums have been transformed into integrals in the thermodynamic limit $N \to \infty$

34.1 Bernoulli numbers

The Bernoulli numbers B_n, $n = 0, 1, \ldots$, a sequence of rational numbers, which are of great importance in number theory, appear in many mathematical applications, for example in the Euler–Maclaurin formula which is the subject matter of this chapter. The values of the first fifteen Bernoulli numbers are

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_3 = 0, B_4 = -\frac{1}{30}, B_5 = 0, B_6 = \frac{1}{42}, B_7 = 0,$$
$$B_8 = -\frac{1}{30}, B_9 = 0, B_{10} = \frac{5}{66}, B_{11} = 0, B_{12} = -\frac{691}{2730}, B_{13} = 0, B_{14} = \frac{7}{6}.$$

Apart from the Euler–Maclaurin formula, the Bernoulli numbers e.g. also appear in the Taylor series expansions of the tangent and hyperbolic tangent functions, in formulas for the sum of powers of the first positive integers, and in connection with the Riemann zeta function.

There are various ways to define to Bernoulli numbers. One of these is via the generating function

$$\frac{ze^{xz}}{e^z - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!} z^n, \tag{34.1}$$

introducing the functions $B_n(x)$ with the Bernoulli numbers given by $B_n = B_n(0)$. The functions $B_n(x)$ are the Bernoulli polynomials

$$B_n(x) = \sum_{\nu=0}^{n} \binom{n}{\nu} B_{n-\nu} x^{\nu},$$

defined on the interval $0 \leq x \leq 1$ and continued periodically beyond that interval. We shall, however, not further discuss the Bernoulli polynomials here, but give one property of the derivative

$$B_n'(x) = nB_{n-1}(x) \tag{34.2}$$

and state explicitly the first two Bernoulli polynomials

$$B_0(x) = 1 \quad \text{and} \quad B_1(x) = x - [x] - 1$$

where $[x]$ is the largest integer n with $n \leq x$.

For our limited purpose to derive a few elementary properties of the Bernoulli numbers, it suffices to consider the generating function (34.1) for the special value $x = 0$.

Then we can write

$$z = (e^z - 1) \sum_{n=0}^{\infty} \frac{B_n}{n!} z^n = \left(\sum_{m=1}^{\infty} \frac{z^m}{m!} \right) \left(\sum_{n=0}^{\infty} \frac{B_n}{n!} z^n \right) = \sum_{\mu=1}^{\infty} \frac{z^\mu}{\mu!} \sum_{n=0}^{\mu-1} \binom{\mu}{n} B_n. \quad (34.3)$$

Exercise 34.1 Bernoulli numbers
Fill in the steps in the derivation of the identity (34.3), using the summation formula

$$\left(\sum_{m=0}^{\infty} a_m z^m \right) \left(\sum_{n=0}^{\infty} b_n z^n \right) = \sum_{\mu=0}^{\infty} z^\mu \sum_{m+n=\mu} a_m b_n.$$

This identity obviously implies

$$\sum_{n=0}^{\mu-1} \binom{\mu}{n} B_n = \begin{cases} 1 & \mu = 1 \\ 0 & \mu \geq 2 \end{cases}.$$

This recursion relation can be evaluated by setting successively $\mu = 1$, $\mu = 2$, and so on: We obtain one by one the Bernoulli numbers as given above.

The generating function (34.1) also allows to prove that, as we may have suspected already, all Bernoulli numbers with odd index $n > 1$ vanish: $B_n = 0$ for $n > 1$ odd. Set again $x = 0$ to obtain

$$\frac{z}{e^z - 1} + \frac{z}{2} = \sum_{n=0, n\neq 1}^{\infty} \frac{B_n}{n!} z^n,$$

where we used $B_1 = -1/2$ when shuffling the term for $n = 1$ of the sum to the left hand side of the equation. Obviously the function on the left hand side is an even function of z. Thus we can conclude that $B_{2k+1} = 0$ for $k = 1, 2, \ldots$.

In the next section, the Bernoulli numbers will play a central role in another piece of applied mathematics: The transformation of a finite sum into an integral plus some remainder.

34.2 The Euler–Maclaurin summation formula

For the derivation of this important formula, we consider the following expression for an arbitrary function $f(x)$ and arbitrary integers $m < n$

$$\begin{aligned}\int_m^n dx B_1(x) f'(x) &= \sum_{\nu=m}^{n-1} \int_\nu^{\nu+1} dx \left(x - [x] - \frac{1}{2}\right) f'(x) \\ &= \sum_{\nu=m}^{n-1} \left(\left(x - \nu - \frac{1}{2}\right) f(x) \Big|_\nu^{\nu+1} - \int_\nu^{\nu+1} dx f(x) \right) \\ &= \sum_{\nu=m}^{n} f(\nu) - \frac{1}{2}\left(f(m) + f(n)\right) - \int_m^n dx f(x).\end{aligned}$$

This can be rearranged to give the Euler–Maclaurin summation formula in its simplest form

$$\sum_{\nu=m}^{n} f(\nu) = \int_m^n dx f(x) + \frac{1}{2}\left(f(m) + f(n)\right) + \int_m^n dx B_1(x) f'(x). \tag{34.4}$$

The formula allows to convert a sum into an integral and two correction terms:

$$\begin{aligned}\frac{1}{2}\left(f(m) + f(n)\right) &: \quad \text{from the summation limits and} \\ \int_m^n dx B_1(x) f'(x) &: \quad \text{a remainder term.}\end{aligned}$$

The Euler–Maclaurin summation formula can be further refined by looking more carefully at the remainder term. Recalling the property (34.2) of the Bernoulli polynomials and that for integer values μ and ν of its argument

$$B_n(\mu) = B_n(\nu) = B_n(0) = B_n \qquad (\mu, \nu \geq 2),$$

we can write the remainder term as

$$\int_m^n dx B_1(x) f'(x) = \frac{B_2}{2!}\left(f'(n) - f'(m)\right) - \int_m^n dx \frac{B_2(x)}{2!} f''(x).$$

This procedure can be iterated such that we arrive at the general form of the Euler–Maclaurin summation formula

$$\begin{aligned}\sum_{\nu=m}^{n} f(\nu) = \int_m^n dx f(x) + \frac{1}{2}\left(f(m) + f(n)\right) + \sum_{j=2}^{k} (-1)^j \frac{B_j}{j!} \left(f^{(j-1)}(n) - f^{(j-1)}(m)\right) \\ + \int_m^n dx (-1)^{k+1} \frac{B_k(x)}{k!} f^{(k)}(x).\end{aligned} \tag{34.5}$$

This form of the Euler–Maclaurin summation formula can be rewritten using an estimate of the remaining integral on the right hand side of (34.5) as

$$\begin{aligned}\sum_{\nu=m}^{n} f(\nu) = \int_m^n dx f(x) + \frac{1}{2}\left(f(m) + f(n)\right) \\ + \sum_{j=1}^{k} \frac{B_{2j}}{(2j)!} \left(f^{(2j-1)}(n) - f^{(2j-1)}(m)\right) \\ + \Theta \frac{B_{2(k+1)}}{(2(k+1))!} \left(f^{(2k+1)}(n) - f^{(2k+1)}(m)\right)\end{aligned} \tag{34.6}$$

for some number $0 \leq \Theta \leq 1$.

This result of applied mathematics, the Euler–Maclaurin summation formula in the form (34.6), provides the first ingredient in needed for the finite–size analysis of the Bethe ansatz equations.

We now turn to the second ingredient which is again provided by applied mathematics, the Wiener–Hopf technique to solve functional equations.

35
Wiener–Hopf technique

The shortest path between two truths in the real domain passes through the complex domain.

Jacques Hadamard

35.1 General method

The Wiener–Hopf method is a mathematical technique for treating functional equations. These equations arise, for example, from the study of certain classes of integral equations, which is, in view of what we discussed at various places and, in particular in the previous chapter, why we are interested in the Wiener–Hopf technique here.

We consider, in this section, integral equations of the type

$$f(x) = \phi(x) + \int_0^\infty dy\, K(x-y) f(y) \quad -\infty < x < \infty, \tag{35.1}$$

where $\phi(x)$ is a given function, and the kernel K of the integral equation has the difference property

$$K(x,y) = K(x-y).$$

This type of integral equation cannot be readily solved by Fourier transformation because of the integral limits which are

$$\int_0^\infty dy$$

i.e. restricted to the real half line $[0,\infty)$.

The strategy to solve this type of integral equation consists in, formally, extending the range of integration to the whole real line $(-\infty,\infty)$. To do this, we split the unknown function $f(x)$ formally into a positive and a negative part by defining

$$f(x) \equiv f_+(x) + f_-(x),$$

where

$$f_+(x) \equiv \begin{cases} 0 & \text{for} \quad x > 0 \\ f(x) & \text{for} \quad x < 0 \end{cases},$$

and

$$f_-(x) \equiv \begin{cases} f(x) & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x < 0 \end{cases}.$$

The integral equation 35.1 then becomes

$$f_+(x) + f_-(x) = \phi(x) + \int_{-\infty}^{\infty} dy\, K(x-y) f_-(y). \tag{35.2}$$

Now that we formally integrate over the whole real line, we can use Fourier transformation. First, we introduce the following conventions[1] for Fourier transforms

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk\, \hat{f}(k) e^{ikx}, \tag{35.3}$$

and

$$\hat{f}(k) = \int_{-\infty}^{\infty} dx\, f(x) e^{-ikx}. \tag{35.4}$$

The integral equation 35.2, upon Fourier transformation, becomes a functional equation for the Fourier transform $\hat{f}(k)$[2]

$$\hat{f}_+(k) + \hat{f}_-(k) = \hat{\phi}(k) + \hat{K}(k)\hat{f}_-(k)$$

or

$$\hat{f}_-(k)\left(1 - \hat{K}(k)\right) + \hat{f}_+(k) = \hat{\phi}(k) \tag{35.5}$$

The problem, obviously, is how to solve this equation which contains two unknown functions. It will turn out, that this is possible if we investigate the analytic properties in the complex plane of the functions involved in the functional equation, and provided that they obey certain conditions. In other words, we invoke the powers of complex analysis.

First, we observe that $\hat{K}(k)$ is an analytic function in the strip

$$-\beta < \Im(k) < \alpha$$

of the complex plane, if

$$|K(x)| \leq \begin{cases} e^{-\alpha x} & \text{for} \quad x > 0 \\ e^{\beta x} & \text{for} \quad x < 0 \end{cases} \quad \text{with} \quad \alpha, \beta > 0.$$

The proof of this statement is left as an easy exercise 35.1.

Exercise 35.1 Analytic properties of $\hat{K}(k)$
Proof the statement about the analyticity of the Fourier transformed kernel $\hat{K}(k)$ in the strip $-\beta < \Im(k) < \alpha$. Use

$$k = \Re k + i\Im k = k' + ik''$$

and split the Fourier transform of $K(x)$, i.e. $\hat{K}(k)$ into two parts, one for $0 < x < \infty$ and the other for $-\infty < x < 0$.

[1] **Attention**: It may well be that there are different conventions in different parts of these lecture notes. Need to harmonize this eventually.

[2] Recall that Wiener–Hopf is really a technique to solve functional eqautions.

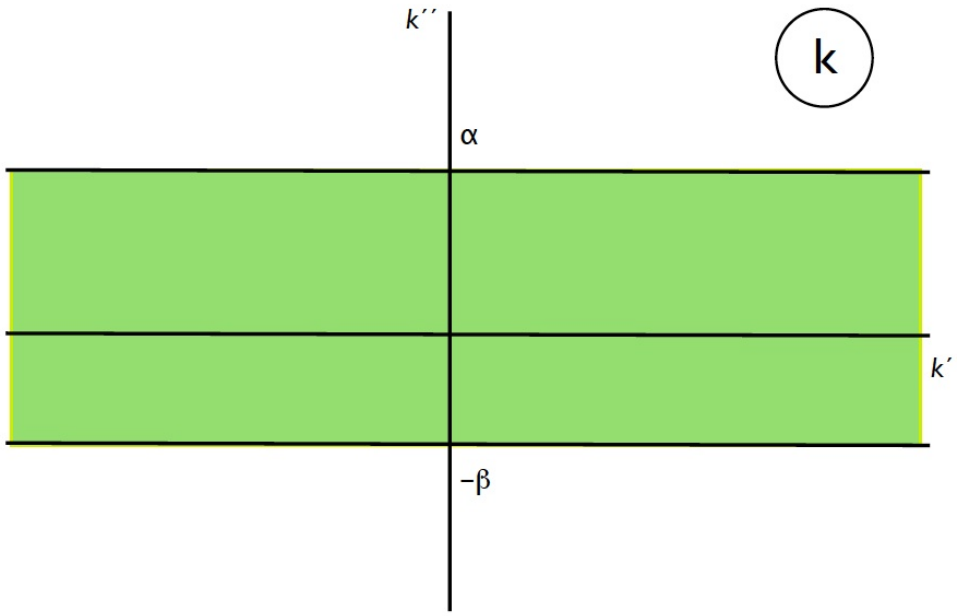

Fig. 35.1 Strip of the complex k–plane where $\hat{K}(k)$ is analytic.

We now make the further assumption that $\hat{f}(k)$ is an analytic function in, at least, a part of the strip $-\beta < \Im(k) < \alpha$, e.g. in $-\gamma < \Im(k) < \delta$ with $\gamma \leq \beta$ and $\delta \leq \alpha$.

To have a more symmetrical functional equation, we define another two complex functions $A(k)$ and $B(k)$ by

$$\frac{A(k)}{B(k)} = 1 - \hat{K}(k),$$

requiring that

$A(k)$ is analytic for $\Im k < \delta$, and
$B(k)$ is analytic for $\Im k > \gamma$.

In other words, we require that in

$$\ln(1 - \hat{K}(k)) = \ln A(k) - \ln B(k)$$

all terms are regular functions.

The functional equation now takes the more symmetrical form

$$A(k)\hat{f}_-(k) + B(k)\hat{f}_+(k) = \hat{\phi}(k)B(k) \equiv C(k) + D(k),$$

where the function $C(k)$ is analytic in the half–plane $\Im k > \gamma$, and the function $D(k)$ is analytic in the half–plane $\Im k < \delta$.

It is important to notice, that all functions in the functional equation have at least the common strip of analyticity $-\gamma < \Im k < \delta$.

Rewriting the functional equation once again, we obtain

$$B(k)\hat{f}_+(k) - C(k) = -A(k)\hat{f}_-(k) + D(k),$$

where now the LHS is analytic in the half–plane $\Im k > \gamma$, whereas the RHS is analytic in the half–plane $\Im k < \delta$. Therefore both sides of the functional equation must be equal to an entire function. It can be shown that this entire function is a polynomial; cf., e.g. Roos (1969) for details.

We stop at this point, illustrating the Wiener–Hopf technique with an example.

35.1.1 Example of the Wiener–Hopf technique

To see the Wiener–Hopf technique at work, consider the integral equation

$$f(x) = e^{-|x|} + \lambda \int_0^\infty dy\, e^{|x-y|} f(y).$$

Here, the inhomogeneous function $\phi(x) = e^{-|x|}$ and the Kernel $K(x-y) = \lambda e^{-|x-y|}$ are essentially the same function, whose Fourier transform is

$$\int_{-\infty}^{\infty} dx\, e^{-|x|} e^{ikx} = \int_0^\infty dx\, e^{-x(ik+1)} \int_{-\infty}^0 dx\, e^{-x(ik-1)} = \frac{2}{k^2+1}$$

which leads to the functional equation

$$f_+(k) + f_-(k) = \frac{2}{k^2+1} + \frac{2\lambda}{k^2+1} f_-(k)$$

or

$$f_+(k) + \frac{k^2 - 2\lambda + 1}{k^2+1} f_-(k) = f_+(k) + \frac{k^2-\xi^2}{k^2+1} f_-(k) = \frac{2}{k^2+1} \tag{35.6}$$

where we assume that ξ with $\xi^2 = 2\lambda - 1$ is real, i.e. $\lambda > \frac{1}{2}$.

Since

$$1 - \hat{K}(k) = \frac{k^2-\xi^2}{k^2+1}$$

has singularities at $k = \pm i$, the strips where it is analytic are

$$-1 < k'' < 0 \quad \text{or} \quad 0 < k'' < 1.$$

Let us concentrate on the strip $-1 < k'' < 0$. The decomposition in this strip yields

$$1 - \hat{K}(k) = \frac{k^2-\xi^2}{k^2+1} = \frac{A(k)}{B(k)} = \frac{k^2-\xi^2}{k-i} \frac{1}{k+i}$$

i.e.

$$A(k) = \frac{k^2-\xi^2}{k-i}$$
$$B(k) = k+i$$

where now $A(k)$ is analytic for $k'' < 0$ and $B(k)$ is analytic for $k'' > -1$.

Now we can write the functional equation as

$$(k+i)f_+(k) + \frac{k^2-\xi^2}{k-i} f_-(k) = \frac{2}{k-i}$$

or

$$(k+i)f_+(k) = -\frac{k^2-\xi^2}{k-i} f_-(k) + \frac{2}{k-i} = E(k)$$

where the left hand side is analytic for $k'' > -1$ while the right hand side is analytic for $k'' < 0$ and $E(k)$ is an entire function for which the simplest choice is a constant which, for convenience, we choose to be $E(k) = ia$.Thus

$$(k+i)f_+(k) = ia \tag{35.7}$$

$$\frac{k^2-\xi^2}{k-i} f_-(k) = \frac{2}{k-i} - ia \tag{35.8}$$

The functions $f_+(x)$ and $f_-(x)$ can now be calculated from (35.7) and (35.8) by Fourier transformation. The results are

$$f_+(x) = ae^x$$

$$f_-(x) = -\frac{2}{\xi}\sin\xi x + a\left(\cos\xi x + \frac{1}{\xi}\sin\xi x\right)$$

Besides the reference to Roos (1969) already mentioned, a short introduction to the Wiener–Hopf technique applied to integral equations including the example can be found in Mathews and Walker (1970) and also in Wyld (1999).

36
Finite Heisenberg quantum spin chain

[. . .] I knew from experience how many sheets of paper go into the waste–paper basket after even a modest calculation: there was no way they could all appear in print.

Rodney Baxter Baxter (1982)

36.1 Bethe ansatz for the finite Heisenberg Spin chain

The first systematic method to calculate finite–size corrections for integrable or Bethe ansatz systems is due to de Vega and Woynarovich (1985). The method deals with systems with a non–vanishing energy gap, e.g. the anisotropic Heisenberg spin chain.

The Hamiltonian of the anisotropic Heisenberg spin chain for an even number N of spins $\frac{1}{2}$ in the form

$$\mathcal{H} = \sum_{n=1}^{N} \left(\sigma_n^x \sigma_{n+1}^x + \sigma_n^y \sigma_{n+1}^y + \sigma_n^z \sigma_{n+1}^z \right) \tag{36.1}$$

describes the easy–axis antiferromagnetic regime for anisotropy $\Delta = \cosh \gamma > 1$. In this regime, which de Vega and Woynarovich (1985) considered, there is a finite energy gap in the excitation spectrum even in the thermodynamic limit $N \to \infty$.

We shall not pursue the analysis for the gapped system, which involves complex solutions of the Bethe ansatz equations (see de Vega and Woynarovich (1985) for details about gapped systems). Instead we concentrate on the Heisenberg spin chain with anisotropy in the range $|\Delta| < 1$ where there is no gap, i.e. the system is critical. In this anisotropy range and for low the lowest excitations in each sector labeled by the magnetization M with total spin $S = \frac{1}{2}N - M$, there are no complex solutions of the Bethe ansatz equations. Thus, the analysis will be more transparent. Moreover, there is an important motivation to study gapless models because the results can be compared to the predictions of conformal invariance for critical, i.e. gapless, systems.

The Bethe ansatz equations for anisotropy $\Delta = \cos\theta$ $(0 < \theta < \pi)$ which determine the rapidities η_α are explicitly given by

$$2N \tan^{-1}\left(\cot\frac{\theta}{2} \tanh\frac{\eta_\alpha}{2} \right) = 2\pi J_\alpha + 2 \sum_{\beta=1}^{M} \tan^{-1}\left(\cot\theta \tanh\frac{\eta_\alpha - \eta_\beta}{2} \right) \tag{36.2}$$

or rewritten

$$2N\phi(\eta_\alpha, \frac{1}{2}\theta) = 2\pi J_\alpha + \sum_{\beta=1}^{M} \phi(\eta_\alpha - \eta_\beta, \theta). \tag{36.3}$$

introducing the useful function

$$\phi(\eta, \theta) = 2\tan^{-1}\left(cot\theta \tanh\frac{\eta}{2}\right). \tag{36.4}$$

The Bethe ansatz quantum numbers J_α

$$-\frac{1}{2}\left(\frac{1}{2}N - (S+1)\right); -\frac{1}{2}\left(\frac{1}{2}N - (S+1)\right) + 1; \ldots; \frac{1}{2}\left(\frac{1}{2}N - (S+1)\right) \tag{36.5}$$

determine the lowest energy state in the sector of total spin

$$S = \frac{1}{2}N - M. \tag{36.6}$$

The ground state corresponds to the lowest energy state in the sector $S = 0$.

In each sector the energy per site is given by

$$E_N^{(S)} = -\frac{1}{N}\sum_{\alpha=1}^{M} \frac{\sin^2\theta}{\cosh\eta_\alpha - \cos\theta} \tag{36.7}$$

once the rapidities have been calculated.

For a finite system this calculation of the rapidities from the Bethe ansatz equations (36.2) or (36.3) is, however, a formidable task except for systems consisting of only a very few spins. For reasonably large but still finite systems, there exists no systematic way of dealing with the Bethe ansatz equations directly.

What we can do, however, is using our knowledge of the thermodynamic limit and attempt to calculate, in a systematic way, corrections to this thermodynamic limit which are due to the finiteness of the system. We shall now pursue this strategy.

36.2 Finite–size corrections

From the Bethe ansatz equations (36.2) and (36.3), respectively, and from our experience with the thermodynamic limit, we are led to introduce the function

$$\begin{aligned} z_N(\eta) &= \frac{1}{2\pi}\left\{2\tan^{-1}\left(\cot\frac{\theta}{2}\tanh\frac{\eta}{2}\right)\right. \\ &\quad \left. -\frac{1}{N}\sum_{\beta=1}^{M} 2\tan^{-1}\left(\cot\theta\tanh\left(\frac{\eta-\eta_\beta}{2}\right)\right)\right\} \qquad (36.8) \\ &= \frac{1}{2\pi}\left\{\phi(\eta_\alpha, \frac{1}{2}\theta) - \frac{1}{N}\sum_{\beta=1}^{M}\phi(\eta_\alpha - \eta_\beta, \theta)\right\} \qquad (36.9) \end{aligned}$$

which by definition satisfies the Bethe ansatz equations

$$z_N(\eta_\alpha) = \frac{J_\alpha}{N} \tag{36.10}$$

if we insert the values of the rapidities $\eta = \eta_\alpha$. The function $z_N(\eta)$ is continuous and increases monotonically for real η.

We could now perform the thermodynamic limit $N \to \infty$ by introducing a density $\sigma_\infty(\eta)$ of Bethe ansatz roots

$$\sigma_\infty(\eta) = \frac{dz(\eta)}{d\eta} \tag{36.11}$$

$$z(\eta) \equiv z_\infty(\eta) = \lim_{N\to\infty} z_N(\eta). \tag{36.12}$$

However, instead of taking the thermodynamic limit, we keep the number of spins N finite and formally define the root density for the finite system by

$$\sigma_N(\eta) = \frac{dz_N(\eta)}{d\eta} \tag{36.13}$$

Using

$$\phi(\eta,\theta) = 2\tan^{-1}\left(\cot\theta \tanh\frac{\eta}{2}\right) \tag{36.14}$$

we arrive at

$$\sigma_N(\eta) = \frac{1}{2\pi}\left\{\phi'(\eta,\frac{\theta}{2}) - \frac{1}{N}\sum_{\beta=1}^{M}\phi'(\eta-\eta_\beta,\theta)\right\} \tag{36.15}$$

where

$$\phi'(\eta,\theta) = \frac{d\phi(\eta,\theta)}{d\eta} = \frac{\sin 2\theta}{\cosh\eta - \cos\theta}. \tag{36.16}$$

Next, we introduce the function

$$S_N(\eta) = \frac{1}{N}\sum_{\beta=1}^{M}\delta(\eta-\eta_\beta) - \sigma_N(\eta). \tag{36.17}$$

with which we can rewrite (36.15) as

$$\sigma_N(\eta) = \frac{1}{2\pi}\left\{\phi'(\eta,\frac{\theta}{2}) - \int_{-\infty}^{\infty} d\eta' S_N(\eta')\phi'(\eta-\eta',\theta)\right. \tag{36.18}$$

$$\left. - \int_{-\infty}^{\infty} d\eta' \sigma_N(\eta')\phi'(\eta-\eta',\theta)\right\}. \tag{36.19}$$

Thus we have arrived at an equation for the root density $\sigma_N(\eta)$ which resembles the integral equation in the thermodynamic limit. We write down both equations to make comparison easier.

The integral equation for root density $\sigma_\infty(\eta)$ in the thermodynamic limit

$$\sigma_\infty(\eta) = \frac{1}{2\pi}\frac{\sin\theta}{\cosh\eta - \cos\theta} - \int_{-\infty}^{\infty}\frac{d\eta}{2\pi}\frac{\sin 2\theta}{\cosh(\eta-\eta') - \cos 2\theta}\sigma_\infty(\eta'). \tag{36.20}$$

The solution of this equation can be found by Fourier transformation and reads (des Cloizeaux and Gaudin, 1966)

$$\sigma_\infty(\eta) = \frac{1}{4\theta}\frac{1}{\cosh\left(\frac{\pi\eta}{2\theta}\right)} \tag{36.21}$$

Exercise 36.1 Solve the integral equation (36.20) by Fourier transformation, i.e. confirm the result (36.21).

With this result for the infinite system under our belt we can now continue our attempt to find finite–size corrections. First, as just announced, we bring the equation for the finite–size root density $\sigma_N(\eta)$, (36.18), in the form

$$\frac{\sin\theta}{\cosh\eta - \cos\theta} = 2\pi\sigma_N(\eta) + \int_{-\infty}^{\infty} d\eta\,\frac{\sin 2\theta}{\cosh(\eta-\eta') - \cos 2\theta}\left(\sigma_N(\eta') + S_N(\eta')\right). \tag{36.22}$$

This equation, which closely resembles an integral equation, can be treated again by Fourier transforming it back and forth. The result is

$$\sigma_N(\eta) = \frac{1}{4\theta}\frac{1}{\cosh\left(\frac{\pi\eta}{2\theta}\right)} - \int_{-\infty}^{\infty} d\eta' F(\eta-\eta')S_N(\eta') \tag{36.23}$$

where the kernel function $F(x)$ is given by

$$F(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \frac{\exp(i\omega x)\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}. \tag{36.24}$$

Exercise 36.2 Perform a Fourier transformation and back transformation of (36.22) to bring it into the form (36.23). The book by Gradshteyn and Ryzhik (1980) will be a useful help for this exercise.

Compiling the results, we can write the equation for finite–size root density as

$$\sigma_N(\eta) = \sigma_\infty(\eta) - \int_{-\infty}^{\infty} d\eta' F(\eta-\eta')S_N(\eta') \tag{36.25}$$

This equation will be the starting point to which to apply the Euler–Maclaurin formula of section 34. As a result, we will obtain a proper integral equation for the finite–size root density $\sigma_N(\eta)$ where the inhomogeneous terms describe finite size corrections to the integral equation (36.20) for the root density in the thermodynamic limit.

Before we go on to do this, let us finish this section by deriving a finite–size expression for the energy per site in the sector of total spin S which we also express in terms of the result in the thermodynamic limit plus a contribution which contains the finite–size quantity $S_N(\eta)$.

This expression will then also be evaluated with the help of the Euler–Maclaurin formula to make the finite–size corrections explicit in a systematic way.

We begin from the energy per site (36.7) which we rewite using $S_N(\eta)$

$$\begin{aligned} E_N^{(S)} &= -\frac{1}{N}\sum_{\alpha=1}^{M}\frac{\sin^2\theta}{\cosh\eta_\alpha-\cos\theta} \equiv -\frac{1}{N}\sum_{\alpha=1}^{M}E_N(\eta_\alpha) \\ &= -\int_{-\infty}^{\infty}d\eta\,E(\eta)S_N(\eta) - \int_{-\infty}^{\infty}d\eta\,E(\eta)\sigma_N(\eta) \end{aligned} \tag{36.26}$$

This expression can be transformed into

$$E_N^{(S)} = E_\infty - \int_{-\infty}^{\infty}d\eta\epsilon(\eta)S_N(\eta), \tag{36.27}$$

where

$$\epsilon(\eta) = \frac{\pi}{2}\frac{\sin\theta}{\theta}\frac{1}{\cosh\left(\frac{\pi\eta}{2\theta}\right)}, \tag{36.28}$$

as we shall see in the following exercise.

Exercise 36.3 Derivation of the finite–size energy formula
Replace $\sigma_N(\eta)$ with (36.23) in (36.26)

Before we discuss the systematic treatment of finite–size corrections, we remark that it suffices, in order to calculate the lowest order corrections, to replace the finite–size root density $\sigma_N(\eta)$ in the function $S_N(\eta)$, (36.17), by the root density $\sigma_\infty(\eta)$ in the thermodynamic limit and the roots η_α by those of the thermodynamic limit, i.e. by those values of η_α which can be determined from

$$z_\infty(\eta_\alpha) = \frac{J_\alpha}{N}$$

in the limit $N\to\infty$. This has been done by Avdeev and Dörfel (1986) for the isotropic case $\Delta = 1$ and by Hamer (1985), (1986) for the anisotropic case $|\Delta| < 1$.

Obviously this truncates the possibility of calculating systematically further finite–size corrections as we shall do in the following section employing the Euler–Maclaurin formula.

36.3 Application of the Euler–Maclaurin formula

Our goal now is to apply the Euler–Maclaurin summation formula (34.6) to expressions of the form

$$\int_{-\infty}^{\infty}d\eta\,g(\eta)S_N(\eta) = \frac{1}{N}\sum_{\beta=1}^{M}g(\eta_\beta) - \int_{-\infty}^{\infty}d\eta\,g(\eta)\sigma_N(\eta)$$

which appear in the equation for the finite–size root density (36.23) or (36.25) and the finite–size energy (36.27). This integral can be rewritten using the definition of the finite–size root density $\sigma_N(\eta) = \frac{dz_N(\eta)}{d\eta}$

$$\int_{-\infty}^{\infty} d\eta\, g(\eta) S_N(\eta) = \frac{1}{N}\sum_{\beta=1}^{M} g(\eta_\beta) - \int_{-\infty}^{\infty} d\eta\, g(z_N)\frac{dz_N}{d\eta}$$
$$= \frac{1}{N}\sum_{\beta=1}^{M} g(\eta_\beta) - \int_{-\infty}^{\infty} dz_N\, g(z_N)$$

We now split the last integral

$$\int_{-\infty}^{\infty} dz_N\, g(z_N) = \int_{-\infty}^{-\Lambda} d\eta\, g(z_N)\frac{dz_N}{d\eta} + \int_{-\infty}^{\Lambda} d\eta\, g(z_N)\frac{dz_N}{d\eta}$$
$$+ \int_{-z_N(-\Lambda)}^{z_N(\Lambda)} dz_N\, g(z_N)$$
$$= \int_{-\infty}^{-\Lambda} d\eta\, g(\eta)\sigma_N(\eta) + \int_{-\infty}^{\Lambda} d\eta\, g(\eta)\sigma_N(\eta)$$
$$+ \int_{-z_N(-\Lambda)}^{z_N(\Lambda)} dz_N\, g(z_N) \qquad (36.29)$$

where Λ is determined from the largest root of the Bethe ansatz equations which is implicitly given by

$$z_N(\Lambda) = \frac{1}{N}\left\{\frac{1}{2}\left(\frac{N}{2} - (S+1)\right)\right\} \qquad (36.30)$$

or equivalently by

$$\int_{\Lambda}^{\infty} d\eta\, \sigma_N(\eta) = \left(1 - \frac{\theta}{\pi}\right)\frac{S}{N} + \frac{1}{2N} \qquad (36.31)$$

The property that the rapidity Λ defined in these equivalent ways is the largest root is guaranteed from the property of $z_N(\eta)$ to be a monotonically increasing function. All roots of the Bethe ansatz equations lie in the interval from $z_N(-\Lambda)$ to $z_N(\Lambda)$ which corresponds to the interval $(-\Lambda, \Lambda)$ in the rapidity variable η.

Exercise 36.4 Equivalence of the implicit definitions of the largest root Λ
Show the equivalence of the two definitions of Λ in (36.30) and (36.31).

It is the last integral in (36.29) to which we now apply the Euler–Maclaurin summation formula in the form

$$\frac{1}{2N}\left\{f\left(\frac{n_1}{N}\right) + 2\sum_{n=n_1+1}^{n_2-1} f\left(\frac{n}{N}\right) + f\left(\frac{n_2}{N}\right)\right\} - \int_{\frac{n_1}{N}}^{\frac{n_2}{N}} dx\, f(x)$$
$$= \frac{1}{12N^2}\left\{f'\left(\frac{n_2}{N}\right) - f'\left(\frac{n_1}{N}\right)\right\} + \mathcal{O}\left(\frac{\max f'''}{N^3}\right)$$

which uses the variable transformation $x \to Nx$ with $f(x) \to f(Nx) \to f(x)$.

We obtain

$$\int_{-\infty}^{\infty} d\eta\, g(\eta) S_N(\eta) = \frac{1}{N}\sum_{\beta=1}^{M} g(\eta_\beta) - \left\{ \int_{-\infty}^{-\Lambda} d\eta\, g(\eta)\sigma_N(\eta) + \int_{-\infty}^{\Lambda} d\eta\, g(z_N)\sigma_N(\eta) \right\} + \frac{1}{12N^2}\left\{ \left.\frac{dg(z_N)}{dz_N}\right|_{z_N=z_N(\Lambda)} - \left.\frac{dg(z_N)}{dz_N}\right|_{z_N=z_N(-\Lambda)} \right\} - \frac{1}{2N}\left\{ g(z_N(\Lambda)) + g(z_N(-\Lambda)) \right\} - \frac{1}{N}\sum_{\beta=1}^{M}{}^{*} g\left(z_N(\eta_\beta) = \frac{J_\beta}{N} \right)$$

where in the last summation, indicated by the $*$, it is understood that the summation does not run over the smallest and largest roots, $-\Lambda$ and Λ, respectively. We observe that the sums in this expression almost cancel leaving only a contribution from the smallest and largest roots, $-\Lambda$ and Λ, respectively, so that we obtain

$$\int_{-\infty}^{\infty} d\eta\, g(\eta) S_N(\eta) = - \left\{ \int_{-\infty}^{-\Lambda} d\eta\, g(z_N)\sigma_N(\eta) + \int_{-\infty}^{\Lambda} d\eta\, g(\eta)\sigma_N(\eta) \right\} + \frac{1}{12N^2}\left\{ \left.\frac{dg(\eta)}{d\eta}\frac{d\eta}{dz_N}\right|_{\eta=\Lambda} - \left.\frac{dg(\eta)}{d\eta}\frac{d\eta}{dz_N}\right|_{\eta=-\Lambda} \right\} - \frac{1}{2N}\left\{ g(\Lambda) + g(-\Lambda) \right\}.$$

Using again $\sigma_N(\eta) = \frac{dz_N(\eta)}{d\eta}$ and its symmetry $\sigma_N(\eta) = \sigma_N(-\eta)$ we can finally write

$$\int_{-\infty}^{\infty} d\eta\, g(\eta) S_N(\eta) = - \left\{ \int_{-\infty}^{-\Lambda} d\eta\, g(\eta)\sigma_N(\eta) + \int_{-\infty}^{\Lambda} d\eta\, g(z_N)\sigma_N(\eta) \right\} + \frac{1}{12N^2\sigma_N(\Lambda)}\left\{ g'(\Lambda) - g'(-\Lambda) \right\} - \frac{1}{2N}\left\{ g(\Lambda) + g(-\Lambda) \right\}. \tag{36.32}$$

This last expression (36.32) is in a form to which we now can apply the second mathematical technique we have introduced at the beginning of this part: The Wiener–Hopf technique. As we shall see in the next section, using (36.32) in the expressions for the finite–size root density (36.23) or (36.25) finally renders them into proper integral equations which are, after some further rewriting, of the Wiener–Hopf type. From these Wiener–Hopf type integral equations we then derive functional equations from whose solutions we use to evaluate the finite–size energy (36.27) which we shall also express with the help of (36.32) in a form to read off finite–size corrections.

36.4 Application of the Wiener–Hopf technique

36.4.1 Derivation of the Wiener–Hopf type integral equation

Using the approximation (36.32), the expressions for the finite–size root density (36.23) reads

$$\begin{aligned}\sigma_N(\eta) - \int_{-\infty}^{-\Lambda} d\eta'\, F(\eta-\eta')\sigma_N(\eta') - \int_{\Lambda}^{\infty} d\eta'\, F(\eta-\eta')\sigma_N(\eta') \\ + \frac{1}{2N}\left(F(\eta-\Lambda)+F(\eta+\Lambda)\right) + \frac{1}{12N^2\sigma_N(\Lambda)}\left(F'(\eta-\Lambda)-F'(\eta+\Lambda)\right) \\ = \frac{1}{4\theta}\frac{1}{\cosh\left(\frac{\pi\eta}{2\theta}\right)}.\end{aligned} \tag{36.33}$$

We now shift the integration limits $\pm\Lambda \to 0$, i.e. first we shift the integration variable $\eta' \to \eta' - \Lambda$ and then we let $\eta' \to -\eta'$

$$\begin{aligned}\int_{-\infty}^{-\Lambda} d\eta'\, F(\eta-\eta')\sigma_N(\eta') &= \int_{-\infty}^{0} d\eta'\, F(\eta-\eta'+\Lambda)\sigma_N(\eta'-\Lambda) \\ &= \int_{0}^{\infty} d\eta'\, F(\eta+\eta'+\Lambda)\sigma_N(\eta'+\Lambda)\end{aligned}$$

where we again used the symmetry $\sigma_N(\eta) = \sigma_N(-\eta)$. Similarly we obtain

$$\int_{\Lambda}^{\infty} d\eta'\, F(\eta-\eta')\sigma_N(\eta') = \int_{0}^{\infty} d\eta'\, F(\eta-\eta'-\Lambda)\sigma_N(\eta'+\Lambda).$$

Thus we arrive at

$$\begin{aligned}\sigma_N(\eta) - \int_{0}^{\infty} d\eta'\, F(\eta-\eta'-\lambda)\sigma_N(\eta'+\Lambda) - \int_{0}^{\infty} d\eta'\, F(\eta+\eta'+\Lambda)\sigma_N(\eta'+\Lambda) \\ + \frac{1}{2N}\left(F(\eta-\Lambda)+F(\eta+\Lambda)\right) + \frac{1}{12N^2\sigma_N(\Lambda)}\left(F'(\eta-\Lambda)-F'(\eta+\Lambda)\right) \\ = \frac{1}{4\theta}\frac{1}{\cosh\left(\frac{\pi\eta}{2\theta}\right)}.\end{aligned} \tag{36.34}$$

The next step of transformation into a recognizable integral equation of Wiener–Hopf type consists of shifting the variable η: $\eta \to \eta + \Lambda$.

$$\begin{aligned}\sigma_N(\eta+\Lambda) - \int_{0}^{\infty} d\eta'\, F(\eta-\eta')\sigma_N(\eta'+\Lambda) - \int_{0}^{\infty} d\eta'\, F(\eta+\eta'+2\Lambda)\sigma_N(\eta'+\Lambda) \\ + \frac{1}{2N}\left(F(\eta)+F(\eta+2\Lambda)\right) + \frac{1}{12N^2\sigma_N(\Lambda)}\left(F'(\eta)-F'(\eta+2\Lambda)\right) \\ = \frac{1}{4\theta}\frac{1}{\cosh\left(\frac{\pi(\eta+\Lambda)}{2\theta}\right)}.\end{aligned} \tag{36.35}$$

The recipe of the Wiener–Hopf technique now redefines the function $\sigma_N(\eta)$

$$\sigma_N^+(\eta) = \begin{cases} \sigma_N(\eta+\Lambda) & \eta > 0 \\ 0 & \eta < 0 \end{cases} \tag{36.36}$$

and

$$\sigma_N^-(\eta) = \begin{cases} 0 & \eta > 0 \\ \sigma_N(\eta+\Lambda) & \eta < 0 \end{cases} \tag{36.37}$$

so that we finally arrive at the desired Wiener–Hopf type integral equation

$$\begin{aligned} \sigma_N^+(\eta) + \sigma_N^-(\eta) - \int_0^\infty d\eta'\, F(\eta-\eta')\sigma_N^+(\eta') - \int_0^\infty d\eta'\, F(\eta+\eta'+2\Lambda)\sigma_N^+(\eta') \\ + \frac{1}{2N}\left(F(\eta)+F(\eta+2\Lambda)\right) + \frac{1}{12N^2\sigma_N(\Lambda)}\left(F'(\eta)-F'(\eta+2\Lambda)\right) \\ = \frac{1}{4\theta}\frac{1}{\cosh\left(\frac{\pi(\eta+\Lambda)}{2\theta}\right)}. \end{aligned} \tag{36.38}$$

This integral equation is now in a proper form to be Fourier transformed which we shall do in the next section.

36.4.2 Fourier transformation of the Wiener–Hopf type integral equation

In order to Fourier transform the Wiener–Hopf type integral equation (36.38) we recall Fourier tarnsform of the kernel function (36.24) which we calculated in exercise 36.2:

$$F(\eta-\eta') = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \frac{\exp(i\omega(\eta-\eta'))\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}.$$

Its derivative with respect to the second variable η' is given by

$$\frac{\partial}{\partial\eta'}F(\eta-\eta') = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega(-i\omega) \frac{\exp(i\omega(\eta-\eta'))\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}.$$

These two equations for the kernel $F(\eta)$ and its derivative suffice, together with the Fourier transform of the right hand side inhomogeneity, to write down the Fourier transform of the Wiener–Hopf type integral equation (36.38) as

$$\begin{aligned} \sigma^+(\omega) + \sigma^-(\omega) - \sigma^+(\omega)\frac{\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)} \\ - \sigma^+(-\omega)\frac{\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}e^{2i\omega\Lambda} \\ + \frac{1}{2N}\frac{1}{2\pi}\frac{\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}\left[1+e^{2i\omega\Lambda}\right] \\ - \frac{1}{12N^2\sigma_N(\Lambda)}\frac{i\omega}{2\pi}\frac{\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}\left[1-e^{2i\omega\Lambda}\right] \\ = \frac{1}{4\pi}\frac{1}{\cosh\omega\theta}e^{i\omega\Lambda}. \end{aligned} \tag{36.39}$$

Let us rearrange this lengthy expression:

$$\begin{aligned}\sigma^-(\omega) + \frac{\sinh\omega\pi}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}\sigma^+(\omega) \\ = \frac{1}{4\pi}\frac{1}{\cosh\omega\theta}e^{i\omega\Lambda} \\ - \frac{1}{2\pi}\frac{\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)}\left\{\left[\frac{1}{2N} - \frac{i\omega}{12N^2\sigma_N(\Lambda)}\right]\right. \\ \left. + \left[\frac{1}{2N} + \frac{i\omega}{12N^2\sigma_N(\Lambda)} - 2\pi\sigma^+(-\omega)\right]e^{2i\omega\Lambda}\right\}. \end{aligned} \tag{36.40}$$

This equation, together with the equation determining the largest root Λ

$$\left(1-\frac{\theta}{\pi}\right)\frac{S}{N} + \frac{1}{2N} = \int_\Lambda^\infty d\eta\,\sigma_N(\eta) = \int_0^\infty d\eta\,\sigma_N(\eta+\Lambda) = 2\pi\sigma^+(0), \tag{36.41}$$

an equation determining $\sigma_N(\Lambda)$

$$\sigma_N(\Lambda) = 2\int_{-\infty}^\infty d\omega\,\sigma^+(\omega), \tag{36.42}$$

and the equation for the finite–size energy per spin (36.27), together with (36.28), completely determine the problem of calculating the finite–size corrections to the energy. Of course, we still have to use (36.32) which we obtained from the application of the Euler–Maclaurin summation formula to the finite–size energy per spin (36.27). This yields

$$\begin{aligned}E_N^{(S)} - E_\infty = \int_\Lambda^\infty d\eta\,\epsilon(\eta)\sigma_N(\eta) + \int_\Lambda^\infty d\eta\,\epsilon(\eta)\sigma_N(\eta) \\ - \frac{1}{2N}\left\{\epsilon(\Lambda)+\epsilon(-\Lambda)\right\} - \frac{1}{12N^2\sigma_N(\Lambda)}\left\{\epsilon'(\Lambda)-\epsilon'(-\Lambda)\right\}. \end{aligned} \tag{36.43}$$

Explicitly we have for the so–called dressed energy $\epsilon(\eta)$ (see (36.28)) which we need for large values of the argument corresponding to the largest root Λ

$$\epsilon(\eta) = \frac{\pi}{2}\frac{\sin\theta}{\theta}\frac{1}{\cosh\left(\frac{\pi\eta}{2\theta}\right)} \to \pi\frac{\sin\theta}{\theta}e^{-\left(\frac{\pi|\eta|}{2\theta}\right)} \quad \text{for} \quad \eta \gg 1. \tag{36.44}$$

The derivative of this expression is

$$\epsilon'(\eta) = \left(\frac{\pi}{2}\right)^2\frac{\sin\theta}{\theta^2}\frac{\sinh\left(\frac{\pi\eta}{2\theta}\right)}{\cosh^2\left(\frac{\pi\eta}{2\theta}\right)} \to \frac{\pi^2}{2}\frac{\sin\theta}{\theta^2}e^{-\left(\frac{\pi|\eta|}{2\theta}\right)} \quad \text{for} \quad \eta \gg 1. \tag{36.45}$$

Exercise 36.5 Prove (36.42) using Fourier transformation and the definition of the step function where $\Theta(0) = \frac{1}{2}$.

The various parts on the right hand side of (36.43) have to be further manipulated. We start with

$$-\frac{1}{2N}\left\{\epsilon(\Lambda)+\epsilon(-\Lambda)\right\}-\frac{1}{12N^2\sigma_N(\Lambda)}\left\{\epsilon'(\Lambda)-\epsilon'(-\Lambda)\right\}= (2\pi)^2\frac{\sin\theta}{\theta}e^{-\left(\frac{\pi\Lambda}{2\theta}\right)}\frac{(-1)}{2\pi}\left[\frac{1}{2N}-\frac{\frac{\pi}{2\theta}}{12N^2\sigma_N(\Lambda)}\right]. \quad (36.46)$$

Furthermore, we have

$$\int_{-\infty}^{-\Lambda}d\eta\,\epsilon(\eta)\sigma_N(\eta)=\int_{-\infty}^{\infty}d\eta\,\epsilon(\eta+\Lambda)\sigma^+(\eta) \quad (36.47)$$

and

$$\int_{\Lambda}^{\infty}d\eta\,\epsilon(\eta)\sigma_N(\eta)=\int_{-\infty}^{\infty}d\eta\,\epsilon(\eta+\Lambda)\sigma^+(\eta), \quad (36.48)$$

i.e. the contributions in (36.46) contribute the same. Moreover, with

$$\sigma^+(\eta)=\int_{-\infty}^{\infty}d\omega\,e^{i\omega\eta}\sigma^+(\omega) \quad (36.49)$$

these contributions become

$$\int_{-\infty}^{\infty}d\eta\,\epsilon(\eta+\Lambda)\sigma^+(\eta)=\pi\frac{\sin\theta}{\theta}e^{-\frac{\pi\Lambda}{2\theta}}\int_{-\infty}^{\infty}d\eta\int_{-\infty}^{\infty}d\omega\,e^{i\omega\eta}\sigma^+(\omega)e^{i\eta(\omega+\frac{i\pi}{2\theta})} = 2\pi^2\frac{\sin\theta}{\theta}e^{-\frac{\pi\Lambda}{2\theta}}\sigma^+\left(-\frac{i\pi}{2\theta}\right). \quad (36.50)$$

We can now collect the contributions to the finite–size energy per spin

$$E_N^{(S)}-E_\infty=(2\pi)^2\frac{\sin\theta}{\theta}e^{-\frac{\pi\Lambda}{2\theta}}\left(\sigma^+\left(-\frac{i\pi}{2\theta}\right)-\frac{1}{2\pi}\left[\frac{1}{2N}-\frac{\frac{\pi}{2\theta}}{12N^2\sigma_N(\Lambda)}\right]\right). \quad (36.51)$$

This expression for the finite–size energy per spin is, of course, only an approximation. It can be shown by considering further terms in the Euler–Maclaurin summation formula which contain higher derivatives of $\epsilon(\Lambda)$, that the corrections are of order $e^{-\frac{3\pi\Lambda}{2\theta}}$.

The expressions (36.40), (36.41), (36.42), and (36.51) form a closed set of equations which we now have to further investigate by applying the Wiener–Hopf technique to the functional equation (36.40).

In order to get acquainted with the method described so far, the following exercise is recommended.

Exercise 36.6 Isotropic case
Redo the calculations for the isotropic antiferromagnetic case $\Delta = 1$ to obtain the expressions corresponding to (36.40), (36.41), (36.42), and (36.51). The results are:

$$\sigma^-(\omega)+\frac{1}{1+e^{-2|\omega|}}\sigma^+(\omega) = \frac{1}{2\pi}\frac{e^{-2|\omega|}}{1+e^{-2|\omega|}} \left\{\left(\frac{1}{2N}-\frac{i\omega}{12N^2\sigma_N(\Lambda)}\right)+\left(\frac{1}{2N}+\frac{i\omega}{12N^2\sigma_N(\Lambda)}-2\pi\sigma^+(-\omega)\right)e^{2\omega\Lambda}\right\}$$

and

$$E_N^{(S)} - E_\infty = (2\pi)^2 e^{-\frac{\pi\Lambda}{2}} \left\{ \sigma^+\left(-\frac{i\pi}{2}\right) - \frac{1}{2\pi}\left[\frac{1}{2N} - \frac{\frac{\pi}{2}}{12N^2\sigma_N(\Lambda)}\right]\right\}$$

while (36.41) and (36.42) remain unchanged.

This case generalized to an asymmetric distribution of Bethe ansatz quantum number J_α is treated in detail in the summer school proceedings contribution of Ferenc Woynarovich (1997).

36.4.3 Functional equation

According to the prescription of the Wiener–Hopf technique, we now need to write the functional equation (36.40) in a form such that both sides of the equation decompose into two additive parts, one being analytic in the upper, the other in the lower half of the complex ω–plane.

As a first step to achieve this goal, we observe that

$$1 - \frac{\sinh\omega(\pi-2\theta)}{2\cosh\omega\theta\sinh\omega(\pi-\theta)} = \frac{\sinh\omega\pi}{2\cosh\omega\theta\sinh\omega(\pi-\theta)} \equiv \frac{1}{G(\omega)G(-\omega)}$$

where (see Gradshteyn and Ryzhik (1980) page 937)

$$\frac{1}{G(\omega)G(-\omega)} = \frac{1}{2(\pi-\theta)} \frac{\Gamma\left(\frac{1}{2}+i\omega\frac{\theta}{\pi}\right)\Gamma\left(1+i\omega\left(1-\frac{\theta}{\pi}\right)\right)\Gamma\left(\frac{1}{2}-i\omega\frac{\theta}{\pi}\right)\Gamma\left(1-i\omega\left(1-\frac{\theta}{\pi}\right)\right)}{\Gamma\left(1+i\omega\right)\Gamma\left(1+i\omega\right)}$$

In order to factorize this expression, we have to make sure that the individual factors do not diverge for $\omega \to \infty$. Rather the individual factors should be finite $G(\omega) \to 1$ for $\omega \to \infty$. This can be achieved by considering the asymptotic of the complex Γ–function (see e.g. Abramowitz and Stegun (1972) page 257). We obtain

$$\frac{1}{G(\omega)} = \frac{1}{\sqrt{2(\pi-\theta)}} \frac{\Gamma\left(\frac{1}{2}+i\omega\frac{\theta}{\pi}\right)\Gamma\left(1+i\omega\left(1-\frac{\theta}{\pi}\right)\right)}{\Gamma\left(1+i\omega\right)} e^{-i\omega\left[\left(1-\frac{\theta}{\pi}\right)\ln\left(1-\frac{\theta}{\pi}\right)+\frac{\theta}{\pi}\ln\frac{\theta}{\pi}\right]}$$

The functional equation can now be written as

$$G(-\omega)\sigma^-(\omega) + \frac{\sigma^+(\omega)}{G(\omega)} = \frac{1}{4\pi}\frac{G(-\omega)}{\cosh\omega\theta}e^{i\omega\Lambda} + \frac{1}{2\pi}\left[\frac{1}{G(\omega)} - G(-\omega)\right]\left[\frac{1}{2N} - \frac{i\omega}{12N^2\sigma_N(\Lambda)}\right] \tag{36.52}$$

where, at this stage of the approximation, in the functional equation (36.40) we have neglected the term proportional to $e^{2i\omega\Lambda} = e^{2i\omega'\Lambda}e^{-2\omega''\Lambda}$.

Of the two terms on the left hand side of the functional equation (36.52) the first can be shown to be analytic in the upper complex ω half plane, while the second is analytic in the lower complex ω half plane.

Exercise 36.7 Show that the statement above is correct by considering the Fourier transform of $\sigma^\pm(\omega)$ and the poles and zeroes of $\frac{1}{G(\pm\omega)}$.

Now we need to write also the right hand side of (36.52) as a sum of two terms, one of which is analytic in the upper, the other in the lower half plane. In our case this can be achieved using Cauchy's integral formula (see Roos (1969), page 169)

This leads for the part of the functional equation in the lower half plane to the expression

$$\sigma^+(\omega) = \lim_{\epsilon\to 0} \quad \frac{(-G(\omega))}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega+i\epsilon}\times \\ \times\left\{\frac{G(-\omega')}{2\cosh\omega'\theta}e^{i\omega'\Lambda}+\left[\frac{1}{G(\omega')}-G(-\omega')\right]\left[\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right]\right\}$$

We deal with the three additive terms on the right hand side one by one. The first term can be evaluated using the residue theorem and reads

$$-\frac{G(\omega)}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega+i0}\,\frac{G(-\omega')}{2\cosh\theta\omega'} = -\frac{1}{4\pi i\theta}\frac{G(-\frac{i\pi}{2\theta})G(\omega)}{\frac{i\pi}{2\theta}-\omega}e^{-\frac{\pi\Lambda}{2\theta}}+\mathcal{O}\left(e^{-\frac{3\pi\Lambda}{2\theta}}\right)$$

We only need this for $\omega = \frac{-i\pi}{2\theta}$ and for $\omega = 0$. For those two values, we obtain

$$\frac{1}{(2\pi)^2}G^2\left(-\frac{i\pi}{2\theta}\right)e^{-\frac{\pi\Lambda}{2\theta}} \tag{36.53}$$

and

$$\frac{1}{(2\pi)^2}G(0)G\left(-\frac{i\pi}{2\theta}\right)e^{-\frac{\pi\Lambda}{2\theta}}, \tag{36.54}$$

respectively.

The second term produces after a lengthy calculation which uses contour integration and again the residue theorem

$$\lim_{\epsilon\to 0}\frac{(-G(\omega))}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega+i\epsilon}\,\frac{1}{G(\omega')}\left[\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right] = \frac{1}{2\pi}\left[\frac{1}{2N}-\frac{\pi}{2\theta}\frac{1}{12N^2\sigma_N(\Lambda)}\right]$$

for $\omega = -\frac{i\pi}{2\theta}$ and

$$\lim_{\epsilon\to 0}\frac{(-G(\omega))}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega+i\epsilon}\,\frac{1}{G(\omega')}\left[\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right] = \frac{1}{2\pi}\frac{1}{2N}$$

for $\omega = 0$.

For the third and last term we use the asymptotic form $G(\omega') \approx e^{\frac{\mu}{12}\frac{1}{i\omega'}} \approx 1+\frac{\mu}{12}\frac{1}{i\omega'}$ where $\mu = \frac{\theta}{\pi-\theta}-\frac{\pi}{2\theta}$ and the expansion $\frac{1}{\left(\omega'+\frac{i\pi}{2\theta}\right)} \approx \frac{1}{\omega'}-\frac{i\pi}{2\theta\,\omega'^2}$

$$\lim_{\epsilon\to 0}\quad\frac{G(\omega)}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega+i\epsilon}\,G(-\omega')\left[\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right] \\ = -\frac{G\left(-\frac{i\pi}{2\theta}\right)}{2\pi}\left[\frac{1}{2N}-\frac{\mu}{12}\frac{1}{12N^2\sigma_N(\Lambda)}-\frac{\pi}{2\theta}\frac{1}{12N^2\sigma_N(\Lambda)}\right]$$

for $\omega = -\frac{i\pi}{2\theta}$ and

$$\lim_{\epsilon\to 0}\frac{G(\omega)}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega+i\epsilon}\,G(-\omega')\left[\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right] = -\frac{G(0)}{2\pi}\left[\frac{1}{2N}-\frac{\mu}{12}\frac{1}{12N^2\sigma_N(\Lambda)}\right]$$

for $\omega = 0$.

Collecting terms, we arrive at the general expression

$$\begin{aligned}\sigma^+(\omega) = & -\frac{1}{2i\theta}\frac{G\left(-\frac{i\pi}{2\theta}\right)}{2\pi}\frac{G(\omega)}{\frac{i\pi}{2\theta}-\omega}e^{-\frac{\pi\Lambda}{2\theta}} \\ & -\frac{G(\omega)}{(2\pi)^2 i}\int_{-\infty}^{\infty}\frac{d\omega'}{\omega'-\omega-i0}\left(\frac{1}{G(\omega')}-G(-\omega')\right)\left(\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right)\end{aligned}$$

which we evaluated for $\omega = -\frac{i\pi}{\theta}$

$$\begin{aligned}\sigma^+\left(-\frac{i\pi}{2\theta}\right) = & \left(\frac{G\left(-\frac{i\pi}{2\theta}\right)}{2\pi}\right)^2 e^{-\frac{\pi\Lambda}{2\theta}} + \frac{1}{2\pi}\left(\frac{1}{2N}-\frac{\pi}{2\theta}\frac{1}{12N^2\sigma_N(\Lambda)}\right) \\ & -\frac{G\left(-\frac{i\pi}{2\theta}\right)}{2\pi}\left(\frac{1}{2N}-\frac{\mu}{12}\frac{1}{12N^2\sigma_N(\Lambda)}-\frac{\pi}{2\theta}\frac{1}{12N^2\sigma_N(\Lambda)}\right) \quad (36.55)\end{aligned}$$

and for $\omega = 0$

$$\begin{aligned}2\pi\sigma^+(0) &= \left(\frac{G\left(-\frac{i\pi}{2\theta}\right)G(0)}{2\pi}\right)e^{-\frac{\pi\Lambda}{2\theta}} + \frac{1}{2N} - G(0)\left(\frac{1}{2N}-\frac{\mu}{12}\frac{1}{12N^2\sigma_N(\Lambda)}\right) \\ &= \left(1-\frac{\theta}{\pi}\right)\frac{S}{N}+\frac{1}{2N}. \quad (36.56)\end{aligned}$$

Furthermore we need to evaluate

$$\begin{aligned}\sigma_N(\Lambda) &= 2\int_{-\infty}^{\infty} d\omega\,\sigma^+(\omega) \\ &= -\frac{1}{i\theta}\frac{G\left(-\frac{i\pi}{2\theta}\right)}{2\pi}e^{-\frac{\pi\Lambda}{2\theta}}\int_{-\infty}^{\infty}d\omega\,\frac{G(\omega)}{\frac{i\pi}{2\theta}-\omega} \\ &\quad -\frac{2}{(2\pi)^2 i}\int_{-\infty}^{\infty}\left(\frac{1}{G(\omega')}-G(-\omega')\right)\left(\frac{1}{2N}-\frac{i\omega'}{12N^2\sigma_N(\Lambda)}\right)\int_{-\infty}^{\infty}d\omega d\omega'\,\frac{G(\omega)}{\omega'-\omega-i0}.\end{aligned}$$

With the help of complex integration and the asymptotic expansions for $G(\omega)$ and $(\omega-\omega')^{-1}$ we have already employed, to lowest order this becomes

$$\sigma_N(\Lambda) = \frac{1}{2\theta}G\left(-\frac{i\pi}{2\theta}\right)e^{-\frac{\pi\Lambda}{2\theta}} + \frac{\mu}{12}\left(\frac{1}{2N}-\frac{\mu}{12}\frac{1}{2}\frac{1}{12N^2\sigma_N(\Lambda)}\right). \quad (36.57)$$

The expressions (36.57), (36.56), and (36.55) suffice to determine the lowest order finite–size corrections to the energy per spin.

36.4.4 Finite–size corrections to the energy

Now employing skillful algebraic manipulation we can extract from (36.57), (36.56), and (36.55) the final formula for the finite–size energy per spin (36.51) which is exact to lowest order in the inverse system size $\frac{1}{N}$

$$E_N^{(S)} - E_\infty = (2\pi)^2\frac{\sin\theta}{\theta}\left(\frac{\left(1-\frac{\theta}{\pi}\right)S^2}{8N^2}-\frac{1}{48N^2}\right) \quad (36.58)$$

For example the ground state finite–size energy correction in the $S = 0$ sector becomes

$$N(E_N^{(0)} - E_\infty) = -\pi^2 \frac{\sin\theta}{\theta} \frac{1}{12N} = -\frac{\pi\zeta}{6N} \tag{36.59}$$

where $\zeta = \frac{\pi}{2}\frac{\sin\theta}{\theta}$ whereas the finite-size correction for general S and $S = 0$ reads

$$N(E_N^{(S)} - E_N^{(S)}) = \pi^2 \frac{\sin\theta}{\theta} \frac{1}{2N} = \frac{2\pi\zeta}{N} x_S \tag{36.60}$$

where $x_S = \left(1 - \frac{\theta}{\pi}\right)\frac{S^2}{2}$.

These two expressions coincide precisely with the predictions of conformal invariance for critical one–dimensional systems (Blöte *et al.*, 1986; Affleck, 1986) (see also e.g.Cardy (1987)).

36.5 Higher order corrections

The task to calculate higher order corrections consists first of all in a careful analysis of the approximations made such that we can figure out what contributions in powers of $\frac{1}{N}$ they will make.

References

Abramowitz, Milton and Stegun, Irene A (1972). *Handbook of mathematical functions*. Dover, New York.

Affleck, Ian (1986). Universal term in the free energy at a critical point and the conformal anomaly. *Phys Rev Lett*, **56**, 746.

Alcaraz, Francisco, Barber, Michael, Batchelor, Murray, Baxter, Rodney, and Quispel, Reinout (1987). Surface exponents of the quantum XXZ, Ashkin-Teller, and Potts models. *J Phys A: Math Gen*, **20**, 6397.

Alcaraz, Francisco C, Barber, Michael N, and Batchelor, Murray T (1988). Conformal Invariance, the XXZ Chain and the Operator Content of Two–Dimensional Critical Systems. *Ann Phys*, **182**, 280.

Altland, Alexander and Simons, Ben D (2010). *Condensed Matter Field Theory* (Second edn). Cambridge University Press.

Amico, Luigi, Frahm, Holger, Osterloh, Andreas, and Ribeiro, G A P (2007). Integrable spin–boson models descending from rational six–vertex models. *Nucl Phys B [FS]*, **787**, 283–300.

Anderson, Philip W (1952). An Approximate Quantum Theory of the Antiferromagnetic Ground State. *Phys Rev*, **86**, 694.

Anderson, Philip W (1972). More is different. *Science*, **177**, 393.

Anderson, Philip W (1987). The Resonating Valence Bond State in La_2CuO_4 and Superconductivity. *Science*, **235**, 1196.

Andrei, Natan (1983). *Integrable models in condensed matter physics*, Volume 6, Series on Modern Condensed Matter Physics, p. 458. World Scientific, Singapore.

Andrei, Natan, Furuya, K, and Lowenstein, John H (1983). Solution of the kondo problem. *Rev Mod Phys*, **55**, 331.

Arovas, Daniel (2013). Lecture Notes on Thermodynamics and Statistical Mechanics (A Work in Progress). `http://www-physics.ucsd.edu/students/courses/spring2010/physics210a/`. Lecture notes, University of California San Diego.

Ashcroft, Neil W and Mermin, N David (1976). *Solid State Physics*. Holt, Rinehart, and Winston, New York.

Auerbach, Assa (1994). *Interacting Electrons and Magnetism*. Springer, New York.

Avdeev, L V and Dörfel, Bernd-D (1986). Finite–size corrections for the XXX antiferromagnet. *J Phys A: Math Gen*, **19**, L13–L17.

Bargmann, Valentine (1961). On a Hilbert Space of Analytic Functions and an Associated Integral Transform, Part I. *Commun Pure Appl Math*, **14**, 187–214.

Batchelor, Murray and Hamer, Chris (1990). Surface energy of integrable quantum spin chains. *J Phys A: Math Gen*, **23**, 761.

Batchelor, Murray T (2007). The Bethe ansatz after 75 years. *Physics Today*, **60**(1), 36.

Batchelor, Murray T and Blöte, Henk W J (1989). Conformal invariance and critical behavior

of the $o(n)$ model on a honeycomb lattice. *Phys Rev B*, **39**, 2391.

Batchelor, Murray T, Nienhuis, Bernard, and Warnaar, Ole S (1989). Bethe–ansatz results for a solvable $o(n)$ model on a square lattice. *Phys Rev Lett*, **62**, 2425.

Baxter, Rodney (1982). *Exactly solved models in statistical mechanics*. Academic Press, London. Reprinted by Dover Publications (2007) with a new chapter on *Subsequent developments* added by the author. The original 1982 edition can be freely downloaded at the web site listed above.

Baym, Gordon and Pethick, Christopher (2004). *Landau Fermi Liquid Theory: Concepts and Applications*. Wiley–VCH.

Berezin, Felix A (1966). *The Method of Second Quantization*. Academic Press, New York and London.

Bernal, John D and Fowler, Ralph H (1933). A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J Chem Phys*, **1**, 515.

Bethe, Hans (1931). Zur Theorie der Metalle. I. Eigenwerte und Eigenfunktionen der linearen Atomkette. *Z Physik*, **71**, 205.

Bethe, Hans A (1997). *Selected Works of Hans A Bethe — With Commentary*. World Scientific, Singapore.

Binney, James J, Dowrick, Nigel J, Fisher, Andrew J, and Newman, Mark E J (1992). *The Theory of Critical Phenomena*. Oxford University Press.

Bloch, Felix (1930). Zur Theorie des Ferromagnetismus. *Z Physik*, **61**, 206.

Bloch, Felix (1934). Inkohärente Röntgenstreuung und Dichteschwankungen eines entarteten Fermigases. *Helv Phys Acta*, **7**, 385.

Blum, Thomas and Shapir, Yonathan (1990). A slouble seven–vertex model for clusters with interfacial bending energy. *J Phys A*, **23**, L511.

Blöte, Henk W, Cardy, John L, and Nightingale, M Peter (1986). Conformal invariance, the central charge, and universal finite-size amplitudes at criticality. *Phys Rev Lett*, **56**, 742.

Bogoliubov, Nicholay M (1996). Exactly solvable models of quantum nonlinear optics. University of Helsinki, Research Institute of Theoretical Physics, Report Series in Theoretical Physics HU–TFT–IR–96–3.

Bogoliubov, Nicholay M, Bullough, Robin K, and Timonen, Jussi (1996). Exact solution of generalized Tavis–Cummings models in quantum optics. *J Phys A*, **29**, 6305–6312.

Bohr, Niels (1911). *Studier over Metallernes Elektrontheori*. Ph.D. thesis, University of Copenhagen.

Braak, Daniel (2011). Integrability of the Rabi Model. *Phys Rev Lett*, **107**, 100401. (see also the online supplement).

Braak, Daniel (2013). A generalized G–function for the quantum Rabi model. *Ann Phys (Berlin)*, **525**, L23–L28.

Brush, Stephen G (1967). History of the Lenz–Ising Model. *Rev Mod Phys*, **39**, 883.

Callen, Herbert B (1960). *Thermodynamics*. John Wiley, New York.

Callen, Herbert B (1985). *Thermodynamics and an Introduction to Thermostatics* (Second edn). John Wiley, New York.

Cardy, John L (1987). Conformal invariance. In *Phase Transitions and Critical Phenomena* (ed. C. Domb and J. L. Lebowitz), Volume 11. Academic Press, London.

Cardy, John L (1996). *Scaling and Renormalization in Statistical Physics*. Cambridge University Press.

Chandler, David (1987). *Introduction to Modern Statistical Mechanics*. Oxford University Press.

Chang, C H (1952). The Spontaneous Magnetization of a Two–Dimensional Rectangular Ising Model. *Phys Rev*, **88**, 1422.

Chilingaryan, S A and Rodríguez-Lara, Blas Manuel (2013). Quantum Rabi model for two qubits. arXiv:1301.4462v3 [quant-ph].

Cohen-Tannoudji, Claude, Diu, Bernard, and Laloë, Franck (1977). *Quantum Mechanics*. Volume 1. John Wiley, New York.

Coleman, Piers (2003). Many–Body Physics: Unfinished Revolution. *Ann Henri Poincaré*, **4**, 1–22.

Cross, Michael (2004). Statistical Physics. `http://www.pma.caltech.edu/~mcc/Ph127/index.html`. Lecture notes of a three term course at the California Institute of Technology.

de Vega, Hector J and Woynarovich, Ferenc (1985). Method for calculating finite size corrections in Bethe ansatz systems: Heisenberg chain and six–vertex model. *Nucl Phys B*, **251[FS13]**, 439–456.

des Cloizeaux, Jacques and Gaudin, Michel (1966). Anisotropic Linear Magnetic Chain. *J Math Phys*, **7**, 1384.

des Cloizeaux, Jacques and Pearson, J J (1962). Spin–Wave Spectrum of the Antiferromagnetic Linear Chain. *Phys Rev*, **128**, 2131.

Dicke, Robert (1954). Coherence in Spontaneous Radiation Processes. *Phys Rev*, **93**, 99–110.

Douçot, Benoit and Zinn-Justin, Jean (ed.) (1995). *Strongly interacting Fermions and high T_c superconductivity*. Elsevier, Amsterdam.

Eckle, Hans-Peter and Hamer, Chris (1991). Non-analytic finite-size corrections for the Heisenberg chain in a magnetic field with free and twisted boundary conditions. *J Phys A: Math Gen*, **24**, 191.

Faddeev, Ludwig D (2006). History and Perspectives of Quantum Groups. *Milan J Math*, **74**, 279–294.

Faddeev, Ludwig D and Takhtajan, Leon A (1981). What is the Spin of a Spin Wave? *Phys Lett A*, **85**, 375.

Faddeev, Ludwig D and Takhtajan, Leon A (1984). Spectrum and scattering of excitations in the one–dimensional isotropic Heisenberg model. *J Math Sciences*, **24**, 241. Translated from the original Russian paper in: Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im V A Steklova AN SSSR **109** (1981) 134.

Fazekas, Patrik (1999). *Lecture Notes on Electron Correlation and Magnetism*. World Scientific, Singapore.

Fetter, Alexander L and Walecka, John Dirk (2003). *Quantum Theory of Many–Particle Systems*. Dover, Mineola.

Feynman, Richard P (1972). *Statistical Mechanics — A Set of Lectures*. Benjamin/Cummings, Reading, Massachusetts.

Feynman, Richard P (1988). *Asia–Pacific Physics News*, **3**, 22.

Ford, George W, Kac, Mark, and Mazur, Peter (1965). Statistical mechanics of assemblies of coupled oscillators. *J Math Phys*, **6**, 504.

Fowler, Michael (1987). Soluble Models. In *Nonlinearity in Condensed Matter* (ed. A. R.

Bishop, D. K. Campbell, P. Kumar, and S. E. Trullinger), Volume 69, Springer Series in Solid–State Science, p. 172. Springer, Berlin.

Fox, Mark (2006). *Quantum Optics — An Introduction*. Oxford University Press.

Fulton, Robert L and Gouterman, Martin (1961). Vibronic Coupling. I. Mathematical treatment for two electronic states. *J Chem Phys*, **35**, 1059–1071.

Gamow, George (1988). *One, Two, Three, . . . , Infinity*. Dover Publications, New York.

Garrison, John C and Chiao, Raymond Y (2008). *Quantum Optics*. Oxford University Press.

Gaudin, Michel (1971). Thermodynamics of the Heisenberg-Ising Ring for $\Delta \geq 1$. *Phys Rev Lett*, **26**, 1301.

Gaudin, Michel (1983). *La fonction d'onde de Bethe*. Masson, Paris.

Gaudin, Michel (2014). *The Bethe wavefunction*. Cambridge University Press.

Geller, Michael R and Cleland, Andrew N (2005). Superconducting qubits coupled to nanoelectromechanical resonators: An architecture for solid-state quantum-information processing. *Physical Review A*, **71**, 032311.

Gerry, Christopher C and Knight, Peter L (2005). *Introductory Quantum Optics*. Cambridge University Press.

Giamarchi, Thierry (2004). *Quantum Physics in One dimension*. Oxford University Press, Oxford.

Giauque, W F and Stout, J W (1936). The entropy of water and the third law of thermodynamics. the heat capacity of ice from 15 to 273°k. *J Am Chem Soc*, **58**(7), 1144–1150.

Giovannini, Bernard (2002). Cours de physique du solide avancée III & IV: Propriétés électronique des métaux. `http://dpmc.unige.ch/gr_giamarchi/Solides/Files/SolidIV.pdf`. Notes rédigées par Damien Stucki et Christophe Berthod, Année académique 2001/2002, Université de Genève, accessed 16 January 2014.

Gogolin, Alexander O, Nersesyan, Alexander A, and Tsvelik, Alexei M (1998). *Bosonization and Strongly Correlated Systems*. Cambridge University Press.

Goldenfeld, Nigel (1992). *Lectures on Phase Transitions and the Renormalization Group*. Addison-Wesley, Reading, Massachusetts.

Goodstein, David L (1985). *States of Matter*. Dover Publications, New York.

Gradshteyn, I S and Ryzhik, I M (1980). *Table of integrals, series, and products*. Academic Press, San Diego.

Griffiths, Robert B (1964*a*). Free Energy of the Antiferromagnetic Linear Chain. *Phys Rev*, **136**, A751.

Griffiths, Robert B (1964*b*). Magnetization Curve at Zero Temperature for the Antiferromagnetic Heisenberg Linear Chain. *Phys Rev*, **133**, A768.

Gunn, Michael (1988). Spin strings and superconductivity. *Physics World*, **1**, 31.

Gutzwiller, Martin C (1963). The effect of correlations on the ferromagnetism of transition metals. *Phys Rev Lett*, **10**, 159.

Haldane, F Duncan M (1980). General Relation of Correlation Exponents and Spectral Properties of One–Dimensional Fermi Systems: Application to an Anisotropic $S = \frac{1}{2}$ Heisenberg Chain. *Phys Rev Lett*, **45**, 1358–1362.

Haldane, F Duncan M (1981). 'Luttinger liquid theory' of one–dimensional quantum fluids: I. Properties of the Luttinger model and their extension to the general 1D interacting spinless Fermi gas. *J Phys C: Solid State Phys*, **14**, 2585–2609.

Haldane, F Duncan M (1988). Exact Jastrow–Gutzwiller resonating-valence-bond ground

state of the spin–$\frac{1}{2}$ antiferromagnetic Heisenberg chain with $1/r^2$ exchange. *Phys Rev Lett*, **60**, 635.

Halmos, Paul R (1978). *Measure Theory*. Springer, Berlin.

Hamer, Chris, Quispel, Reinout, and Batchelor, Murray (1987). Conformal anomaly and surface energy for Potts and Ashkin-Teller quantum chains. *J Phys A: Math Gen*, **20**, 5677.

Hamer, Chris J (1985). Finite–size corrections for ground states of the XXZ Heisenberg chain in the critical region. *J Phys A: Math Gen*, **18**, L1133–L1137.

Hamer, Chris J (1986). Finite–size corrections for ground states of the XXZ Heisenberg chain. *J Phys A: Math Gen*, **19**, 3335–3351.

Haroche, Serge and Raimond, Jean-Michel (2006). *Exploring the Quantum — Atoms, Cavities, and Photons*. Oxford University Press.

Heilmann, O J and Lieb, Elliott H (1970). Violation of the Non–Crossing Rule: the Hubbard Hamiltonian for Benzene. *Trans NY Acad Sci*, **33**, 116. (Also in Ann NY Acad Sci **172** (1971) 584.).

Heisenberg, Werner (1928). Zur Theorie des Ferromagnetismus. *Z Physik*, **49**, 619–636.

Hermann, Armin (1976). *Werner Heisenberg*. Rowohlt, Reinbeck.

Hubbard, John (1963). Electron correlations in narrow energy bands. *Proc Roy Soc (London)*, **A276**, 238.

Hulthén, Lamek (1938). Über das Austauschproblem eines Kristalles. *Ark Mat Astron Fys A*, **26**, 1.

Ising, Ernst (1925). Beitrag zur Theorie des Ferromagnetismus. *Z Physik*, **31**, 253.

Izyumov, Yurii A and Skryabin, Yurii N (1988). *Statistical Mechanics of Magnetically Ordered Systems*. Consultants Bureau, New York.

Jackson, John D (1999). *Classical Electrodynamics* (Third edn). John Wiley, New York.

Jaynes, Edwin T and Cummings, Frederick W (1963). Comparison of quantum and semiclassical radiation theories with application to the beam maser. *Proceedings of the IEEE*, **51**, 89–109.

Johnson, James D (1981). A survey of analytic results for 1–d heisenberg magnets. *J Appl Phys*, **52**, 1991.

Jordan, Pascal and Wigner, Eugene (1928). Über das Paulische Äquivalenzverbot. *Z Physik*, **47**, 631.

Judd, B R (1979). Exact solutions to a class of Jahn-Teller systems. *J Phys C*, **12**, 1685.

Kanamori, J (1963). Electron Correlation and Ferromagnetism of Transition Metals. *Prog Theor Phys*, **30**, 275.

Kaufman, Bruria (1949). Crystal Statistics. II. Partition Function Evaluated by Spinor Analysis. *Phys Rev*, **76**, 1232.

Kaufman, Bruria and Onsager, Lars (1949). Crystal Statistics. III. Short–Range Order in a Binary Ising Lattice. *Phys Rev*, **76**, 1244.

Kennedy, Thomas (2008). Introduction to Mathematical Physics, Chapter 1. `http://math.arizona.edu/~tgk/541/index.html`. Lecture notes of course Math 541, Department of Mathematics, University of Arizona (last accessed 10 February 2013).

Kivelson, Steven A, Rokhsar, Daniel S, and Sethna, James P (1987). Topology of the resonating valence–bond state: Solitons and high–T_c superconductivity. *Phys Rev B*, **35**, 8865.

Kolb, Edward W and Turner, Michael S (1989). *The Early Universe*. Addisson–Wesley, Redwood City.

Kubo, Ryogo (1988). *Statistical Mechanics — An Advanced Course with Problems and Solutions* (second edn). North–Holland, Amsterdam.

Kulish, Petr P and Sklyanin, Evgeni K (1991). The general $U_q[sl(2)]$ invariant XXZ integrable quantum spin chain. *J Phys A: Math Gen*, **24**, L435.

Landau, Lev D (1957a). The Theory of a Fermi Liquid. *Soviet Phys JETP*, **3**, 920–925.

Landau, Lev D (1957b). Oscillations in a Fermi Liquid. *Soviet Phys JETP*, **5**, 101–108.

Landau, Lev D (1959). On the Theory of the Fermi Liquid. *Soviet Phys JETP*, **8**, 70–74.

Laroche, Dominique, Gervais, Guillaume, Lilly, Michael P, and Reno, John L (2014). 1D–1D Coulomb Drag Signature of a Luttinger Liquid. *Science*, **343**, 631.

Laughlin, Robert B (1998). Nobel Prize Lecture. `http://large.stanford.edu/prizes/nobel/lecture/`. Overheads from Nobel Prize Lecture, Stanford University (last accessed 14 March 2013).

Laughlin, Robert B and Pines, David (2000). The theory of everything. *Proc. Nat. Acad. Sci.*, **97**, 28–31.

Leeuwen, Hendrika Johanna van (1921). Problème de la théorie électronique du magnétisme. *J Physique et le Radium*, **2**, 361– 377.

Lenz, Wilhelm (1920). Beitrag zum Verständnis der magnetischen Eigenschaften in festen Körpern. *Physikalische Zeitschrift*, **21**, 613.

Lieb, Elliott H (1963). Exact Analysis of an Interacting Bose Gas. II. The Excitation Spectrum. *Phys Rev*, **130**, 1616.

Lieb, Elliott H (1965). *The Bose fluid*, Volume VII C, Lectures in Theoretical Physics. University of Colorado Press.

Lieb, Elliott H (1967*a*). Exact solution of the problem of the entropy of two-dimensional ice. *Phys Rev Lett*, **18**, 692.

Lieb, Elliott H (1967*b*). Residual entropy of square ice. *Phys Rev*, **162**, 162.

Lieb, Elliott H and Liniger, Werner (1963). Exact Analysis of an Interacting Bose Gas. I. The General Solution and the Ground State. *Phys Rev*, **130**, 1605.

Lieb, Elliott H and Mattis, Daniel C (1962). Ordering Energy Levels of Interacting Spin Systems. *J Math Phys*, **3**, 749.

Lieb, Elliott H, Schultz, Theodore D, and Mattis, Daniel C (1961). Two Soluble Models of an Antiferromagnetic Chain. *Ann Phys*, **16**, 407.

Lieb, Elliott H and Wu, Fa-Yueh (1968). Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension. *Phys Rev Lett*, **20**, 1445.

Lieb, Elliott H and Wu, Fa-Yueh (2003). The one–dimensional Hubbard model: a reminiscence. *Physica A*, **321**, 1.

Luttinger, Joaquin M (1963). An Exactly Soluble Model of a Many–Fermion System. *J Math Phys*, **4**, 1154–1162.

Ma, Shang-keng (1976). *Modern Theory of Critical Phenomena*. Benjamin/Cummings, Reading, Massachusetts.

Madelung, Otfried (1978). *Introduction to Solid–State Theory*. Springer, Berlin.

Mahan, Gerald D (2000). *Many–Particle Physics* (Third edn). Kluwer, New York.

Marder, Michael P (2010). *Condensed Matter Physics* (Second edn). John Wiley, New York.

Marshall, W (1955). Antiferromagnetism. *Proc Roy Soc (London)*, **A232**, 48.

Mathews, Jon and Walker, Walker, Robert L (1970). *Mathematical Methods of Physics* (Second edn). Benjamin–Cummings.

Mattis, Daniel C (1993). *The Many–Body Problem — An Encyclopedia of Exactly Solved Models in One Dimension*. World Scientific, Singapore.

Mattis, Daniel C and Lieb, Elliott H (1965). Exact Solution of a Many–Fermion System and its Associated Boson Field. *J Math Phys*, **6**, 304–312.

Mezincescu, Luca and Nepomechie, Rafael I (1992). Introduction to the thermodynamics of spin chains. Proceedings of the NSERC-CAP Workshop on *Quantum Groups, Integrable Models and Statistical Systems*, Kingston, Canada.

Montroll, Elliott W, Potts, Renfrey B, and Ward, John C (1963). Correlations and Spontaneous Magnetization of the Two–Dimensional Ising Model. *J Math Phys*, **4**, 308.

Nagle, John F (1966). Lattice statistics of hydrogen bond crystals i. the residual entropy of ice. *J Math Phys*, **7**, 1484.

Nazarov, Yuli V and Danon, Jeroen (2013). *Advanced Quantum Mechanics*. Cambridge University Press.

Negele, John W and Orland, Henri (1998). *Quantum Many–Particle Systems*. Westview Press, Boulder.

Nepomechie, Raphael (1999). A spin chain primer. *Int J Mod Phys B*, **13**, 2973. `arXiv:hep-th/9810032v15Oct1998`.

Nielsen, Michael A and Chuang, Isaac L (2010). *Quantum computation and quantum information* (10th anniversary edn). Cambridge University Press, Cambridge.

Nienhuis, Bernard (2006). Brief summary of ensemble theory. Unpublished lecture notes.

Nishimori, Hidetoshi and Ortiz, Gerardo (2011). *Elements of Phase Transitions and Critical Phenomena*. Oxford University Press.

Onsager, Lars (1944). Crystal Statistics. I. A Two–Dimensional Model with an Order–Disorder Transition. *Phys Rev*, **65**, 117.

Onsager, Lars (1949). Discussion Remark. *Nuovo Cimento Suppl*, **6**, 261.

Orbach, Raymond (1958). Linear Antiferromagnetic Chain with Anisotropic Coupling. *Phys Rev*, **112**, 309.

Pariser, R and Parr, R G (1953). A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Hydrocarbons I and II. *J Chem Phys*, **21**, 466,767.

Pasquier, Vincent and Saleur, Hubert (1990). Common structures between finite systems and conformal field theories through quantum groups. *Nucl Phys B*, **330**, 523.

Pauling, Linus (1935). The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *J Am Chem Soc*, **57**(12), 2680–2684.

Pauling, Linus (1960). *The Nature of the Chemical Bond* (Third edn). Cornell University Press, Ithaca, New York.

Paz, Robert (1989). Feynman's Office: The Last Blackboards. *Physics Today*, **42**(2), 88. Special Issue: Richard Feynman.

Peierls, Rudolf (1936). On Ising's Model of Ferromagnetism. *Proc Cambridge Phil Soc*, **32**, 477.

Pfeuty, Pierre (1970). The one–dimensional Ising model with a transverse field. *Ann Phys*, **57**, 79.

Pople, J R (1953). Electron Interaction in Unsaturated Hydrocarbons. *Trans Faraday Soc*, **49**, 1375.

Rabi, Isidore I (1936). On the Process of Space Quantization. *Phys Rev*, **49**, 324–328.

Rabi, Isidore I (1937). Space Quantization in a Gyrating Magnetic Field. *Phys Rev*, **51**, 652–654.

Razumov, A V and Stroganov, Yu G (2001). Spin chains and combinatorics. *J Phys A: Math Gen*, **34**, 3185.

Rickayzen, G (1980). *Green's Functions and Condensed Matter*. Academic Press, London.

Roos, Bernard W (1969). *Analytic Functions and Distributions in Physics and Engineering*. John Wiley, New York.

Schoelkopf, Robert J and Girvin, Steven M (2008). Wiring up quantum systems. *Nature*, **451**, 664–669.

Schrödinger, Erwin (1935). Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, **23**, 807, 823, 844.

Schwabl, Franz (2006). *Statistical Mechanics* (Second edn). Springer, Berlin.

Sethna, James P (2006). *Statistical Mechanics — Entropy, Order Parameters, and Complexity*. Oxford University Press.

Shastry, B Sriram (1988). Exact solution of an $S = \frac{1}{2}$ Heisenberg antiferromagnetic chain with long-ranged interactions. *Phys Rev Lett*, **60**, 639.

Shenker, Stephen H (1982). Field theories and phase transitions. In *Recent advances in field theories and statistical mechanics* (ed. J.-B. Zuber and R. Stora), Les Houches Session XXXIX. North–Holland.

Simmons, George F (1992). *Calculus Gems*. McGraw–Hill, New York.

Sklyanin, Evgeni K (1988). Boundary conditions for integrable quantum systems. *J Phys A: Math Gen*, **21**, 2375.

Solé, Ricard V (2011). *Phase Transitions*. Princeton University Press.

Sólyom, Jenő (2010). *Fundamentals of the Physics of Solids — Normal, Broken-Symmetry, and Correlated Systems*. Volume 3. Springer, Berlin.

Sommerfeld, Arnold and Bethe, Hans (1933). *Elektronentheorie der Metalle*. Handbuch der Physik **34/2**. Springer, Berlin. Reprint 1967.

Sommerfeld, Arnold and Bethe, Hans (1967). *Elektronentheorie der Metalle*. Springer, Berlin. (Reprint of volume 24/2 of the Handbuch der Physik by Hans Geiger and Karl Scheel, 1933).

Sornette, Didier (2003). *Why Stock Markets Crash — Critical Events in Complex Financial Systems*. Princeton University Press.

Straumann, Norbert (2004). Cosmological phase transitions. `arXiv:astro-ph/0409042`. Invited lecture at the third Summer School on Condensed Matter Research, 7–14 August 2004, Zuoz, Switzerland.

Stroganov, Yu G (2001). The importance of being odd. *J Phys A: Math Gen*, **34**, L179–L185.

Sutherland, Bill (1985). An introduction to the Bethe ansatz. In *Exactly solvable problems in condensed matter and relativistic field theory* (ed. B. S. Shastry, S. S. Jha, and V. Singh), Number 242 in Lecture Notes in Physics, p. 1. Springer, Berlin.

Sutherland, Bill (2004). *Beautiful Models — 70 Years of Exactly Solved Quantum Many–Body Problems*. World Scientific, Singapore.

Takahashi, Minoru (1971*a*). One–Dimensional Heisenberg Model at Finite Temperature. *Prog Theo Phys*, **46**, 401.

Takahashi, Minoru (1971*b*). Thermodynamics of the Heisenberg–Ising model for $|\Delta| < 1$ in one dimension. *Phys Lett A*, **36**, 325.

Takahashi, Minoru (1997). *Thermodynamical Bethe Ansatz and Condensed Matter*, pp. 204–250. Conformal Field Theories and Integrable Models: Lectures Held at the Eötvös University Graduate Course, Budapest, Hungary, 1996, Lecture Notes in Physics 498, Zalán Horváth and László Palla (Editors). Springer, Berlin.

Takahashi, Minoru (1999). *Thermodynamics of One–Dimensional Solvable Models*. Cambridge University Press.

Takahashi, Minoru and Suzuki, Masuo (1972). One–Dimensional Anisotropic Heisenberg Model at Finite Temperatures. *Prog Theo Phys*, **48**, 2187.

Takhtajan, Leon A (1985). Introduction to algebraic bethe ansatz. In *Exactly Solvable Problems in Condensed Matter and Relativistic Field Theory* (ed. B. S. Shastry, S. S. Jha, and V. Singh), Lecture Notes in Physics 242. Springer, Berlin.

Tavis, Michael and Cummings, Frederick (1968). Exact Solution for an N–Molecule–Radiation–Field Hamiltonian. *Physical Review*, **170**, 379–384.

Tavis, Michael and Cummings, Frederick (1969). Approximate Solutions for an N-Molecule-Radiation-Field Hamiltonian. *Physical Review*, **188**, 692.

Tian, Lin (2011). Cavity cooling of a mechanical resonator in the presence of a two-level-system defect. *Physical Review B*, **84**, 035417.

Tomonaga, Sin-itiro (1950). Remarks on Bloch's Method of Sound Waves applied to Many–Fermion Problems. *Prog Theor Phys*, **5**, 544–569.

Tong, David (2012). Statistical Physics. `http://www.damtp.cam.ac.uk/user/tong/statphys.html`. Lecture notes, University of Cambridge.

Truong, Tuong T (1987). Six–vertex model. unpublished lecture notes.

Tsvelik, Alexei M and Wiegmann, Paul B (1983). Exact results in the theory of magnetic alloys. *Adv Phys*, **32**, 453.

Voit, Johannes (1994). One–dimensional fermi liquids. *Rep Prog Phys*, **57**, 977–1116.

Voit, Johannes (2005). *The Statistical Mechanics of Financial Markets* (Third edn). Springer, Berlin.

Walker, L R (1959). Antiferromagnetic Linear Chain. *Phys Rev*, **116**, 1089.

Wen, Xiao-Gang (2004). *Quantum field theory of many–body systems*. Oxford University Press.

Woynarovich, Ferenc (1982*a*). On the eigenstates of a Heisenberg chain with complex wavenumbers not forming strings. *J Phys C: Solid State Phys*, **15**, 6397.

Woynarovich, Ferenc (1982*b*). On the $S^z = 0$ excited states of an anisotropic Heisenberg chain. *J Phys A: Math Gen*, **15**, 2985.

Woynarovich, Ferenc (1997). *Introduction to coordinate–space Bethe ansatz*, pp. 151–203. Conformal Field Theories and Integrable Models: Lectures Held at the Eötvös University Graduate Course, Budapest, Hungary, 1996, Lecture Notes in Physics 498, Zalán Horváth and László Palla (Editors). Springer, Berlin.

Wyld, Henry W (1999). *Mathematical Methods for Physics* (Second edn). Perseus.

Yamada, Tomoji (1969). Fermi–Liquid Theory of Linear Antiferromagnetic Chains. *Prog Theo Phys*, **41**, 880.

Yang, Chen Ning (1952). The Spontaneous Magnetization of a Two–Dimensional Ising Model. *Phys Rev*, **85**, 808.

Yang, Chen Ning (1967). Some Exact Results for the Many-Body Problem in one Dimension with Repulsive Delta-Function Interaction. *Phys Rev Lett*, **19**, 1312.

Yang, Chen Ning and Yang, C P (1966*a*). One-Dimensional Chain of Anisotropic Spin-Spin Interactions. I. Proof of Bethe's Hypothesis for Ground State in a Finite System. *Phys Rev*, **150**, 321.

Yang, Chen Ning and Yang, C P (1966*b*). One-Dimensional Chain of Anisotropic Spin-Spin Interactions. II. Properties of the Ground-State Energy Per Lattice Site for an Infinite System. *Phys Rev*, **150**, 327.

Yang, Chen Ning and Yang, C P (1966*c*). One-Dimensional Chain of Anisotropic Spin-Spin Interactions. III. Applications. *Phys Rev*, **151**, 258.

Yang, Chen Ning and Yang, C P (1969). Thermodynamics of a One–Dimensioanl System of Bosons with Repulsive Delta–Function Interaction. *J Math Phys*, **10**, 1115.