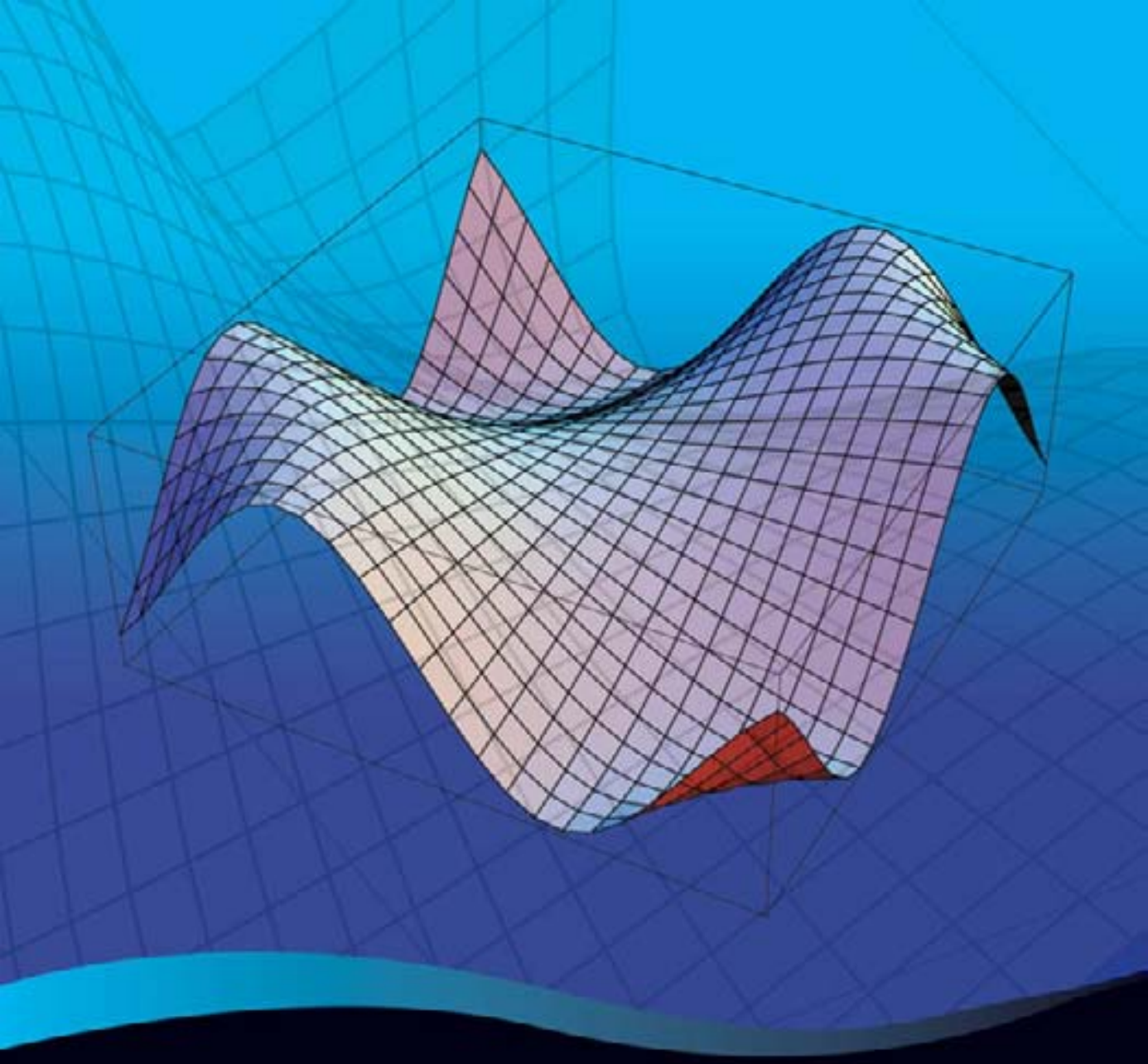


Statistical Field Theory

An Introduction to Exactly Solved Models in Statistical Physics

Giuseppe Mussardo

OXFORD GRADUATE TEXTS

Statistical Field Theory

This page intentionally left blank

Statistical Field Theory

An Introduction to Exactly Solved Models in Statistical Physics

Giuseppe Mussardo
International School of Advanced Studies, Trieste

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2010

British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Mussardo, G.
Statistical field theory : an introduction to exactly solved models in statistical
physics / Giuseppe Mussardo.
p. cm.—(Oxford graduate texts)
ISBN 978–0–19–954758–6 (hardback)
1. Field theory (Physics)—Statistical methods. I. Title.
QA173.7.M87 2009
530.14–dc22
2009026995

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by
CPI Antony Rowe, Chippenham, Wiltshire

ISBN 978–0–19–954758–6 (Hbk.)

1 3 5 7 9 10 8 6 4 2

Ulrich thought that the general and the particular are nothing but two faces of the same coin.

Robert Musil, *Man without Quality*

This page intentionally left blank

Preface

This book is an introduction to statistical field theory, an important subject of theoretical physics that has undergone formidable progress in recent years. Most of the attractiveness of this field comes from its profound interdisciplinary nature and its mathematical elegance; it sets outstanding challenges in several scientific areas, such as statistical mechanics, quantum field theory, and mathematical physics.

Statistical field theory deals, in short, with the behavior of classical or quantum systems consisting of an enormous number of degrees of freedom. Those systems have different phases, and the rich spectrum of the phenomena they give rise to introduces several questions: What is their ground state in each phase? What is the nature of the phase transitions? What is the spectrum of the excitations? Can we compute the correlation functions of their order parameters? Can we estimate their finite size effects? An ideal guide to the fascinating area of phase transitions is provided by a remarkable model, the Ising model.

There are several reasons to choose the Ising model as a pathfinder in the field of critical phenomena. The first one is its simplicity – an essential quality to illustrate the key physical features of the phase transitions, without masking their derivation with worthless technical details. In the Ising model, the degrees of freedom are simple boolean variables $\sigma_{\vec{i}}$, whose values are $\sigma_{\vec{i}} = \pm 1$, defined on the sites $\vec{i}$ of a d-dimensional lattice. For these essential features, the Ising model has always played an important role in statistical physics, both at the pedagogical and methodological levels.

However, this is not the only reason of our choice. The simplicity of the Ising model is, in fact, quite deceptive. Despite its apparent innocent look, the Ising model has shown an extraordinary ability to describe several physical situations and has a remarkable theoretical richness. For instance, the detailed analysis of its properties involves several branches of mathematics, quite distinguished for their elegance: here we mention only combinatoric analysis, functions of complex variables, elliptic functions, the theory of nonlinear differential and integral equations, the theory of the Fredholm determinant and, finally, the subject of infinite dimensional algebras. Although this is only a partial list, it is sufficient to prove that the Ising model is an ideal playground for several areas of pure and applied mathematics.

Equally rich is its range of physical aspects. Therefore, its study offers the possibility to acquire a rather general comprehension of phase transitions. It is time to say a few words about them: phase transitions are remarkable collective phenomena, characterized by sharp and discontinous changes of the physical properties of a statistical system. Such discontinuities typically occur at particular values of the external parameters (temperature or pressure, for instance); close to these critical values, there is a divergence of the mean values of many thermodynamical quantities, accompanied by anomalous fluctuations and power law behavior of correlation functions. From an experimental point of view, phase transitions have an extremely rich phenomenology, ranging from the superfluidity of certain materials to the superconductivity of others,

from the mesomorphic transformations of liquid crystals to the magnetic properties of iron. Liquid helium He^4, for instance, shows exceptional superfluid properties at temperatures lower than $T_c = 2.19\,K$, while several alloys show phase transitions equally remarkable, with an abrupt vanishing of the electrical resistance for very low values of the temperature.

The aim of the theory of phase transitions is to reach a general understanding of all the phenomena mentioned above on the basis of a few physical principles. Such a theoretical synthesis is made possible by a fundamental aspect of critical phenomena: their universality. This is a crucial property that depends on two basic features: the internal symmetry of the order parameters and the dimensionality of the lattice. In short, this means that despite the differences that two systems may have at their microscopic level, as long as they share the two features mentioned above, their critical behaviors are surprisingly identical.[1] It is for these universal aspects that the theory of phase transitions is one of the pillars of statistical mechanics and, simultaneously, of theoretical physics. As a matter of fact, it embraces concepts and ideas that have proved to be the building blocks of the modern understanding of the fundamental interactions in Nature. Their universal behavior, for instance, has its natural demonstration within the general ideas of the renormalization group, while the existence itself of a phase transition can be interpreted as a spontaneously symmetry breaking of the hamiltonian of the system. As is well known, both are common concepts in another important area of theoretical physics: quantum field theory (QFT), i.e. the theory that deals with the fundamental interactions of the smallest constituents of the matter, the elementary particles.

The relationship between two theories that describe such different phenomena may appear, at first sight, quite surprising. However, as we will see, it will become more comprehensible if one takes into account two aspects: the first one is that both theories deal with systems of infinite degrees of freedom; the second is that, close to the phase transitions, the excitations of the systems have the same dispersion relations as the elementary particles.[2] Due to the essential identity of the two theories, one should not be surprised to discover that the two-dimensional Ising model, at temperature T slightly away from T_c and in the absence of an external magnetic field, is equivalent to a fermionic neutral particle (a Majorana fermion) that satisfies a Dirac equation. Similarly, at $T = T_c$ but in the presence of an external magnetic field B, the two-dimensional Ising model may be regarded as a quantum field theory with eight scalar particles of different masses.

The use of quantum field theory – i.e. those formalisms and methods that led to brilliant results in the study of the fundamental interactions of photons, electrons, and all other elementary particles – has produced remarkable progress both in the understanding of phase transitions and in the computation of their universal quantities. As will be explained in this book, our study will significantly benefit from such a possibility: since phase transitions are phenomena that involve the long distance scales of

[1] This becomes evident by choosing an appropriate combination of the thermodynamical variables of the two systems.

[2] The explicit identification between the two theories can be proved by adopting for both the path integral formalism.

the systems – the infrared scales – the adoption of the continuum formalism of field theory is not only extremely advantageous from a mathematical point of view but also perfectly justified from a physical point of view. By adopting the QFT approach, the discrete structure of the original statistical models shows itself only through an ultraviolet microscopic scale, related to the lattice spacing. However, it is worth pointing out that this scale is absolutely necessary to regularize the ultraviolet divergencies of quantum field theory and to implement its renormalization.

The main advantage of QFT is that it embodies a strong set of constraints coming from the compatibility of quantum mechanics with special relativity. This turns into general relations, such as the completeness of the multiparticle states or the unitarity of their scattering processes. Thanks to these general properties, QFT makes it possible to understand, in a very simple and direct way, the underlying aspects of phase transitions that may appear mysterious, or at least not evident, in the discrete formulation of the corresponding statistical model.

There is one subject that has particularly improved thanks to this continuum formulation: this is the set of two-dimensional statistical models, for which one can achieve a classification of the fixed points and a detailed characterization of their classes of universality. Let us briefly discuss the nature of the two-dimensional quantum field theories.

Right at the critical points, the QFTs are massless. Such theories are invariant under the conformal group, i.e. the set of geometrical transformations that implement a scaling of the length of the vectors while preserving their relative angle. But, in two dimensions conformal transformations coincide with mappings by analytic functions of a complex variable, characterized by an infinite-dimensional algebra known as a Virasoro algebra. This enables us to identify first the operator content of the models (in terms of the irreducible representations of the Virasoro algebra) and then to determine the exact expressions of the correlators (by solving certain linear differential equations). In recent years, thanks to the methods of conformal field theory, physicists have reached the exact solutions of a huge number of interacting quantum theories, with the determination of all their physical quantities, such as anomalous dimensions, critical exponents, structure constants of the operator product expansions, correlation functions, partition functions, etc.

Away from criticality, quantum field theories are, instead, generally massive. Their analysis can often be carried out only by perturbative approaches. However, there are some favorable cases that give rise to integrable models of great physical relevance. The integrable models are characterized by the existence of an infinite number of conserved charges. In such fortunate circumstances, the exact solution of the off-critical models can be achieved by means of S-matrix theory. This approach makes it possible to compute the exact spectrum of the excitations and the matrix elements of the operators on the set of these asymptotic states. Both these data can thus be employed to compute the correlation functions by spectral series. These expressions enjoy remarkable convergence properties that turn out to be particularly useful for the control of their behaviors both at large and short distances. Finally, in the integrable cases, it is also possible to study the exact thermodynamical properties and the finite size effects of the quantum field theories. Exact predictions for many universal quantities

can also be obtained. For the two-dimensional Ising model, for instance, there are two distinct integrable theories, one corresponding to its thermal perturbation (i.e. $T \neq T_c$, $B = 0$), the other to the magnetic deformation ($B \neq 0$, $T = T_c$). In the last case, a universal quantity is given, for instance, by the ratio of the masses of the lowest excitations, expressed by the famous golden ratio $m_2/m_1 = 2\cos(\pi/5) = (\sqrt{5}+1)/2$.

In addition to their notable properties, the exact solution provided by the integrable theories is an important step towards the general study of the scaling region close to the critical points. In fact, they permit an efficient perturbative scheme to study non-integrable effects, in particular to follow how the mass spectrum changes by varying the coupling constants. Thanks to this approach, new progress has been made in understanding several statistical models, in particular the class of universality of the Ising model by varying the temperature T and the magnetic field B. Non-integrable field theories present an extremely interesting set of new physical phenomena, such as confinement of topological excitations, decay processes of the heavier particles, the presence of resonances in scattering processes, or false vacuum decay, etc. The analytic control of such phenomena is one of the most interesting results of quantum field theory in the realm of statistical physics.

This book is a long and detailed journey through several fields of physics and mathematics. It is based on an elaboration of the lecture notes for a PhD course, given by the author at the International School for Advances Studies (Trieste). During this elaboration process, particular attention has been paid to achieving a coherent and complete picture of all surveyed topics. The effort done to emphasize the deep relations among several areas of physics and mathematics reflects the profound belief of the author in the substantial unity of scientific knowledge.

This book is designed for students in physics or mathematics (at the graduate level or in the last year of their undergraduate courses). For this reason, its style is greatly pedagogical; it assumes only some basis of mathematics, statistical physics, and quantum mechanics. Nevertheless, we count on the intellectual curiosity of the reader.

Structure of the Book

In this book many topics are discussed at a fairly advanced level but using a pedagogical approach. I believe that a student could highly profit from some exposure to such treatments.

The book is divided in four parts.

Part I: **Preliminary notions** (Chapters 1, 2, and 3)

The first part deals with the fundamental aspects of phase transitions, illustrated by explicit examples coming from the Ising model or similar systems.

Chapter 1: a straighforward introduction of essential ideas on second-order phase transitions and their theoretical challenge. Our attention focuses on some important issues, such as order parameters, correlation length, correlation functions, scaling behavior, critical exponents, etc. A short discussion is also devoted to the Ising model and its most significant developments during the years of its study. The chapter also contains two appendices, where all relevant results of classical and statistical mechanics are summarized.

Chapter 2: this deals with one-dimensional statistical models, such as the Ising model and its generalizations (Potts model, systems with $O(n)$ or Z_n symmetry, etc.). Several methods of solution are discussed: the recursive method, the transfer matrix approach or series expansion techniques. General properties of these methods – valid on higher dimensional lattices – are also enlighted. The contents of this chapter are quite simple and pedagogical but extremely useful for understanding the rest of the book. One of the appendices at the end of the chapter is devoted to a famous problem of topology, i.e. the four-color problem, and its relation with the two-dimensional Potts model.

Chapter 3: here we discuss the approximation schemes to approach lattice statistical models that are not exactly solvable. In addition to the mean field approximation, we also consider the Bethe–Peierls approach to the Ising model. Moreover, there is a thorough discussion of the gaussian model and its spherical version – two important systems with several points of interest. In one of the appendices there is a detailed analysis of the random walk on different lattices: apart from the importance of the subject on its own, it is shown that the random walk is responsible for the critical properties of the spherical model.

Part II: **Two-dimensional lattice models** (Chapters 4, 5, and 6)

This part provides a general introduction to the key ideas of equilibrium statistical mechanics of discrete systems.

Chapter 4: at the beginning of this chapter there is the Peierls argument (it permits us to prove the existence of a phase transition in the two-dimensional Ising model). The rest of the chapter deals with the duality transformations that link the low- and the high-temperature phases of several statistical models. Particularly important is the proof of the so-called *star–triangle identity*. This identity will be crucial in the later discussion of the transfer matrix of the Ising model (Chapter 6).

Chapter 5: two exact combinatorial solutions of the two-dimensional Ising model are the key topics of this chapter. Although no subsequent topic depends on them, both the mathematical and the physical aspects of these solutions are elegant enough to deserve special attention.

Chapter 6: this deals with the exact solution of the two-dimensional Ising model achieved through the transfer matrix formalism. A crucial role is played by the commutativity properties of the transfer matrices, which lead to a functional equation for their eigenvalues. The exact free energy of the model and its critical point can be identified by means of the lowest eigenvalue. We also discuss the general structure of the Yang–Baxter equation, using the six-vertex model as a representative example.

Part III: **Quantum field theory and conformal invariance** (Chapters 7–14)

This is the central part of the book, where the aims of quantum field theory and some of its fundamental results are discussed. A central point is the bootstrap method of conformal field theories. The main goal of this part is to show the extraordinary efficiency of these techniques for the analysis of critical phenomena.

Chapter 7: the main reasons for adopting the methods of quantum field theory to study the critical phenomena are emphasized here. Both the canonical quantization and the path integral formulation of the field theories are presented, together with the analysis of the perturbation theory. Everything in this chapter will be needed sooner or later, since it highlights most of the relevant aspects of quantum field theory.

Chapter 8: the key ideas of the renormalization group are introduced here. They involve the scaling transformations of a system and their implementations in the space of the coupling constants. From this analysis, one gets to the important notion of relevant, irrelevant and marginal operators and then to the universality of the critical phenomena.

Chapter 9: a crucial aspect of the Ising model is its fermionic nature and this chapter is devoted to this property of the model. In the continuum limit, a Dirac equation for neutral Majorana fermions emerges. The details of the derivation are

much less important than understanding why it is possible. The simplicity and the exactness of the result are emphasized.

Chapter 10: this chapter introduces the notion of conformal transformations and the important topic of the massless quantum field theories associated to the critical points of the statistical models. Here we establish the important conceptual result that the classification of all possible critical phenomena in two dimensions consists of finding out all possible irreducible representations of the Virasoro algebra.

Chapter 11: the so-called minimal conformal models, characterized by a finite number of representations, are discussed here. It is shown that all correlation functions of these models satisfy linear differential equations and their explicit solutions are given by using the Coulomb gas method. Their exact partition functions can be obtained by enforcing the modular invariance of the theory.

Chapter 12: free theories are usually regarded as trivial examples of quantum systems. This chapter proves that this is not the case of the conformal field theories associated to the free bosonic and fermionic fields. The subject is not only full of beautiful mathematical identities but is also the source of deep physical concepts with far reaching applications.

Chapter 13: the conformal transformations may be part of a larger group of symmetry and this chapter discusses several of their extensions: supersymmetry, Z_n transformations, and current algebras. In the appendix the reader can find a self-contained discussion on Lie algebras.

Chapter 14: the identification of a class of universality is one of the central questions in statistical physics. Here we discuss in detail the class of universality of several models, such as the Ising model, the tricritical Ising model, and the Potts model.

Part IV: **Away from criticality** (Chapters 15–21)

This part of the book develops the analysis of the statistical models away from criticality.

Chapter 15: here is introduced the notion of *the scaling region* near the critical points, identified by the deformations of the critical action by means of the relevant operators. The renormalization group flows that originate from these deformations are subjected to important constraints, which can be expressed in terms of sum rules. This chapter also discusses the nature of the perturbative series based on the conformal theories.

Chapter 16: the general properties of the integrable quantum field theories are the subject of this chapter. They are illustrated by means of significant examples, such as the Sine–Gordon model or the Toda field theories based on the simple roots of a Lie algebra. For the deformations of a conformal theory, it is shown how to set up an efficient counting algorithm to prove the integrability of the corresponding model.

Chapter 17: this deals with the analytic theory of the S-matrix of the integrable models. Particular emphasis is put on the dynamical principle of the *bootstrap*, which gives rise to a recursive structure of the amplitudes. Several dynamical quantities, such as mass ratios or three-coupling constants, have an elegant mathematic formulation, which also has an easy geometrical interpretation.

Chapter 18: the Ising model in a magnetic field is one of the most beautiful example of an integrable model. In this chapter we present its exact S-matrix and the exact spectrum of its excitations, which consist of eight particles of different masses. Similarly, we discuss the exact scattering theory behind the thermal deformation of the tricritical Ising model and the unusual features of the exact S-matrix of the non-unitary Yang-Lee model. Other important examples are provided by $O(n)$ invariant models: when $n = 2$, one obtains the important case of the Sine–Gordon model. We also discuss the quantum-group symmetry of the Sine–Gordon model and its reductions.

Chapter 19: the thermodynamic Bethe ansatz permits us to study finite size and finite temperature effects of an integrable model. Here we derive the integral equations that determine the free energy and we give their physical interpretation.

Chapter 20: at the heart of a quantum field theory are the correlation functions of the various fields. In the case of integrable models, the correlators can be expressed in terms of the spectral series based on the matrix elements on the asymptotic states. These matrix elements, also known as form factors, satisfy a set of functional and recursive equations that can be exactly solved in many cases of physical interest.

Chapter 21: this chapter introduces a perturbative technique based on the form factors to study non-integrable models. Such a technique permits the computation of the corrections to the mass spectrum, the vacuum energy, the scattering amplitudes, and so on.

Problems Each chapter of this book includes a series of problems. They have different levels of difficulty: some of them relate directly to the essential material of the chapters, other are instead designed to introduce new applications or even new topics. The problems are an integral part of the course and their solution is a crucial step for the understanding of the whole subject.

Mathematical aspects Several chapters have one or more appendices devoted to some mathematical aspects encountered in the text. Far from being a collection of formulas, these appendices aim to show the profound relationship that links mathematics and physics. Quite often, they also give the opportunity to achieve comprehension of mathematical results by means of physical intuition. Some appendices are also devoted to put certain ideas in their historical perspective in one way or another.

References At the end of each chapter there is an annotated bibliography. The list of references, either books or articles, is by no means meant to be a comprehensive

survey of the present literature. Instead it is meant to guide the reader a bit deeper if he/she wishes to go on. It also refers to the list of material consulted in preparing the chapters. There are no quotations of references in the text, except for a few technical points.

Acknowledgements

Over the years I have had the pleasure of collaborating and discussing many of the themes of this book with several colleagues and friends.

First of all, I would like especially to thank Gesualdo Delfino for the long and profitable collaboration on the two-dimensional quantum field theory, and for sharing his deep understanding of many aspects of the theory. Similarly, I would like to thank John Cardy: his extraordinary scientific vision has been over the years a very precious guide. I have also a special debt with Adam Schwimmer and Vladimir Rittenberg, for their constant encouragement and for arousing enthusiam to face any scientific themes always with great perspicacity and smartness. I am also particularly grateful to Aliosha Zamolodchikov, with whom I have had the privilege to discuss many important topics and to enjoy his friendship.

I have been fortunate in having the benefit of collaboration and discussion with numerous collegues who have generously shared their insights, in particular Olivier Babelon, Denis Bernard, Andrea Cappelli, Ed Corrigan, Boris Dubrovin, Patrick Dorey, Fabian Essler, Paul Fendley, Vladimir Kravstov, Andre' LeClair, Alexander Nersesyan, Paul Pearce, Hubert Saleur, Giuseppe Santoro, Kareljan Schoutens, Fedor Smirnov, Sasha Zamolodchikov, and Jean-Bernard Zuber. I would also like to mention and thank my collaborators Carlo Acerbi, Daniel Cabra, Filippo Colomo, Alessandro De Martino, Davide Fioravanti, Anne Koubek, Marco Moriconi, Paola Mosconi, Alessandro Mossa, Silvia Penati, Alessandro Silva, Prospero Simonetti, Galen Sotkov, Roberto Tateo, and Valentina Riva.

Contents

Part I

Preliminary Notions

This page intentionally left blank

1
Introduction

La sapienza è figliola della sperienza.

Leonardo da Vinci, *Codice Forster III, 14 recto*

In this chapter we introduce some general concepts of statistical mechanics and phase transitions, in order to give a rapid overview of the different topics of the subject and their physical relevance. For the sake of clarity and simplicity, we will focus our attention on magnetic systems but it should be stressed that the concepts discussed here are of a more general nature and can be applied to other systems as well. We will analyze, in particular, the significant role played by the *correlation length* in the phase transitions and the important properties of *universality* observed in those phenomena. As we will see, near a phase transition the thermodynamic quantities of a system present an anomalous power law behavior, parameterized by a set of *critical exponents*. The universal properties showed by phase transitions is manifested by the exact coincidence of the critical exponents of systems that share the same symmetry of their hamiltonian and the dimensionality of their lattice but may be, nevertheless, quite different at a microscopic level. From this point of view, the study of phase transitions consists of the classification of all possible *universality classes*. This important property will find its full theoretical justification in the context of the renormalization group ideas, a subject that will be discussed in one of the following chapters.

In this chapter we will also introduce the Ising model and recall the most significant progress in the understanding of its features: (i) the duality transformation found by H.A. Kramers and G.H. Wannier for the partition function of the bidimensional case in the absence of a magnetic field; (ii) the exact solution of the lattice model given by L. Onsager; and (iii) the exact solution provided by A.B. Zamolodchikov (with methods borrowed from quantum field theory) of the bidimensional Ising model in a magnetic field at the critical value T_c of the temperature.

In the appendices at the end of the chapter one can find the basic notions of the various ensembles used in statistical mechanics, both at the classical and quantum level, with a discussion of their physical properties.

1.1 Phase Transitions

1.1.1 Competitive Principles

The atoms of certain materials have a magnetic dipole, due either to the spin of the orbital electrons or to the motion of the electrons around the nucleus, or to both of

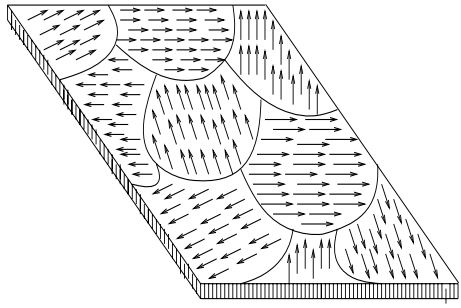


Fig. 1.1 *Magnetic domains for* $T > T_c$.

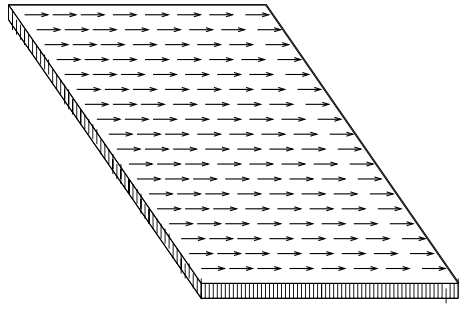


Fig. 1.2 *Alignment of the spins for* $T < T_c$.

them. In many materials, the magnetic dipoles of the atoms are randomly oriented and the total magnetic field produced by them is then zero, as in Fig. 1.1. However, in certain compounds or in substances like iron or cobalt, for the effect of the interactions between the atomic dipoles, one can observe a macroscopic magnetic field different from zero (Fig. 1.2). In those materials, which are called *ferromagnetic*, this phenomenon is observed for values of the temperature less than a critical value T_c, known as the *Curie temperature*, whose value depends on the material in question. At $T = T_c$ these materials undergo a *phase transition*, i.e. there is a change of the physical properties of the system: in our example, this consists of a spontaneous magnetization on macroscopic scales, created by the alignment of the microscopic dipoles.

The occurrence of a phase transition is the result of two competitive instances: the first tends to minimize the energy while the second tends to maximize the entropy.

- **Principle of energy minimization**
  In ferromagnetic materials, the configuration of the magnetic dipoles of each atom (which we denote simply as *spins*) tend to minimize the total energy of the system. This minimization is achieved when all spins are aligned. The origin of the atomic dipole, as well as their interaction, is due to quantum effects. In the following, however, we focus our attention on the classical aspects of this problem, i.e. we

will consider as given the interaction among the spins, and those as classical degrees of freedom. In this framework, the physical problem can be expressed in a mathematical form as follows: first of all, to each spin, placed at the site i of a d-dimensional lattice, is associated a vector $\vec{S}_i$; secondly, their interaction is described by a hamiltonian $\mathcal{H}$. The simplest version of these hamiltonians is given by

$$\mathcal{H} = -\frac{J}{2} \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j, \tag{1.1.1}$$

where $J > 0$ is the coupling constant and the notation $\langle ij \rangle$ stands for a sum to the neighbor spins. The lowest energy configurations are clearly those in which all spins are aligned along one direction.

If the minimization of the energy was the only principle that the spins should follow, we would inevitably observe giant magnetic fields in many substances. The reason why this does not happen is due to another competitive principle.

- **Principle of entropy maximization**
  Among the extraordinarily large number of configurations of the system, the ones in which the spins align with each other along a common direction are quite special. Hence, unless a great amount of energy is needed to orientate, in a different direction, spins that are at neighbor sites, the number of configurations in which the spins are randomly oriented is much larger that the number of the configurations in which they are completely aligned. As is well known, the measure of the disorder in a system is expressed by the *entropy* S: if we denote by $\omega(E)$ the number of states of the system at energy E, its definition is given by the Boltzmann formula

$$S(E) = k \log \omega(E), \tag{1.1.2}$$

  where k is one of the fundamental constants in physics, known as *the Boltzmann constant*.

  If the tendency to reach the status of maximum disorder was the only physical principle at work, clearly we could never observe any system with a spontaneous magnetization.

Classification scheme of phase transitions
In the modern classification scheme, phase transitions are divided into two broad categories: first-order and second-order phase transitions. First-order phase transitions are those that involve a latent heat. At the transition point, a system either absorbs or releases a fixed amount of energy, while its temperature stays constant.

First-order phase transitions are characterized by a finite value of the correlation length. In turn, this implies the presence of a mixed-phase regime, in which some parts of the system have completed the transition and others have not. This is what happens, for instance, when we decrease the temperature of water to its freezing value T_f: the water does not instantly turn into ice, but forms a mixture of water and ice domains. The presence of a latent heat signals that the structure of the material is drastically changing at $T = T_f$: above T_f, there is no crystal lattice and the water molecules can wander around in a disordered path, while below T_f there is the lattice of ice crystals, where the molecules are packed into a face-centered cubic lattice. In addition to the phase transition of water, many other important phase transitions fall into this category, including Bose–Einstein condensation.

The second class of phase transitions consists of the continuous phase transitions, also called second-order phase transitions. These have no associated latent heat and they are also characterized by the divergence of the correlation length at the critical point. Examples of second-order phase transitions are the ferromagnetic transition, superconductors, and the superfluid transition. Lev Landau was the first to set up a phenomenological theory of second-order phase transitions. Several transitions are also known as infinite-order phase transitions. They are continuous but break no symmetries. The most famous example is the Kosterlitz–Thouless transition in the two-dimensional XY model. Many quantum phase transitions in two-dimensional electron gases also belong to this class.

As the example of the magnetic dipoles has shown, the macroscopic physical systems in which there is a very large number of degrees of freedom are subjected to two different instances: one that tends to order them to minimize the energy, the other that tends instead to disorder them to maximize the entropy. However, to have real competition between these two different tendencies, one needs to take into account another important physical quantity, i.e. the *temperature* of the system. Its role is determined by the laws of statistical mechanics.

1.1.2 Partition Function

One of the most important advances witnessed in nineteenth century physics has been the discovery of the exact probabilistic function that rules the microscopic configurations of a system at equilibrium. This is a fundamental law of statistical mechanics.[1] To express such a law, let us denote by $\mathcal{C}$ a generic state of the system (in our example, a state is specified once the orientation of each magnetic dipole is known). Assume that the total number N of the spins is sufficiently large (we will see that a phase transition may occur only when $N \to \infty$). Moreover, assume that the system is at thermal

[1] In the following we will mainly be concerned with the laws of classical statical mechanics. Moreover, we will use the formulation of statistical mechanics given by the *canonical ensemble*. The different ensembles used in statistical mechanics, both in classical and quantum physics, can be found in the appendix of this chapter.

equilibrium, namely that the spins and the surrounding environment exchange energy at a common value T of the temperature. Within these assumptions, the probability that a given configuration $\mathcal{C}$ of the system is realized, is given by the Boltzmann law

$$P[\mathcal{C}] = \frac{e^{-E(\mathcal{C})/kT}}{Z}, \tag{1.1.3}$$

where $E(\mathcal{C})$ is the energy of the configuration $\mathcal{C}$ while T is the absolute temperature. A common notation is $\beta = 1/kT$. The expectation value of any physical observable $\mathcal{O}$ is then expressed by the statistical average on all configurations, with weights given by the Boltzmann law

$$\langle \mathcal{O} \rangle \; = \; Z^{-1} \sum_{\mathcal{C}} \mathcal{O}(\mathcal{C})\, e^{-\beta E(\mathcal{C})}. \tag{1.1.4}$$

The quantity Z in the denominator is the *partition function* of the system, defined by

$$Z(N, \beta) \; = \; \sum_{\mathcal{C}} e^{-\beta\, E(\mathcal{C})}. \tag{1.1.5}$$

It ensures the proper normalization of the probabilities, $\sum_{\mathcal{C}} P[\mathcal{C}] = 1$. For its own definition, this quantity contains all relevant physical quantities of the statistical system at equilibrium. By making a change of variable, it can be expressed as

$$\begin{aligned} Z(N, \beta) \quad &= \quad \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} = \sum_{E} \omega(E)\, e^{-\beta E} = \sum_{E} e^{-\beta E + \log \omega(E)} \\ &= \sum_{E} e^{\beta [TS - E]} \; \equiv e^{-\beta\, F(N,\beta)}, \end{aligned} \tag{1.1.6}$$

where $F(N, \beta)$ is the *free energy* of the system. This is an extensive quantity, related to the *internal energy* $U = \langle H \rangle$ and the *entropy* $S = -\langle \left(\frac{\partial F}{\partial T}\right)_N \rangle$ by the thermodynamical relation

$$F \; = \; U - T\, S. \tag{1.1.7}$$

Namely, we have

$$\begin{aligned} \langle U \rangle \; &= \; \frac{\partial}{\partial \beta} (\beta F) \\ \\ \langle S \rangle \; &= \; \beta^2\, \frac{\partial F}{\partial \beta}. \end{aligned} \tag{1.1.8}$$

The extensive property of F comes from the definition of $Z(N, \beta)$, because if the system is made of two weakly interacting subsystems, $Z(N, \beta)$ is given by the product of their partition functions.[2] The proof of eqn (1.1.7) is obtained starting with the

[2]This is definitely true if the interactions are short-range, as we assume hereafter. In the presence of long-range forces the situation is more subtle and the extensivity property of the free energy may be violated.

identity

$$\sum_{\mathcal{C}} e^{\beta[F(N,\beta)-E(\mathcal{C})]} = 1.$$

Taking the derivative with respect to β of both terms we have

$$\sum_{\mathcal{C}} e^{\beta[F(N,\beta)-E(\mathcal{C})]} \left[F(N,\beta) - E(\mathcal{C}) + \beta \left(\frac{\partial F}{\partial \beta} \right)_N \right] = 0,$$

i.e. precisely formula (1.1.7). This equation enables us to easily understand the occurrence of different phases in the system by varying the temperature. In fact, moving T, there is a different balance in the free energy between the entropy (that favors disorder) and the energy (that privileges there order). Therefore may exist a critical value $T = T_c$ at which there is a perfect balance between the two different instances. To distinguish in a more precise way the phases of a system it is necessary to introduce the important concept of *order parameter*.

1.1.3 Order Parameters

To characterize a phase transition we need an order parameter, i.e. a quantity that has a vanishing thermal average in one phase (typically the high-temperature phase) and a non-zero average in the other phases. Hence, such a quantity characterizes the onset of order at the phase transition. It is worth stressing that there is no general procedure to identify the proper order parameter for each phase transition. Its definition may require, in fact, a certain amount of skill or ingenuity. There is, however, a close relation between the order parameter of a system and the symmetry properties of its hamiltonian. In the example of the magnetic dipoles discussed so far, a physical quantity that has a zero mean value for $T > T_c$ and a finite value for $T < T_c$ is the total magnetization, $\vec{\mathcal{M}} = \sum_i \vec{S}_i$. Hence, a local order parameter for such a system is identified by the vector $\vec{S}_i$ since we have

$$\langle \vec{S}_i \rangle = \begin{cases} 0; & T > T_c \\ \vec{\mathcal{S}}_0 \neq 0; & T < T_c. \end{cases} \tag{1.1.9}$$

When the system is invariant under translations, the mean value of the spin is the same for all sites.

For what concerns the symmetry properties, it is easy to see that the hamiltonian (1.1.1) is invariant under an arbitrary global rotation $\mathcal{R}$ of the spins. As is well known, the set of rotations forms a group. In the case of vectors with three components,[3] the group is denoted by $SO(3)$ and is isomorphic to the group of orthogonal matrices 3×3 with determinant equal to 1, with the usual rule of multiplication of matrices.

In the range $T > T_c$, there is no magnetization and the system does not have any privileged direction: in this phase the symmetry of its hamiltonian is perfectly respected. Vice versa, when $T < T_c$, the system acquires a special direction, identified by the vector $\vec{\mathcal{S}}_0 = \langle \vec{S}_i \rangle$ along which the majority of the spins are aligned. In this

[3] It will become useful to generalize this example to the situation in which the spins are made up of n components. In this case the corresponding symmetry group is denoted by $SO(n)$.

case, the system is in a phase which has less symmetry of its hamiltonian and one says that a *spontaneously symmetry breaking* has taken place. More precisely, in this phase the symmetry of the system is restricted to the subclass of rotations along the axis identified by the vector $\vec{\mathcal{S}}_0$, i.e. to the group $SO(2)$. One of the tasks of the theory of phase transitions is to provide an explanation for the phenomenon of spontaneously symmetry breaking and to study its consequences.

1.1.4 Correlation Functions

The main source of information on phase transitions comes from scattering experiments. They consist of the study of scattering processes of some probe particles sent to the system (they can be photons, electrons, or neutrons). In liquid mixtures, near the critical point, the fluid is sufficiently hot and diluted that the distinction between the liquid and gaseous phases is almost non-existent. The phase transition is signaled by the remarkable phenomenon of critical opalescence, a milky appearance of the liquid, due to density fluctuations at all possible wavelengths and to the anomalous diffusion of light.[4] For magnetic systems, neutrons provide the best way to probe these systems: first of all, they can be quite pervasive (so that one can neglect, to a first approximation, their multiple scattering processes) and, secondly, they couple directly to the spins of the magnetic dipoles. The general theory of the scattering processes involves in this case the *two-point correlation function* of the dipoles

$$G^{(2)}(\vec{i},\vec{j}) \;=\; \langle \vec{S}_i \cdot \vec{S}_j \rangle. \tag{1.1.10}$$

When there is a translation invariance, this function depends on the distance difference $\vec{i}-\vec{j}$. Moreover, if the system is invariant under rotations, the correlator is a function of the absolute value of the distance $r = \mid \vec{i}-\vec{j} \mid$ between the two spins, so that $G^{(2)}(\vec{i},\vec{j}) = G^{(2)}(r)$. Strictly speaking, any lattice is never invariant under translations and rotations but we can make use of these symmetries as long as we analyze the system at distance scales much larger than the lattice spacing a.

As is evident by its own definition, $G^{(2)}(r)$ measures the degree of the relative alignment between two spins separated by a distance r. Since for $T < T_c$ the spins are predominantly aligned along the same direction, to study their fluctuations it is convenient to subtract their mean value, defining the *connected correlation function*

$$G_c^{(2)}(r) \;=\; \langle (\vec{S}_i - \vec{\mathcal{S}}_0)\cdot(\vec{S}_j - \vec{\mathcal{S}}_0)\rangle \;=\; \langle \vec{S}_i \cdot \vec{S}_j\rangle - \mid \vec{\mathcal{S}}_0 \mid^2 . \tag{1.1.11}$$

When $T > T_c$, the mean value of the spin vanishes and $G_c^{(2)}(r)$ coincides with the original definition of $G^{(2)}(r)$.

Nearby spins usually tend to be correlated. Away from the critical point, $T \neq T_c$, their correlation extends to a certain distance ξ, called the *correlation length*. This is the typical size of the regions in which the spins assume the same value, as shown in Fig. 1.3. The correlation length can be defined more precisely in terms of the

[4] Smoluchowski and Einstein were the first to understand the reason of this phenomenon: the fluctuations in the density of the liquid produce anologous fluctuations in its refraction index. In particular, Einstein showed how these fluctuations can be computed and pointed out their anomalous behavior near the critical point.

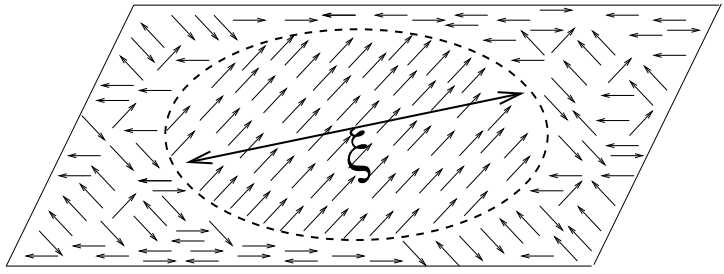


Fig. 1.3 *The scale of the magnetic domains is given by the correlation length* $\xi(T)$.

asymptotic behavior of the correlation function[5]

$$G_c^{(2)}(r) \simeq e^{-r/\xi}, \qquad r \gg a, \qquad T \neq T_c. \tag{1.1.12}$$

At the critical point $T = T_c$, there is a significant change in the system and the two-point correlation function takes instead a power law behavior

$$G_c^{(2)}(r) \simeq \frac{1}{r^{d-2+\eta}}, \qquad r \gg a, \qquad T = T_c. \tag{1.1.13}$$

The parameter η in this formula is the *anomalous dimension* of the order parameter. This is the first example of *critical exponents*, a set of quantities that will be discussed thoroughly in the next section. The power law behavior of $G_c^{(2)}(r)$ clearly shows that, at the critical point, fluctuations of the order parameter are significantly correlated on all distance scales. Close to a phase transition, the correlation length diverges:[6] denoting by t the relative displacement of the temperature from the critical value, $t = (T - T_c)/T_c$, one observes that, near the Curie temperature, ξ behaves as (see Fig. 1.4)

$$\xi(T) \;=\; \begin{cases} \xi_+ \; t^{-\nu}, & T > T_c; \\ \xi_- \; (-t)^{-\nu}, & T < T_c, \end{cases} \tag{1.1.14}$$

where ν is another critical exponent.

The two different behaviors of the correlation functions – at the critical point and away from it – can be summed up in a single expression

$$G_c^{(2)}(r) \;=\; \frac{1}{r^{d-2+\eta}} \, f\left(\frac{r}{\xi}\right). \tag{1.1.15}$$

This formula involves the *scaling function* $f(x)$ that depends only on the dimensionless ratio $x = r/\xi$. For large x, this function has the asymptotic behavior $f(x) \sim e^{-x}$, while its value at $x = 0$ simply fixes the normalization of this quantity, which can always be

[5]This asymptotic behavior of the correlator can be deduced by quantum field theory methods, as shown in Chapter 8.

[6]This is the significant difference between a phase transition of second order and one of first order. In phase transitions of first order the correlation length is finite also at the critical point.

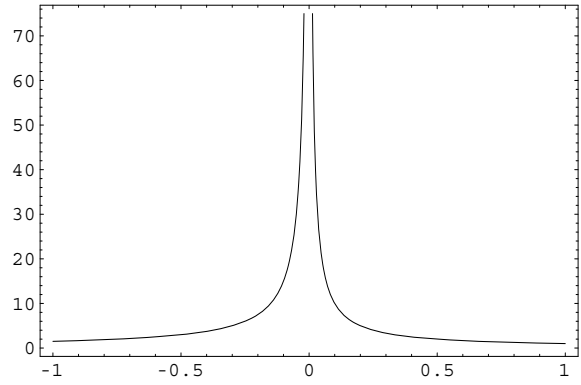


Fig. 1.4 *Behavior of the correlation length as a function of the temperature near* $T = T_c$.

chosen as $f(0) = 1$. It is worth stressing that the temperature enters the correlation functions only through the correlation length $\xi(T)$.

Aspects of phase transitions. It is now useful to stop and highlight the aspects of phase transitions that have emerged so far. The most important property is that, at $T = T_c$, the fluctuations of the order parameter extend significantly to the entire system, while they are exponentially small away from the critical point. This means that the phase transition taking place at T_c is the result of an extraordinary collective phenomenon that involves all the spins of the system at once.

This observation poses the obvious theoretical problem to understand how the short-range interactions of the spins can give rise to an effective interaction that extends to the entire system when $T = T_c$. There is also another consideration: if one regards the correlation length ξ as a measure of the effective degrees of freedom involved in the dynamics, its divergence at the critical point implies that the study of the phase transitions cannot be faced with standard perturbative techniques. Despite these apparent difficulties, the study of phase transitions presents some conceptual simplifications that are worth underlining. The first simplification concerns the *scale invariance* present at the critical point, namely the symmetry under a dilatation of the length-scale

$$a \;\to\; \lambda\, a.$$

Under this transformation, the distance between two points of the system gets reduced as

$$r \;\to\; r/\lambda.$$

The correlation function (1.1.13), thanks to its power law behavior, is invariant under this transformation as long as the order parameter transforms as

$$\vec{S} \;\to\; \lambda^{(d-2+\eta)/2}\, \vec{S}. \tag{1.1.16}$$

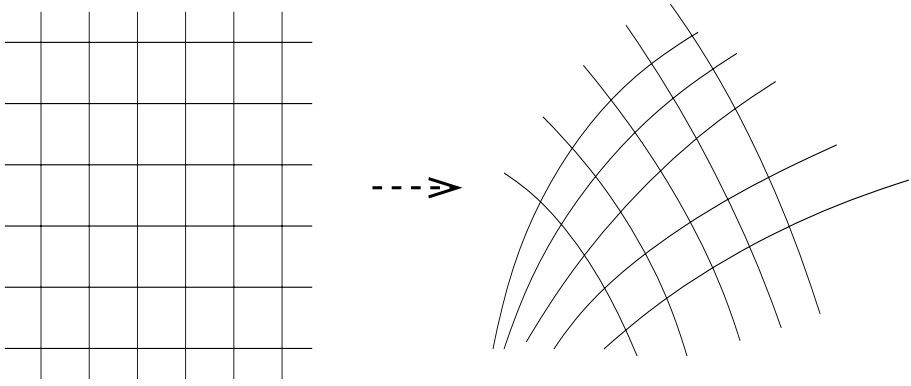


Fig. 1.5 *Conformal transformation. It leaves invariant the angles between the lines.*

Expressed differently, at the critical point there is complete equivalence between a change of the length-scale and the normalization of the order parameter. The divergence of the correlation length implies that the system becomes insensitive to its microscopic scales[7] and becomes scale invariant. Moreover, in Chapter 11 we will prove that, under a set of general hypotheses, the global dilatation symmetry expressed by the transformation $a \to \lambda\, a$ can be further extended to the local transformations $a \to \lambda(\vec{x})\, a$ that change the lengths of the vectors but leave invariant their relative angles. These are the *conformal transformations* (see Fig. 1.5). Notice that in the two-dimensional case, the conformal transformations coincide with the mappings provided by the analytic functions of a complex variable: studying the irreducible representations of the associated infinite dimensional algebra, one can reach an exact characterization of the bidimensional critical phenomena.

The second simplification – strictly linked to the scaling invariance of the critical point – is the *universality* of phase transitions. It is an experimental fact that physical systems of different nature and different composition often show the same critical behavior: it is sufficient, in fact, that they share the same symmetry group $\mathcal{G}$ of the hamiltonian and the dimensionality of the lattice space. Hence the critical properties are amply independent of the microscopic details of the various interactions, so that the phenomenology of the critical phenomena falls into different *classes of universality*. Moreover, thanks to the insensitivity of the microscopic details, one can always characterize a given class of universality by studying its simplest representative. We will see later on that all these remarkable universal properties find their elegant justification in the *renormalization group* formulation. In the meantime, let's go on and complete our discussion of the anomalous behavior near the critical point by introducing other critical exponents.

[7] Although the system has fluctuations on all possible scales, it is actually impossible to neglect completely the existence of a microscopic scale. In the final formulation of the theory of the phase transitions this scale is related to the renormalization of the theory and, as a matter of fact, is responsible of the anomalous dimension of the order parameter.

1.1.5 Critical Exponents

Close to a critical point, the order parameter and the response functions of a statistical system show anomalous behavior. Directly supported by a large amount of experimental data, these anomalous behaviors are usually expressed in terms of power laws, whose exponents are called *critical exponents*. In addition to the quantities η and ν previously defined, there are other critical exponents directly related to the order parameter. To define them, it is useful to couple the spins to an external magnetic field $\vec{B}$

$$\mathcal{H} = -\frac{J}{2}\sum_{\langle ij\rangle}\vec{S}_i\cdot\vec{S}_j - \vec{B}\cdot\sum_i \vec{S}_i. \tag{1.1.17}$$

To simplify the notation, let's assume that $\vec{B}$ is along the z axis, with its modulus equals B. In the presence of B, there is a net magnetization of the system along the z axis with a mean value given by[8]

$$\mathcal{M}(B,T) = \langle S_i^z\rangle \equiv \frac{1}{Z}\sum_{\mathcal{C}} S_i^z\, e^{-\beta\mathcal{H}} = -\frac{\partial F}{\partial B}. \tag{1.1.18}$$

The *spontaneous magnetization* is a function of T alone, defined by

$$M(T) = \lim_{B\to 0}\mathcal{M}(B,T), \tag{1.1.19}$$

and its typical behavior is shown in Fig. 1.6. Near T_c, M has an anomalous behavior, parameterized by the critical exponent β

$$M = M_0(-t)^{\beta}, \tag{1.1.20}$$

where $t = (T - T_c)/T_c$.

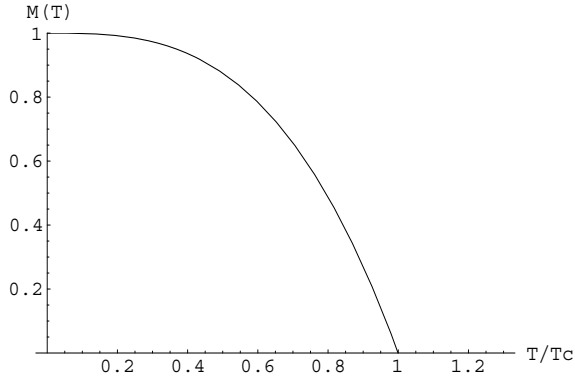


Fig. 1.6 *Spontaneous magnetization versus temperature.*

[8]By translation invariance, the mean value is the same for all spins of the system.

Another critical exponent δ is defined by the anomalous behavior of the magnetization when the temperature is kept fixed at the critical value T_c but the magnetic field is different from zero

$$\mathcal{M}(B, T_c) \;=\; \mathcal{M}_0 B^{1/\delta} \;. \tag{1.1.21}$$

The *magnetic susceptibility* is the response function of the system when we switch on a magnetic field

$$\chi(B, T) \;=\; \frac{\partial \mathcal{M}(B, T)}{\partial B}. \tag{1.1.22}$$

This quantity presents a singularity at the critical point, expressed by the critical exponent γ

$$\chi(0, T) \;=\; \begin{cases} \chi_+ \; t^{-\gamma}, & T > T_c; \\ \chi_- \; (-t)^{-\gamma}, & T < T_c. \end{cases} \tag{1.1.23}$$

Finally, the last critical exponent that is relevant for our example of a magnetic system is associated to the critical behavior of the specific heat. This quantity, defined by

$$C(T) \;=\; \frac{\partial U}{\partial T}, \tag{1.1.24}$$

has a singularity near the Curie temperature parameterized by the exponent α

$$C(T) \;=\; \begin{cases} C_+ \; t^{-\alpha}, & T > T_c; \\ C_- \; (-t)^{-\alpha}, & T < T_c. \end{cases} \tag{1.1.25}$$

A summary of the critical exponents of a typical magnetic system is given in Table 1.1.

The critical exponents assume the same value for all statistical systems that belong to the same universality class while, varying the class of universality, they change correspondingly. Hence they are important fingerprints of the various universality classes. In Chapter 8 we will see that the universality classes can also be identified by the so-called *universal ratios*. These are dimensionless quantities defined in terms of the various response functions: simple examples of universal ratios are given by ξ_+/ξ_i, χ_+/χ_-, or C_+/C_-. Other universal ratios will be defined and analyzed in Chapter 8.

Let's end our discussion of the critical behavior with an important remark: a statistical system can present a phase transition (i.e. anomalous behavior of its free energy and its response functions) only in its thermodynamic limit $N \to \infty$, where N is the number of particles of the system. Indeed, if N is finite, the partition function is a

Table 1.1: *Definition of the critical exponents.*

Exponent	Definition	Condition
α	$C \sim \mid T - T_c \mid^{-\alpha}$	$B = 0$
β	$M \sim (T_c - T)^{-\beta}$	$T < T_c, \;\; B = 0$
γ	$\chi \sim \mid T - T_c \mid^{-\gamma}$	$B = 0$
δ	$B \sim \mid M \mid^{\delta}$	$T = T_c$
ν	$\xi \sim \mid T - T_c \mid^{-\nu}$	$B = 0$
η	$G_c^{(2)} \sim r^{-(d-2+\eta)}$	$T = T_c$

regular function of the temperature, without singular points at a finite value of T, since it is expressed by a sum of a finite number of terms.

1.1.6 Scaling Laws

The exponents α, β, γ, δ, η, and ν, previously defined, are not all independent. Already at the early stage of the study on phase transitions, it was observed that they satisfy the algebraic conditions[9]

$$\begin{aligned} &\alpha + 2\beta + \gamma = 2; \\ &\alpha + \beta\,\delta + \beta = 2; \\ &\nu(2-\eta) = \gamma; \\ &\alpha + \nu\, d = 2, \end{aligned} \tag{1.1.26}$$

so that it is sufficient to determine only two critical exponents in order to fix all the others.[10] Moreover, the existence of these algebraic equations suggests that the thermodynamic quantities of the system are functions of B and T in which these variables enter only *homogeneous combinations*, i.e. they satisfy scaling laws.

An example of a scaling law is provided by the expression of the correlator, eqn (1.1.15). It is easy to see that this expression, together with the divergence of the correlation length (1.1.14), leads directly to the third equation in (1.1.26). To prove this, one needs to use a general result of statistical mechanics, known as the *fluctuation-dissipation theorem*, that permits us to link the response function of an external field (e.g. the magnetic susceptibility) to the connected correlation function of the order parameter coupled to such a field. For the magnetic susceptibility, the fluctuation-dissipation theorem leads to the identity

$$\begin{aligned} \chi = \frac{\partial \mathcal{M}(B,T)}{\partial B} &= \frac{\partial}{\partial B}\left[\frac{1}{Z}\sum_{\mathcal{C}} S_i^z e^{-\beta\mathcal{H}}\right] \\ &= \beta \sum_j \left(\langle S_j^z S_i^z\rangle - |\,\langle S_i^z\rangle\,|^2\right) = \beta \sum_r G_c^{(2)}(r), \end{aligned} \tag{1.1.27}$$

which can be derived by using eqns (1.1.17) and (1.1.5) for the hamiltonian and the partition function, together with the definition of the mean value, given by eqn (1.1.4).

Substituting in (1.1.27) the scaling law (1.1.15) of the correlation function, one has

$$\begin{aligned} \chi = \beta \sum_r G_c^{(2)}(r) \;&=\; \beta \sum_r \frac{1}{r^{d-2+\eta}} f\left(\frac{r}{\xi}\right) \\ &\simeq \int dr\, r^{d-1}\, \frac{1}{r^{d-2+\eta}} f\left(\frac{r}{\xi}\right) = A\,\xi^{2-\eta}, \end{aligned} \tag{1.1.28}$$

[9] The last of these equations, which involves the dimensionality d of the system, generally holds for d less of d_c, known as *upper critical dimensions*.

[10] As discussed in Chapter 8, the critical exponents are not the most fundamental theoretical quantities. As a matter of fact, they can all be derived by a smaller set of data given by the *scaling dimensions* of the relevant operators.

where A is a constant given by the value of the integral obtained by the substitution $r \to \xi z$

$$A = \int dz\, z^{1-\eta}\, f(z).$$

Using the anomalous behavior of $\xi(t)$ given by (1.1.14), we have

$$\chi \simeq \xi^{2-\eta} \simeq t^{-\nu(2-\eta)}, \tag{1.1.29}$$

and, comparing to the anomalous behavior of χ expressed by (1.1.23), one arrives at the relation

$$\nu(2-\eta) = \gamma.$$

A scaling law can be similarly written for the singular part of the free energy $F_s(B,T)$, expressed by a homogeneous function of the two variables

$$F_s(B,T) = t^{2-\alpha}\, \mathcal{F}\left(\frac{B}{t^{\beta\delta}}\right). \tag{1.1.30}$$

It is easy to see that this expression implies the relation

$$\alpha + \beta\delta + \beta = 2, \tag{1.1.31}$$

i.e. the second equation in (1.1.26). In fact, the magnetization is given by the derivative of the free energy $F_s(B,T)$ with respect to B

$$M = \left.\frac{\partial F_s}{\partial B}\right|_{B=0} = t^{2-\alpha-\beta\delta}\, F'(0).$$

Comparing to eqn (1.1.20), one recovers eqn (1.1.31). Scaling relations for other thermodynamic quantities can be obtained in a similar way.

In Chapter 8 we will see that the homegeneous form assumed by the thermodynamic quantities in the vicinity of critical points has a theoretical justification in the renormalization group equations that control the scaling properties of the system.

1.1.7 Dimensionality of the Space and the Order Parameters

Although the world in which we live is three-dimensional, it is however convenient to get rid from this slavery and to consider instead the dimensionality d of the space as a variable like any other. There are various reasons to adopt this point of view.

The first reason is of a phenomenological nature: there are many systems that, by the particular nature of their interactions or their composition, present either one-dimensional or two-dimensional behavior. Systems that can be considered one-dimensional are those given by long chains of polymers, for instance; in particular if the objects of study are the monomers along the chain. Two-dimensional systems are given by those solids composed of weakly interacting layers, as happens in graphite. Another notable example of a two-dimensional system is provided by the quantum Hall effect, where the electrons of a thin metallic bar are subjected to a strong magnetic field in the vertical direction at very low temperatures. Examples of two-dimensional

critical phenomena are also those relative to surface processes of absorption or phenomena that involve the thermodynamics of liquid films.

It is necessary to emphasize that the effective dimensionality shown by critical phenomena can depend on the thermodynamic state of the system. Namely there could be a dimensional transmutation induced by the variation of the thermodynamic parameters, such as the temperature: there are materials that in some thermodynamic regimes appear as if they were bidimensional, while in other regimes they have instead a three-dimensional dynamics. Consider, for instance, a three-dimensional magnetic system in which the interaction along the vertical axis J_z is much smaller than the interaction J among the spins of the same plane, i.e. $J_z \ll J$. In the high-temperature phase (where the correlation length $\xi(T)$ is small), one can neglect the coupling between the next neighbor planes, so that the system appears to be a two-dimensional one. However, decreasing the temperature, the correlation length $\xi(T)$ increases and, in each plane, there will be large areas in which the spins become parallel and behave as a single spin but of a large value. Hence, even though the coupling J_z between the planes was originally small, their interaction can be quite strong for the large values of the effective dipoles; correspondingly, the system presents at low temperatures a three-dimensional behavior.

There is, however, a more theoretical reason to regard the dimensionality d of a system as an additional parameter. First of all, the existence of a phase transition of a given hamiltonian depends on the dimensionality of the system. The fluctuations become stronger by decreasing d and, because they disorder the system, the critical temperature decreases correspondingly. Each model with a given symmetry selects a *lower critical dimension* d_i such that, for $d < d_i$ its phase transition is absent. For the Ising model (and, more generally, for all models with a discrete symmetry) $d_i = 1$. For systems with a continuous symmetry, the fluctuations can disorder the system much more easily, since the order parameter can change its value continuously without significantly altering the energy. Hence, for many of these systems we have $d_i = 2$.

The critical exponents depend on d and, for each system, there is also a *higher critical dimension* d_s: for $d > d_s$, the critical exponents take the values obtained in the mean field approximation that will be discussed in Chapter 3. For the Ising model, we have $d_s = 4$. The range

$$d_i < d < d_s$$

of a given system is therefore the most interesting interval of dimensions, for it is the range of d in which one observes the strongly correlated nature of the fluctuations. This is another reason to regard d as a variable of statistical systems. In fact, the analysis of their critical behavior usually deals with divergent integrals coming from the large fluctuations of the critical point. To regularize such integrals, a particularly elegant method is provided by the so-called *dimensional regularization*, as discussed in a problem at the end of the chapter. This method permits us, in particular, to define an expansion parameter $\epsilon = d - d_s$ and to express the critical exponents in power series in ϵ. Further elaboration of these series permits us to obtain the critical exponents for finite values of ϵ, i.e. those that correspond to the actual value of d for the system under consideration.

1.2 The Ising Model

After the discussion on the phenomenology of the phase transitions of the previous section, let us now introduce the Ising model. This is the simplest statistical model that has a phase transition. The reason to study this model comes from two different instances: the first is the need to simplify the nature of the spins in order to obtain a system sufficiently simple to be solved exactly, while the second concerns the definition of a model sufficiently realistic to be compared with the experimental data.

The simplification is obtained by considering the spins σ_i as scalar quantities with values ± 1 rather than the vector quantities $\vec{S}_i$ previously introduced. In this way, the hamiltonian of the Ising model is given by

$$\mathcal{H} = -\frac{J}{2} \sum_{\langle i,j \rangle} \sigma_i \, \sigma_j \; - \; B \sum_i \sigma_i, \quad \sigma_i = \pm 1. \tag{1.2.1}$$

When $B = 0$, it has a global discrete symmetry Z_2, implemented by the transformation $\sigma_i \to -\sigma_i$ on all the spins.

Even though the Ising model may appear as a caricature of actual ferromagnetic substances, it has nevertheless a series of advantages: it is able to provide useful information on the nature of phase transition, on the effects of the cooperative dynamics, and on the role of the dimensionality d of the lattice. In the following chapters we will see, for instance, that the model has a phase transition at a finite value T_c of the temperature when $d \geq 2$ while it does not have any phase transition when $d = 1$. Moreover, the study of this model helps to clarify the aspects of the phase transitions that occur in lattice gases or, more generally, in all those systems in which the degrees of freedom have a binary nature.

The elucidation of the mathematical properties of the Ising model has involved a large number of scientists since 1920, i.e. when it was originally introduced by Wilhelm Lenz.[11] The first theoretical results are due to Ernst Ising, a PhD student of Lenz at the University of Hamburg, who in 1925 published a short article based on his PhD studies in which he showed the absence of a phase transition in the one-dimensional case. Since then, the model has been known in the literature as the Ising model.

After this first result, it is necessary to reach 1936 to find ulterior progress in the understanding of the model. In that year, using an elementary argument, R. Peierls showed the existence of a critical point in the two-dimensional case, so that the Ising model became a valid and realistic tool for investigating phase transitions. The exact value of the critical temperature T_c on a two-dimensional square lattice was found by H.A. Kramers and G.H. Wannier in 1941, making use of an ingenious technique. They showed that the partition function of the model can be expressed in a systematic way as a series expansion both in the high- and in the low-temperature phases, showing that the two series were related by a duality transformation. In more detail, in the high-temperature phase the variable entering the series expansion is given by βJ,

[11] One has only to read Ising's original paper to learn that the model was previously proposed by Ising's research supervisor, Wilhelm Lenz. It is rather curious that Lenz's priority has never been recognized by later authors. Lenz himself apparently never made any attempt later on to claim credit for suggesting the model and also never published any papers on it.

while in the low-temperature phase the series is in powers of the variable $e^{-2\beta J}$. The singularity present in both series, together with the duality relation that links one to the other, allowed them to determine the exact value of the critical temperature of the model on the square lattice, given by the equation $\sinh(J/kT_c) = 1$.

The advance of H.A. Kramers and G.H. Wannier was followed by the fundamental contribution of Lars Onsager, who announced at a meeting of the New York Academy of Science, on 28 February 1943, the solution for the partition function of the two-dimensional Ising model at zero magnetic field. The details were published two years later. The contribution of Onsager constitutes a milestone in the field of phase transitions. The original solution of Onsager, quite complex from a mathematical point of view, has been simplified with the contribution of many authors and, in this respect, it is important to mention B. Kaufman and R.J. Baxter. Since then, there have been many other results concerning several aspects, such as the analysis of different two-dimensional lattices, the computation of the spontaneous magnetization, the magnetic susceptibility and, finally, the correlation functions of the spins. In 1976, B. McCoy, T.T. Wu, C. Tracy, and E. Barouch, in a remarkable theoretical *tour de force*, showed that the correlation functions of the spins can be determined by the solution of a nonlinear differential equation, known in the literature as the Painleve' equation. A similar result was also obtained by T. Miwa, M. Jimbo, and their collaborators in Kyoto: in particular, they showed that the monodromy properties of a particular class of differential equation can be analyzed by using the spin correlators of the Ising model.

In the years immediately after the solution proposed by Onsager, in the community of researchers there was considerable optimism of being able to extend his method to the three-dimensional lattice as well as to the bidimensional case but in the presence of an external magnetic field. However, despite numerous efforts and numerous attempts that finally proved to be premature or wrong, for many years only modest progress has been witnessed on both the arguments.

An exact solution of the three-dimensional case is still unknown, although many of its properties are widely known thanks to numerical simulations and series expansions – methods that have been improved during the years with the aid of faster and more efficient computers. The critical exponents or the equations of state, for instance, are nowadays known very accurately and their accuracy increases systematically with new publications on the subject. It is a common opinion among physicists that the exact solution of the three-dimensional Ising model is one of the most interesting open problems of theoretical physics.

The analysis of the two-dimensional Ising model in the presence of a magnetic field has received, on the contrary, a remarkable impulse since 1990, and considerable progress in the understanding of its properties has been witnessed. This development has been possible thanks to methods of quantum field theory and the analytic S-matrix, which have been originally proposed in this context by Alexander Zamolodchikov. By means of these methods it was possible to achieve the exact determination of the spectrum of excitations of the Ising model in a magnetic field and the identification of their interactions. Subsequently G. Delfino and G. Mussardo determined the two-point correlation function of the spins of the Ising model in a magnetic field while Delfino and Simonetti calculated the correlation functions that involve the energy operator of the

model. In successive work, G. Delfino, G. Mussardo, and P. Simonetti systematically studied the properties of the model by varying the magnetic field and the temperature. This analysis was further refined in a following paper by P. Fonseca and A. B. Zamolodchikov, which led to the thorough study of the analytic structure of the free energy in the presence of a magnetic field and for values of temperature different from the critical value. Besides these authors, many others have largely contributed to the developments of the subject and, in the sequel, there will be ample possibility to give them proper credit.

In the following chapters we will discuss the important aspects of the Ising model and its generalizations. In doing so, we will emphasize their physical properties and to put in evidence their mathematical elegance. As we will see, this study will bring us face to face with many important arguments of theoretical physics and mathematics.

Ernst Ising

The Ising model is one of the best known models in statistical mechanics, as is confirmed by the 12 000 articles published on it or referring to it from 1969 to 2002. Therefore it may appear quite paradoxical that the extraordinary notoriety of the model is not accompanied by an analogous notoriety of the scientist to whom the model owes its name. The short biographical notes that follow underline the singular history, entangled with the most dramatic events of the twentieth century, of this humble scientist who became famous by chance and remained unaware of his reputation for many years of his life.

Ernst Ising was born in Cologne on the 10 May 1900. His family, of Jewish origin, moved later to Bochum in Westfalia where Ernst finished his high school studies. In 1919 he started his university studies at Goettingen in mathematics and physics and later he moved to Hamburg. Here, under the supervision of Wilhelm Lenz, he started the study of the ferromagnetic model proposed by Lenz. In 1925 he defended his PhD thesis, devoted to the analysis of the one-dimensional case of the model that nowadays bears his name, and in 1926 he published his results in the journal *Zeitschrift fur Phyisk*. After his PhD, Ising moved to Berlin and during the years 1925 and 1926 he worked at the Patent Office of the Allgemeine Elektrizitatsgesell Schaft. Not satisfied with this employment, he decided to take up a teaching career and he taught for one year at a high school in Salem, near the Lake Costance. In 1928 he decided to return to university to study philosophy and pedagogy.

After his marriage with Johanna Ehmer in 1930, he moved to Crossen as a teacher in the local grammar school. However, when Hitler came to power in 1933, the citizens of Jewish origin were removed from public posts and Ising lost his job in March of that year. He remained unemployed for approximately one year, except for a short period spent in Paris as a teacher in a school for foreign children. In 1934 he found a new job as a teacher at the school opened from the Jewish community near Caputh, a city close to Potsdam, and in 1937 he became the dean of the same school. On 10 November 1938 he witnessed the devastation of the premises of the school by the boys and the inhabitants of Caputh, urged by local politicians to follow the example of the general pogrom in action against the Hebrew population throughout Germany.

In 1939 Ernst and Johanna Ising were caught in Luxemburg while they were trying to emigrate to the United States. Their visa applications were rejected due to the limits put on immigration flows. They decided though to remain there, waiting for the approval of their visa that was expected for the successive year. However, just on the day of his 40th birthday, the Germans invaded Luxemburg, and all consular offices were closed: this cut off any possibility of expatriation. Despite all the troubles, Ising and his family succeeded, however, surviving the horrors of the war, even though from 1943 until the liberation of 1944, Ernst Ising was forced to work for the German army on the railway lanes.

It was only two years after the end of the war that Ising and his wife left Europe on a cargo ship directed to United States. There he initially taught at the State Teacher's College of Minot and then at Bradley University, where he was Professor of Physics from 1948 till 1976. He became an American citizen in 1953 and in 1971 he was rewarded as best teacher of the year. Ernst Ising died on 11 May of 1998 in his house at Peoria, in the state of the Illinois.

The life and the career of Ernst Ising were seriously marked by the events of the Nazi dictatorship and of the Second World War: after his PhD thesis, he never came back to research activity. He lived quite isolated for many years, almost unaware of the new scientific developments. However, his article published in 1925 had a different fate. It was first quoted in an article by Heisenberg in 1928, devoted to the study of exchange forces between magnetic dipoles. However the true impulse to its reputation came from a famous article of Peierls, published in 1936, whose title read *On the Model of Ising for the Ferromagnetics*. Since then, the scientific literature has seen a large proliferation of articles on this model.

In closing these short biographical notes, it is worth adding that it was only in 1949 that Ising became aware of the great fame of his name and of his model within the scientific community.

Appendix 1A. Ensembles in Classical Statistical Mechanics

Statistical mechanics is the field of physics mainly interested in the thermodynamic properties of systems made of an enormous number of particles, typically of the order of the Avogadro number $N_A \sim 10^{23}$. To study such systems, it is crucial to make use of probabilistic methods for it is generally impossible to determine the trajectory of each particle and it is nevertheless meaningless to use them for deriving the thermodynamic properties. On the contrary, the approaches based on probability permit us to compute in a easier way the mean values of the physical quantities and their fluctuations.

The statistical mechanics of a system at equilibrium can be formulated in three different ways, which are based on the *microcanonical ensemble, canonical ensemble,* or *grand-canonical ensemble*. For macroscopic systems, the three different ensembles give the same final results. The choice of one or another of them is then just a question of what is the most convenient for the problem at hand. In this appendix we will recall the formulation of the three ensembles of classical statistical mechanics while in the next appendix we will discuss their quantum version.

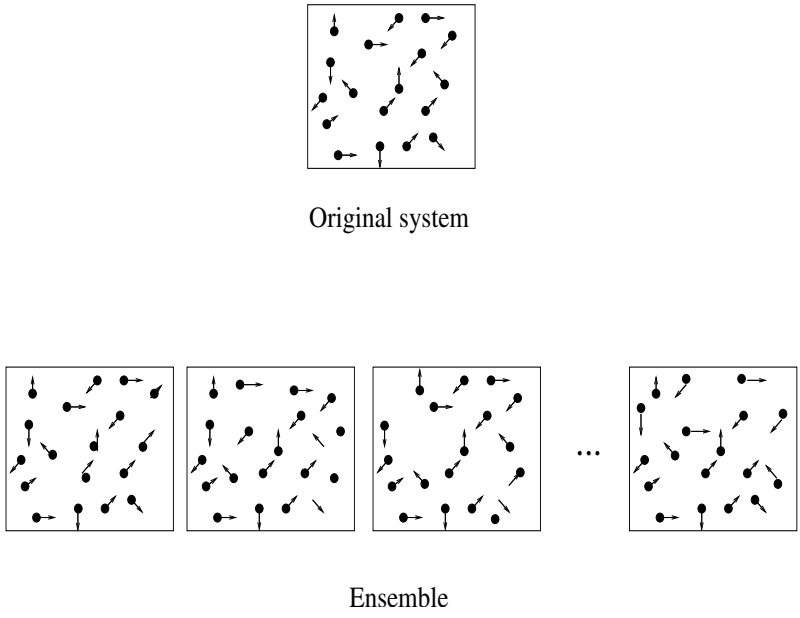


Fig. 1.7 *From the initial system to the ensemble.*

It is convenient to introduce the phase space Γ of the system. Let's assume that the system is made of N particles, each of them identified by a set of d coordinates q_i and d momenta p_i. The phase space Γ is the vector space of $2d \times N$ dimensions, given by the tensor product of the coordinates and momenta of all the particles. In the phase space, the system is identified at any given time by a point and its motion is associated to a curve in this space. If the system is isolated, its total energy E is conserved: in this case the motion takes place along a curve of the surface of Γ defined by the equation $H(q_i, p_i) = E$, where $H(q_i, p_i)$ is the hamiltonian of the system.

For a system with a large number of particles not only is it impossible to follow its motion but it is also useless. The only thing that matters is the possibility to predict the average properties of the system that are determined by the macroscopic constraints to which the system is subjected, such as its volume V, the total number N of particles, and its total energy E. Since there is generally a huge number of microscopic states compatible with a given set of macroscopic constraints, it is natural to assume that the system will visit all of them during its temporal evolution.[12] Instead of considering the time evolution of the system, it is more convenient to consider an infinite number of copies of the same system, with the same macroscopic constraints. This leads to the idea of *statistical ensembles* (see Fig. 1.7). By using an analogy, this is equivalent to looking at an infinite number of snapshots of a single movie rather than the movie itself. The ensembles then provide a statistical sampling of the system.

[12]The validity of these considerations is based on an additional assumption, namely the ergodicity of the system under consideration. By definition a system is ergodic if its motion passes arbitrarily close to all points of the surfaces of the phase space identified by the macroscopic conditions alone. The motion of systems that have additional conservation laws is usually not ergodic, since it takes place only on particular regions of these surfaces.

Since each system is represented by a single point in phase space, the set of systems associated to the ensemble corresponds to a swarm of points in phase space. Because the Liouville theorem states that the density of the points at any given point remains constant during the time evolution,[13] a probability density $\tilde{\rho}_i(q,p)$ is naturally defined in Γ. Hence, we can determine expectation values of physical quantities in terms of expectation values on the ensemble (a procedure that is relatively easy) rather than as a time average of an individual system (a procedure that is instead rather complicated). If the system is ergodic we have in fact the fundamental identity

$$\langle A\rangle \;=\; \lim_{t\to\infty}\frac{1}{t}\int_0^t d\tau\, A\,[q(\tau),p(\tau)] \;=\; \int dq\,dp\,A(p,q)\;\tilde{\rho}(q,p).$$

The different ensembles are defined by the different macroscopic conditions imposed on the system. Let's discuss the three cases that are used most often.

Microcanonical ensemble. The microcanonical ensemble is defined by the following macroscopic conditions: a fixed number N of particles, a given volume V, and a given value of the energy in the range E and $E+\Delta$. In this ensemble the mean values are computed in terms of the probability density $\rho(q,p)$ defined by

$$\rho(q,p) \;=\; \begin{cases} 1 & \text{if } E < H(p,q) < E+\Delta s, \\ 0 & \text{otherwise} \end{cases} \tag{1.A.1}$$

i.e. for any physical quantity A we have

$$\langle A\rangle \;=\; \frac{\int dq\,dp\,A(q,p)\,\rho(q,p)}{\int dq\,dp\,\rho(q,p)}.$$

The fundamental physical quantity in this formulation is the entropy. Once this quantities is known, one can recover all the rest of the thermodynamics. The entropy is a function of E and V, defined by

$$S(E,V) \;=\; k\,\log\Omega(E,V), \tag{1.A.2}$$

where k is the Boltzmann constant and Ω is the volume in the phase space Γ of the microcanonical ensemble

$$\Omega(E,V) \;=\; \int dq\,dp\,\rho(q,p).$$

The absolute temperature is then given by

$$\frac{1}{T} \;=\; \frac{\partial S(E,V)}{\partial E},$$

[13] According to a theorem by Liouville, $\frac{dD}{dt} = 0$, hence the density D satisfies the differential equation $\frac{\partial D}{\partial l} = -\{H,D\}$. At equilibrium, the density D does not vary with time and then satisfies $\{H,D\}=0$. This means that it is only a function of the integrals of motion of the system.

while the pressure P is defined by

$$P \,=\, T\,\frac{\partial S(E,V)}{\partial V}.$$

For the differential of S we have

$$dS(E,V) \,=\, \frac{\partial S}{\partial E}\,dE + \frac{\partial S}{\partial V}\,dV \,=\, \frac{1}{T}(dE + PdV),$$

i.e. the first law of the thermodynamics.

Canonical ensemble. The canonical ensemble permits us to deal with the statistical properties of a system that is in contact with a thermal bath much larger than the system itself. In this ensemble, the assigned macroscopic conditions are given by the total number N of the particles, the volume V of the system, and its temperature T. In this ensemble we cannot fix a priori the value of the energy, for it can be freely exchanged between the system and the thermal bath. These conditions are considered to be more closely related to the actual physical situations, since the temperature of a system can be easily tuned while it is more difficult to ensure the isolation of a system and the constant value of its energy. The probability density of the canonical ensemble takes the form of the Gibbs distribution

$$\rho(q,p) \,=\, e^{-\beta\,H(q,p)},$$

with $\beta = 1/kT$. The partition function is given by

$$Z_N(V,T) \,=\, \int dq\,dp\,e^{-\beta\,H(q,p)}.$$

The mean values are computed according to the formula

$$\langle A\rangle \,=\, \frac{1}{Z_N}\int dq\,dp\,A(q,p)e^{-\beta\,H(q,p)}.$$

As discussed in the text, the partition function Z_N permits us to recover the thermodynamics of the system. The equivalence between the microcanonical and the canonical ensembles can be proved by analyzing the fluctuations of the energy

$$\Delta E^2 \,=\, \langle H^2\rangle - \langle H\rangle^2.$$

A simple calculation gives

$$\langle H^2\rangle - \langle H\rangle^2 \,=\, kT^2\,\frac{\partial\langle H\rangle}{\partial T} \,=\, kT^2\,C_V,$$

where C_V is the specific heat. Since in a macroscopic system $\langle H\rangle \,\propto\, N$ but also $C_V \,\propto\, N$ (by the extensive nature of both quantities), the fluctuations of the energy are of gaussian type, namely in the limit $N \to \infty$ we have

$$\lim_{N\to\infty}\frac{\Delta E^2}{\langle H\rangle^2} = 0.$$

In other words, even though in the canonical ensemble the energy is a quantity that is not fixed but is subjected to fluctuations, as a matter of fact it assumes the same value

in the utmost majority of the systems of the ensemble. This proves the equivalence between the two ensembles.

Grand canonical ensemble. With the reasoning that we used to introduce the canonical ensemble, i.e. the possibility to control the temperature rather than its conjugate variable given by the energy, to introduce the grand canonical ensemble one argues that it is not realistic to assume that the total number N of the particles of a system is known a priori. In fact, experiments can usually determine only the mean value of this quantity. Hence, in the grand canonical ensemble one posits that the system can have an arbitrary number of particles, with its mean value determined by its macroscopic conditions. By introducing the quantity $z = e^{\beta\mu}$, where μ is the *fugacity*, the probability density of the grand canonical ensemble is given by

$$\rho(q,p,N) \;=\; \frac{1}{N!}\, z^N e^{-\beta\, H(q,p)}. \tag{1.A.3}$$

The term $N!$ in this formula takes into account the identity of the configurations obtained by the permutation of N identical particles. By integrating over the coordinates and the momenta present in (1.A.3), we arrive at the probability density relative to N particles. In its normalized form, it is expressed by

$$\rho(N) \;=\; \frac{1}{\mathcal{Z}}\, \frac{z^N}{N!}\, Z_N(V,T),$$

where $Z_N(V,T)$ is the partition function of the canonical ensemble with N particles, whereas the denominator of this formula defines the grand canonical partition function

$$\mathcal{Z}(z,V,T) \;=\; \sum_{N=0}^{\infty} \frac{z^N}{N!}\, Z_N(V,T).$$

The mean value of the number of particles of the system can be computed by the formula

$$\langle N\rangle \;=\; \sum_{N=0}^{\infty} N\,\rho(N) \;=\; z\,\frac{\partial}{\partial z}\,\log \mathcal{Z}(z,V,T). \tag{1.A.4}$$

The fundamental formula of the grand canonical ensemble links the pressure P to the partition function $\mathcal{Z}$

$$P \;=\; \frac{1}{\beta V}\,\log \mathcal{Z}(z,V,T). \tag{1.A.5}$$

The equation of state, i.e. the relationship among P, V, and $\langle N\rangle$, is obtained by expressing z by using eqn (1.A.4) and substituting it in (1.A.5).

The equivalence of this ensemble to the previous ones can be proved by showing that the fluctuations of the number of particles are purely gaussian. It is easy to prove that, in an infinite volume and away from the critical points of the system, one has in fact

$$\lim_{V\to\infty} \frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle^2} \;=\; 0.$$

This equation shows that, even though the number of particles of the system is not fixed a priori, it has the same value in almost all copies of the ensemble.

Appendix 1B. Ensembles in Quantum Statistical Mechanics

In this appendix we will recall the main formulas of statistical mechanics in the context of quantum theory. In quantum mechanics any observable A is associated with a hermitian operator that acts on a Hilbert space. At each time t, the state of an isolated system is identified by a vector $\mid \Psi(t)\rangle$ that evolves according to the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \mid \Psi(t)\rangle = H \mid \Psi(t)\rangle, \tag{1.B.1}$$

where H is the hamiltonian. By using the linear superposition principle, each state of the system can be expressed in terms of a complete set of states $\mid \psi_n\rangle$ provided by the orthonormal eigenvectors of any observable A

$$A \mid \psi_n\rangle = a_n \mid \psi_n\rangle, \quad \langle \psi_n \mid \psi_m\rangle = \delta_{n,m}.$$

This means that $\mid \Psi\rangle$ is given by

$$\mid \Psi\rangle = \sum_n c_n \mid \psi_n\rangle. \tag{1.B.2}$$

For the completeness relation of these states,

$$\sum_n \mid \psi_n\rangle \langle \psi_n \mid = 1.$$

The coefficients c_n of the expansion (1.B.2) are expressed by the scalar product $c_n = \langle \Psi \mid \psi_n\rangle$, and the square of their modulus $\mid c_n \mid^2$ expresses the probability to obtain the eigenvalues a_n as a result of the measurement of the observable A on the state $\mid \Psi\rangle$. Hence

$$\langle \Psi \mid \Psi\rangle = \sum_n \mid c_n \mid^2 = 1.$$

Let's now discuss the statistical properties of quantum systems. As in the classical case, in the presence of a large number of particles it is highly unrealistic to determine the behavior of a system by solving the Schrödinger equation: first of all, this is an impossible goal to pursue in almost all systems and, secondly, it cannot be used to predict the thermodynamic properties. Hence, also in the quantum case, one needs to use a statistical formulation: one has to take into account the incomplete information on the state of the system and extract the predictions only on the mean values of the observables. To do so, let us imagine that the system under study can be considered as a subsystem of a larger one (external world) and in thermodynamic equilibrium. Denote by $\mathcal{H}$ the hamiltonian of such subsystem, E_n the spectrum of its eigenvalues, and $\mid \varphi_n\rangle$ its eigenvectors (without the temporal term). We can use $\mid \varphi_n\rangle$ to express the states of the system, as in eqn (1.B.2), but in this case the coefficients $c_n(t)$ have the meaning of wavefunctions of the external world.

Suppose we consider at a given time instant the quantum mean value of an observable $\mathcal{O}$ on the state $\mid \Psi\rangle$. According to the rules of quantum mechanics, this is given by the expectation value

$$\langle \Psi(t) \mid \mathcal{O} \mid \Psi(t) \rangle \;=\; \sum_{n,m} c_n^*(t)\, c_m(t)\, \langle \varphi_n \mid \mathcal{O} \mid \varphi_m \rangle \;=\; \sum_{n,m} c_n^*(t)\, c_m(t)\, \mathcal{O}_{n,m}, \qquad (1.B.3)$$

where $\mathcal{O}_{n,m} = \langle \varphi_n \mid \mathcal{O} \mid \varphi_m \rangle$. Since we have only partial information on the system, we have to take a statistical average. Under the hypothesis of ergodicity,[14] this is equivalent to taking the *time average* of (1.B.3). Defining

$$\rho_{m,n} \;=\; \overline{c_m(t)\, c_n^*(t)} \;\equiv\; \lim_{t\to\infty} \frac{1}{t} \int_0^t c_m(\tau)\, c_n^*(\tau)\; d\tau, \qquad (1.B.4)$$

the *statistical average* of the observable $\mathcal{O}$ can be expressed by the formula

$$\overline{\langle \mathcal{O} \rangle} \;=\; \overline{\langle \Psi \mid \mathcal{O} \mid \Psi \rangle} \;=\; \sum_{n,m} \rho_{m,n} \mathcal{O}_{n,m} \;=\; \mathrm{Tr}(\rho\, \mathcal{O}), \qquad (1.B.5)$$

where the operator ρ, defined by its matrix elements (1.B.4), is the *density matrix*. Since the trace of an operator is independent of the basis, the final result (1.B.5) does not depend on the basis of the eigenvectors that we used to expand the state $\mid \Psi\rangle$. It should be stressed that the average (1.B.5) that involves the density matrix has two aspects: from one side, it includes the quantum average on the state, but, on the other hand, it performs the statistical average on the wavefunctions of the environment. Both averages are simultaneously present in the formula (1.B.5).

In quantum statistical mechanics, the density matrix corresponds to the probability distribution of classical statistical mechanics. Hence, also in this case, we can introduce three different ensembles.

Microcanonical ensemble. As in the classical case, the microcanonical ensemble is defined by the following macroscopic conditions: a fixed number N of particles, a fixed volume V, and the energy of the system in the range E and $E + \Delta$. Correspondingly, the density matrix assumes the form

$$\rho_{n,m} \;=\; \delta_{n,m}\, w_n, \qquad w_n \;=\; \begin{cases} 1; & E < E_n < E + \Delta \\ 0; & \text{otherwise} \end{cases}$$

and the thermodynamics is derived starting from the entropy

$$S(E, V) \;=\; k \,\log\, \Omega(E, V),$$

where

$$\Omega(E, V) \;=\; \mathrm{Tr}\, \rho.$$

[14] In quantum mechanics this implies the absence of non-trivial integrals of motion, i.e. a set of observables that commute with the hamiltonian and that can be simultaneously diagonalized with it.

Canonical ensemble. In this ensemble the macroscopic variables are given by the fixed number N of particles, the volume V, and the temperature T. The corresponding expression the density matrix is given by

$$\rho_{n,m} \;=\; \delta_{n,m}\, e^{-\beta E_n},$$

with the partition function expressed by

$$Z_N(V,T) \;=\; \mathrm{Tr}\,\rho \;=\; \sum_n e^{-\beta E_n}.$$

In this ensemble, the thermodynamics is derived starting from the free energy

$$F_N(V,T) \;=\; -\beta^{-1}\,\log\, Z_N(V,T).$$

Grand canonical ensemble. In the grand canonical ensemble the macroscopic variables are the volume V and the temperature T. In this case the density matrix acts on a Hilbert space with an indefinite number of particles. Denoting by $E_{n,N}$ the n-th energy level with N particles, the density matrix is expressed by

$$\rho_{n,N} \;=\; z^N\, e^{-\beta\, E_{n,N}},$$

where $z = e^{\beta\mu}$. The equation of state is similar to the classical one

$$P \;=\; \frac{1}{\beta V}\,\log \mathcal{Z}(z,V,T),$$

where $\mathcal{Z}(z,V,T)$ is the grand canonical partition function

$$\mathcal{Z}(z,V,T) \;=\; \sum_{N,n} z^N\, e^{-\beta\, E_{n,N}}.$$

Indistinguishable particles and statistics. A central idea of quantum theory is the concept of indistinguishable particles: for a system with many identical particles, an operation that exchanges two of them, swapping their positions, leaves the physics invariant. This symmetry is represented by a unitary transformation acting on the many-body wavefunction. In three spatial dimension, there are only two possible symmetry operations: the wavefunction of bosons is symmetric under exchange while that of fermions is antisymmetric. The limitation to one of the two possible kinds of quantum symmetry comes from a simple topological argument: a process in which two particles are adiabatically interchanged twice is equivalent to a process in which one of the particles is adiabatically taken around the other. Wrapping one particle around another is then topologically equivalent to having a loop. In three dimensions, such a loop can be safely shrink to zero and, therefore, the wavefunction should be left unchanged by two such interchanges of particles. The only two possibilities are that the wavefunction changes by a $\pm$ sign under a single interchange, corresponding to the cases of bosons and fermions, respectively.

For the same topological reason, the concept of identical-particle statistics becomes ambiguous in one spatial dimension. In this case, for swapping the positions of two particles, they need to pass through one another and it becomes impossible to disentangle the statistical properties from the interactions. If the wavefunction changes sign when two identical particles swap their positions, one could say that the particles are non-interacting fermions or, equivalently, that the particles are interacting bosons, where the change of sign is induced by the interaction as the particles pass through one another. This the main reason at the root of the possibility to adopt the *bosonization procedure* for describing one-dimensional fermions in terms of bosons and vice versa, as we will see in Chapter 12.

In two dimensions, a remarkably rich variety of particle statistics is possible: here there are indistinguishable particles that are neither bosons nor fermions, and they are called *anyons*. In *abelian anyons*, the two-particle wavefunction can change by an arbitrary phase when one particle is exchanged with the other

$$\psi(r_1, r_2) \to e^{i\theta}\, \psi(r_1, r_2). \tag{1.B.6}$$

There could also be *non-abelian anyons*. In this case there is a degenerate set of g states $\psi_a(r_1, \ldots, r_n)$ $(a = 1, 2, \ldots, g)$, with anyons at the positions $r_1, r_2, \ldots, r_n$. The interchanges of two particles are elements of a group, called *the braid group* (see Problem 15). If β_i is the operation that interchanges particles i and $i + 1$, it can be represented by a $g \times g$ unitary matrix $\gamma(\beta_i)$ that acts on these states as

$$\psi_a \to [\gamma(\beta_i)]_{ab}\, \psi_b.$$

The set of the $(n - 1)$ matrices $\gamma(\beta_i)$ $(i = 1, 2, \ldots, n - 1)$ satisfy the Artin relations, discussed in Problem 15.

The situation of non-abelian anyons is realized, for instance, by trapping electrons in a thin layer between two semiconductor slabs. At a sufficiently strong magnetic field in the orthogonal direction and at a sufficiently low temperature, the wavefunction of the two-dimensional electron gas describes a deeply entangled ground state. The excitations above the ground state carry electron charges that are fractions of the original electron charge and have unusual statistical properties under the interchange of two of them. The anyons of this system give rise to the spectacular transport effects of the fractional quantum Hall effect.

Free particles. An important example of quantum statistical mechanics is provided by a system of free particles. This system can be described by the states of a single particle, here denoted by the index ν. Since the particles are indistinguishable at the quantum level, to specify a state of the system it is sufficient to state the occupation number n_ν of each of its modes. If ϵ_ν is the energy of the ν-th mode, the total energy of the system is given by

$$E = \sum_\nu n_\nu\, \epsilon_\nu,$$

while the total number of particles is

$$N = \sum_\nu n_\nu.$$

For three-dimensional systems, there are only two cases: the first is relative to Fermi–Dirac (FD) statistics, the second to Bose–Einstein (BE) statistics. In the first case, each mode can be occupied by at most one particle, so that the possible values of n_ν are

$$n_\nu = 0, 1 \qquad \text{Fermi–Dirac}$$

while, in the second case, each mode can be occupied by an arbitrary number of particles. In this case the possible values of n_ν coincide with the natural numbers

$$n_\nu = 0, 1, 2, \ldots \qquad \text{Bose–Einstein.}$$

The most convenient ensemble to describe the thermodynamics of this system is the grand canonical one. The corresponding partition function is

$$\mathcal{Z}(z, V, T) = \sum_{N=0}^{\infty} \sum_{\substack{\{n_\nu\} \\ \sum n_\nu = N}} z^N \, e^{-\beta \sum n_\nu \epsilon_\nu} = \sum_{N=0} \sum_{\substack{\{n_\nu\} \\ \sum n_\nu = N}} \prod_\nu \left(z \, e^{-\beta \epsilon_\nu} \right)^{n_\nu} .$$

To perform the double sums, it is sufficient to sum *independently* on each index n_ν, for every term in one case appears once and only once in the other, and vice versa. Hence

$$\begin{aligned} \mathcal{Z}(z, V, T) &= \sum_{n_0} \sum_{n_1} \cdots \left[\left(z e^{-\beta \epsilon_0} \right)^{n_0} \left(z e^{-\beta \epsilon_1} \right)^{n_1} \cdots \right] \\ &= \left[\sum_{n_0} \left(z e^{-\beta \epsilon_0} \right)^{n_0} \right] \left[\sum_{n_1} \left(z e^{-\beta \epsilon_1} \right)^{n_1} \right] \cdots \\ &= \prod_\nu \left[\sum_n \left(z e^{-\beta \epsilon_\nu} \right)^n \right], \end{aligned} \tag{1.B.7}$$

where the final sum is on the values $0, 1$ for the fermionic case and on all the integers for the bosonic case. In the first case we have

$$\mathcal{Z}_F(z, V, T) = \prod_\nu \left[1 + z e^{-\beta \epsilon_\nu} \right],$$

while, in the second case, one has a geometrical series

$$\mathcal{Z}_B(z, V, T) = \prod_\nu \left[\frac{1}{1 - z e^{-\beta \epsilon_\nu}} \right].$$

The two expressions can be unified by the formula

$$\mathcal{Z}(z, V, T) = \prod_\nu \left(1 \pm z \, e^{-\beta \, \epsilon_\nu} \right)^{\pm 1},$$

where the + sign referes to Fermi–Dirac statistics whereas the − sign refers to Bose–Einsten. The equation of state of both cases is

$$\beta\, P\, V \;=\; \log\, \mathcal{Z}(z,V,T) \;=\; \pm \sum_{\nu} \log\left(1 \pm z\, e^{-\beta\, \epsilon_\nu}\right),$$

where the variable z is related to the average number of particles by the equation

$$N \;=\; z\, \frac{\partial}{\partial z}\, \log\, \mathcal{Z}(z,V,T) \;=\; \sum_{\nu} \frac{z\, e^{-\beta \epsilon_\nu}}{1 \pm z\, e^{-\beta\, \epsilon_\nu}}. \tag{1.B.8}$$

The last expression shows that the occupation average of each mode is given in both cases by

$$\langle n_\nu \rangle \;=\; \frac{z\, e^{-\beta \epsilon_\nu}}{1 \pm z\, e^{-\beta \epsilon_\nu}}. \tag{1.B.9}$$

Let's briefly discuss the main features of the Fermi–Dirac and Bose–Einstein distributions.

Fermi–Dirac. As is well known, the Fermi–Dirac distribution of free particles turns out to be a surprisingly good model for the behavior of conduction electrons in a metal or for understanding, in the relativistic case, the existence of an upper limit of the mass of the dwarf stars (Chandrasekhar limit).

In order to discuss the fermion system in more detail, let's put $z = e^{\beta\mu}$ and let's consider the occupation average $n(\epsilon)$ in the limit $T \to 0$

$$n(\epsilon) \;=\; \frac{1}{e^{(\epsilon-\mu)/kT} + 1} \;\longrightarrow\; \begin{cases} 1, \text{ if } \epsilon < \mu \\ 0, \text{ if } \epsilon > \mu. \end{cases} \tag{1.B.10}$$

Note that in general the chemical potential depends on temperature. Its zero temperature value is the called the *Fermi energy*, $\epsilon_F = \mu(T = 0)$. The physical origin of the sharp shape of the limit expression (1.B.10) is the Pauli exclusion principle that posits that no two particles can be in the same level of the system. At zero temperature, the particles occupy the lowest possible energy levels up to a finite energy level ϵ_F. In momentum space, the particles fill a sphere of radius p_F, called the *Fermi sphere*. In this regime the gas is said to be *degenerate*. To compute ϵ_F, let's consider the gas inside a cube of side L with periodic boundary conditions, for simplicity. The energy of a single particle is just the kinetic energy $E = \frac{\mathbf{p}^2}{2m}$ and the components p_i of the momentum are quantized as

$$p_i \;=\; \frac{2\pi\hbar}{L}\, q_i, \qquad q_i \;=\; 0, \pm 1, \pm 2, \ldots$$

For large L it is natural to replace the sum (1.B.8) with an integral, according to the rule

$$\sum_{q} \;\longrightarrow\; \frac{V}{(2\pi\hbar)^3} \int d\vec{p}, \tag{1.B.11}$$

where $V = L^3$. If the spin of a particle is s, for a given momentum $\vec{p}$ there are $2s + 1$ single particle states with the same energy $\epsilon(p)$ and the normalization condition at

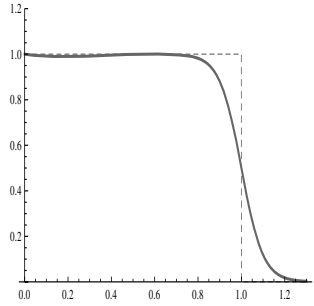


Fig. 1.8 *Fermi-Dirac distribution at $T = 0$ (dashed line) and at $T \neq 0$ (continuous line).*

$T = 0$ becomes

$$N \;=\; (2s+1)\,\frac{V}{(2\pi\hbar)^3}\int_{\epsilon<\epsilon_F} d^3p \;=\; (2s+1)\frac{V}{(2\pi\hbar)^3}\,\frac{4\pi}{3}\,p_F^3. \tag{1.B.12}$$

Hence,

$$\epsilon_F \;=\; \frac{\hbar^2}{2m}\left(\frac{6\pi^2}{2s+1}\,\frac{N}{V}\right)^{2/3}. \tag{1.B.13}$$

We can define a *Fermi temperature* T_F by $\epsilon_F \equiv kT_F$. The Fermi energy and temperature provide useful energy and temperature scales for understanding the properties of fermion systems. For instance, the conduction electron density for metals is typically of order 10^{22} per cubic centimeter, which corresponds to a Fermi temperature of order 10^5 kelvin. This implies that at room temperature the system can be reasonably approximated by the degenerate distribution (1.B.10). Furthermore, notice that the Fermi energy (1.B.13) increases by increasing the density of the gas and, at sufficiently high density, ϵ_F can be higher than any energy scale ϵ_I associated to the interactions between the particles. This means that, counter-intuitively, in fermion systems the free particle approximation becomes better at higher values of the density!

At finite temperatures but smaller than the Fermi temperature $T < T_F$, $n(\epsilon)$ differs from its zero-temperature form only in a small region about μ of width a few kT, as shown in Fig. 1.8. In computing integrals of the form $J = \int_0^\infty f(\epsilon)n(\epsilon)d\epsilon$, the way they differ from the zero temperature values $\int_0^{\mu=\epsilon_F} f(\epsilon)n(\epsilon)d\epsilon$ depends on the form of $f(\epsilon)$ near μ. Integrating by parts, such integrals can be expressed as

$$J \;=\; -\int_0^\infty g(\epsilon)\,n'(\epsilon)\,d\epsilon, \tag{1.B.14}$$

where $g(\epsilon) = f'(\epsilon)$. Note that $n'(\epsilon)$ is sharply peaked at $\epsilon = \mu$, particularly at low temperature. If $g(\epsilon)$ does not vary rapidly in an interval of order kT near μ, the value of the integral can thus be estimated by replacing $g(\epsilon)$ with the first few term of its Taylor expansion about $\epsilon = \mu$

$$g(\epsilon) \;=\; \sum_{n=0}^{\infty}\frac{1}{n!}\frac{d^n g(\mu)}{d\epsilon^n}\,(\epsilon-\mu)^n.$$

Substituting this expansion the integral (1.B.14) becomes

$$J \;=\; -\sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n g(\mu)}{d\epsilon^n} \int_0^{\infty} n'(\epsilon)\,(\epsilon-\mu)^n\, d\epsilon.$$

The various integrals can be evaluated with the substitution $x = (\epsilon - \mu)/kT$ and since n' vanishes away from $\epsilon = \mu$, the lower limit of the integrals can be enlarged to $-\infty$ without significant error. So

$$\int_0^{\infty} n'(\epsilon)(\epsilon-\mu)^n\, d\epsilon \;=\; -(kT)^n\, I_n$$

where

$$I_n \;=\; \int_{-\infty}^{\infty} \frac{x^n\, e^x}{(e^x+1)^2}\, dx.$$

Since $e^x/(e^x+1) = 2/\cosh(x/2)$ is an even function, for n odd I_n vanish. The even ones can be expressed in terms of the Riemann function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ as

$$I_{2n} \;=\; (2n)!(2-2^{2-2n})\zeta(2n).$$

The first representatives are

$$I_0 \;=\; 1, \qquad I_2 \;=\; \frac{\pi^2}{3}, \qquad I_4 \;=\; \frac{7\pi^4}{15}.$$

In this way we recover the so-called *Sommerfeld expansion* of the integral J (where we have inserted the original function $f(\epsilon)$)

$$\begin{aligned} &\int_0^{\infty} f(\epsilon) n(\epsilon) d\epsilon \\ &= \int_0^{\mu} f(\epsilon) d\epsilon + \frac{\pi^2}{6}(kT)^2 f'(\mu) + \frac{7\pi^4}{360}(kT)^4 f'''(\mu) + \mathcal{O}\left(\frac{kT}{\mu}\right). \end{aligned}$$

Applying the formula above, it is possible to compute the dependence of the chemical potential on the temperature

$$\mu \;=\; \epsilon_F \left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\epsilon_F}\right)^2 - \frac{\pi^4}{80}\left(\frac{kT}{\epsilon_F}\right)^4 + \cdots\right],$$

and the expression for the internal energy

$$U \;=\; \frac{3}{5} N\epsilon_F \left[1 + \frac{5}{3}\left(\frac{kT}{\epsilon_F}\right)^2 + \cdots\right],$$

For the pressure we have

$$P \;=\; \frac{2}{5}\frac{N}{V}\epsilon_F \left[1 + \frac{5\pi^2}{12}\left(\frac{kT}{\epsilon_F}\right)^2 + \cdots\right].$$

This formula shows that even at zero temperature there is a non-zero value of the pressure, another manifestation of the Pauli principle.

Bose–Einstein. In three dimensions the boson gas presents the interesting phenomenon of Bose–Einstein condensation, i.e. a first-order phase transition. This phenomenon was predicted by Einstein in 1924. The condensation was achieved for the first time in atomic gases in 1995: the group of E. Cornell and C. Wieman was first, with ^{87}Rb atoms, followed by the group of W. Ketterle with ^{23}Na atoms and the group of R. Hulet with ^{7}Li atoms. In these experiments the atomic gas was confined by a magnetic and/or optical trap to a relatively small region of space and at a temperature of order nanokelvins. In order to discuss this remarkable aspect of bosons in more detail, let's consider, as before, the gas inside a cube of side L with periodic boundary conditions. The components p_i of the momentum are quantized as

$$p_i \,=\, \frac{2\pi\hbar}{L}\, q_i, \qquad q_i \,=\, 0, \pm 1, \pm 2, \ldots$$

and the energy of a single particle is $E = \frac{\mathbf{p}^2}{2m}$. Since the mean value (1.B.9) of the number of particles for each mode ν has to be positive (in particular, the mode relative to the zero energy), for the variable z we have

$$0 \le z \le 1.$$

To compute the mean value of the density of the particle in the limit $L \to \infty$, it seems natural to replace the sum (1.B.8) with an integral, according to the rule (1.B.11). In this way, we have

$$\frac{N}{V} \,=\, \int \frac{d\vec{p}}{\hbar^3}\, \frac{1}{z^{-1}\, e^{\beta p^2/2m} - 1}, \tag{1.B.15}$$

which, by a change of variable, can be written as

$$N \,=\, \frac{V}{\lambda^3}\, g(z), \tag{1.B.16}$$

where

$$g(z) \,=\, \frac{4}{\sqrt{\pi}} \int_0^\infty dx\, \frac{x^2\, e^{-x^2}}{z^{-1} - e^{-x^2}} \,=\, \sum_{n=1}^{\infty} \frac{z^n}{n^{3/2}}.$$

The quantity

$$\lambda \,=\, \sqrt{\frac{2\pi\hbar^2}{mkT}}$$

has the dimension of a length and it is called the *thermal wavelength*, for it expresses the order of magnitude of the de Broglie wavelength associated to a particle of mass m and energy kT. λ can be regarded as the position uncertainty associated with the thermal momentum distribution. The lower the temperature, the longer λ. When atoms are cooled to the point where λ is comparable to the interatomic separation, the atomic wavepackets overlap and the indistinguishability of particles becomes an important physical effect. The function $g(z)$ is an increasing function of z, as shown in Fig. 1.9. At $z = 1$ the function reaches its highest value, expressed in terms of the

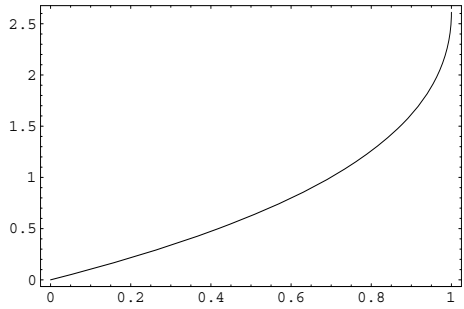


Fig. 1.9 *Plot of the function $g(z)$.*

Riemann function $\zeta(x)$ by

$$g(1) \,=\, \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} \,=\, \xi\left(\frac{3}{2}\right) \simeq 2.612...$$

and, for all values of z between 0 and 1, the function $g(z)$ satisfies the inequality

$$g(z) \,\leq\, g(1) = 2.612...$$

From eqn (1.B.16) the conclusion seems then to be that there exists a critical density of the system given by

$$N_{max} \,=\, g(1)\,\frac{V}{\lambda^3}.$$

But this is impossible due to the bosonic nature of the gas. In fact, if we had reached this critical density, what prevents us adding further particles to the system? Hence, there should be a mistake in the previous derivation, particularly in the substitution of the sum (1.B.8) with the integral (1.B.15). The cure of this drawback is to isolate the zero-mode before making the substitution of the sum with the integral. This is given by

$$n_0 \,=\, \frac{z}{z-1},$$

and for $z \to 1$, it is evident that it can be arbitrarily large, i.e. comparable with the sum of the entire series. Instead of (1.B.16), the correct version of the formula is then

$$N \,=\, \frac{V}{\lambda^3}\,g(z) + \frac{z}{z-1}.$$

Expressing it as

$$\lambda^3\,\frac{n_0}{V} \,=\, \lambda^3\,\frac{N}{V} - g(z),$$

it is easy to see that $n_0/V > 0$ when the temperature and the density of the particles satisfy the condition

$$\lambda^3 \frac{N}{V} \,\geq\, g(1) = 2.612\ldots \tag{1.B.17}$$

In this case, a finite fraction of the total number of the particles occupies the lowest energy level and a condensation phenomenon takes place. The system undergoes a

phase transition from a normal gas state to a Bose–Einstein condensation, in which there is a macroscopic manifestation of the quantum nature of the system. The phase transition (which is of first order) is realized when we have

$$\lambda^3 \frac{N}{V} = g(1).$$

This equation defines a curve in the space of the variables P-n-T. In particular, keeping fixed the density $d = N/V$, this equation identifies a critical temperature T_c given by

$$kT_c = \frac{2\pi\hbar^2}{m[d\, g(1)]^{2/3}},$$

Notice that T_c decreases when the mass of the particles increases. As previously mentioned, the Bose–Einstein condensation was realized for the first time in 1995 by using alkaline gases and, since then, it has become a research field under rapid development.

References and Further Reading

Statistical mechanics enjoys a surfeit of excellent texts. We especially recommend the following books as an introduction to many basic ideas and applications:

L.D. Landau, E.M. Lifshitz, L.P. Pitaevskij, *Statistical Physics*, Pergamon Press, Oxford, 1978.

K. Huang, *Statistical Mechanics*, John Wiley, New York, 1963.

R.P. Feynman, *Statistical Mechanics*, W.A. Benjamin, New York, 1972.

L.E. Reichl, *Modern Course in Statistical Physics*, Arnold Publishers, London, 1980.

R. Kubo, *Statistical Mechanics*, North Holland, Amsterdam 1965.

D.C. Mattis, *The Theory of Magnetism Made Simple*, World Scientific, Singapore, 2006.

Phase transitions are discussed in several monographs. A superb introduction to the modern theory of these phenomena is:

A.Z. Patasinskij, V.L. Pokrovskij, *Fluctuation Theory of Phase Transitions*, Pergamon Press, Oxford, 1979.

The classical volume by Baxter is an excellent in depth introduction to a large class of exactly solvable models of statistical mechanics and to the methods of solution:

R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, New York, 1982.

The book by H.E. Stanley is a standard reference for the phenomenology of critical phenomena and a general overview of the subject:

H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Oxford Science Publications, Oxford, 1971.

Some relevant papers and books about the Ising model are:

E. Ising, *Beitrag zur Theorie des Ferromagnetismus*, Zeit. für Physik 31 (1925), 253.

R. Peierls, *On Ising's model of ferromagnetism*, Proc. Camb. Phil. Soc. 32 (1936), 477.

L. Onsager, *Crystal statistics. I A two-dimensional model with an order-disorder transition*, Phys. Rev. 65 (1944), 117.

B. Kaufman, *Crystal statistics. II partition function evaluated by spinor analysis*, Phys. Rev. 76 (1949), 1232.

B.M. McCoy and T.T. Wu, *The Two Dimensional Ising Model*, Harvard University Press, Cambridge MA, 1973.

T.T. Wu, B.M. McCoy, C. Tracy and E. Barouch, *Spin-spin correlation functions for the two-dimensional Ising model: Exact theory in the scaling region*, Phys. Rev. B 13 (1976), 316.

M. Sato, T. Miwa and M. Jimbo, *Aspects of Holonomic Quantum Fields*, Lecture Notes in Physics Vol. 126, Springer Berlin, 1980.

A.B. Zamolodchikov, *Integrals of motion of the (scaled) $T = T_c$ Ising model with magnetic field*, Int. J. Mod. Phys. A 4 (1989), 4235.

G. Delfino and G. Mussardo, *The spin-spin correlation function in the two-dimensional Ising model in a magnetic field at $T = T_c$*, Nucl. Phys. B 455 (1995), 724.

G. Delfino and P. Simonetti, *Correlation functions in the two-dimensional Ising model in a magnetic field at $T = T_c$*, Phys. Lett. B 383 (1996), 450.

G. Delfino, G. Mussardo and P. Simonetti, *Non-integrable quantum field theories as perturbations of certain integrable models*, Nucl. Phys. B 473 (1996), 469.

P. Fonseca and A. Zamolodchikov, *Ising field theory in a magnetic field: Analytic properties of the free energy*, J. Stat. Phys. 110 (2003), 527.

The computational tractability frontier for the partition functions of several Ising models and their relationship with NP-complete problem is discussed in:

S. Istrail, *Statistical mechanics, three-dimensionality and NP-completeness: I. Universality of intractability of the partition functions of the Ising model across non-planar lattices* in Proceedings of the 32nd ACM Symposium on the Theory of Computing (STOC00) (Portland, Oregon, May 2000), ACM Press, pp. 87–96.

The history of the Ising model is discussed in:

M. Niss, *History of the Lenz–Ising model 1920–1950: From ferromagnetic to cooperative phenomena*, Arch. Hist. Exact Sci. 59 (2005), 267.

S. Brush, *Histroy of the Lenz–Ising model*, Rev. Mod. Phys. 39 (1967), 883.

Our presentation of the statistical properties of quantum particles is very schematic. A more complete treatment, in particular of the two-dimensional case, can be found in the book:

F. Wilczek, *Fractional Statistics and Anyon Superconductivity*, World Scientific, Singapore, 1990, and references therein.

Bose–Einstein condensation is a rapidly developing field. For an overall view on this subject, see:

L. Pitaevskii and S. Stringari, *Bose–Einstein Condensation*, Oxford University Press, Oxford, (2003).

C. Pethick and H. Smith, *Bose–Einstein Condensation in Dilute Gases*, Cambridge University Press, Cambridge, 2008.

Problems

1. Lattice gas

Consider a lattice gas in which the particles occupy the sites of a d-dimensional lattice, with the constraint that each site cannot be occupied by more than one particle. Let e_i be a variable that takes values $\{0, 1\}$: 0 when the site is vacant and 1 when it is occupied. The interaction energy of each configuration is given by

$$\mathcal{H} = J \sum_{\langle ij \rangle} e_i \, e_j .$$

Show that the grand canonical partition function of the lattice gas can be put in correspondence with the canonical partition function of the Ising model. Argue that the phase transition of the lattice gas, which consists of the condensation of the particles, belongs to the same universality class of the Ising model.

2. Potts model

In the Potts model, the spin variable σ_i assumes q values, as $\{0, 1, \ldots, q-1\}$. The energy of the configurations is given by

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \delta_{\sigma_i, \sigma_j} ,$$

where

$$\delta_{a,b} = \begin{cases} 1 \text{ if } a = b \\ 0 \text{ if } a \neq b. \end{cases}$$

a Identify the symmetry transformations of the spins that leave the hamiltonian invariant.

b Show that for $q = 2$, the Potts model is equivalent to the Ising model.

c Discuss the configuration of the minimum energy in the antiferromagnetic limit $J \to -\infty$.

3. Theorem of equipartition

Consider a classical one-dimensional harmonic oscillator, with hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2},$$

a Determine the surface $E = constant$ in the phase space and derive the thermodynamics of the system by using the microcanonical ensemble.

b Put the system in contact with a thermal bath at temperature T. Compute the partition function in the canonical ensemble and show that the mean value of the energy is independent both of the frequency and the mass of the particle, i.e.

$$\left\langle \frac{p^2}{2m} \right\rangle = \left\langle \frac{m\omega^2 x^2}{2} \right\rangle = \frac{1}{2}\langle H \rangle = \frac{1}{2} kT.$$

c Show that

$$\langle (E - \langle E \rangle)^2 \rangle = (kT)^2.$$

4. Equation of state for homogeneous potentials

Consider a system of classical particles whose interaction potential is given by a homogeneous function of degree η

$$U(\lambda \vec{r}_1, \lambda \vec{r}_2, \ldots, \lambda \vec{r}_N) = \lambda^\eta U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N).$$

Show that the equation of state of such a system assumes the form

$$P\, T^{-1+3/\eta} = f\left(\frac{V}{N} T^{-3/\eta}\right),$$

where, in principle, the function $f(x)$ can be computed once the explicit expression U of the potential is known.

5. Zeros of the partition function

Consider a classical system with only two states of magnetization, both proportional to the volume V of the system: $M = \pm \alpha V$. In the presence of an external magnetic field B, the Hamiltonian is given by

$$\mathcal{H} = B\, M.$$

a Compute the partition function in the canonical ensemble and determine its zeros in the complex plane of the temperature. Show that in the thermodynamic limit $V \to \infty$, there is an accumulation of zeros at $T = \infty$.

b Compute $\langle M \rangle$ as a function of B and study the limit of this function when $V \to \infty$.

6. Two-state systems

Consider a system of N free classical particles. The energy of each particle can take only two values: 0 and $E(E > 0)$. Let n_0 and n_1 be the occupation numbers of the two energy levels and U the total energy of the system.

a Determine the entropy of the system.

b Determine $\langle n_0 \rangle$, $\langle n_1 \rangle$ and their fluctuations.

c Express the temperature T as a function of U and show that it can take negative values.

d Discuss what happens when a system at negative temperature is put in thermal contact with a system at positive temperature.

7. Scaling laws

Given the equation of state of a magnetic system in the form

$$B \;=\; M^{\delta}\, \mathcal{Q}\left(\frac{t}{M^{1/\beta}}\right),$$

a prove that the parameters β and δ in the expression above are the critical exponents of the system, as defined in this chapter;

b Show the identity $\gamma = \beta(\delta - 1)$.

8. First-order phase transitions

In second-order phase transitions, the state with the lowest value of the free energy changes continuously when the system crosses its critical point. On the contrary, in a first-order phase transition, the order parameter changes discontinuously.

a Study the behaviour of the minima of the free energy

$$F(x) \;=\; a(T)\, x^2 + x^4$$

by varying the temperature T as $a(T) = (T - T_c)$ and determine if we are in the presence of a first- or second-order phase transition.

b Analyze the same questions for the free energy given by

$$F(x) \;=\; (x^2 - 1)^2\, (x^2 + a(T))$$

with the same expression for $a(T)$.

9. Ergodic system

Consider a classical dynamical system with a phase space ($0 < q < 1$; $0 < p < 1$) and equation of motion given by

$$q(t) = q_0 + t; \quad p(t) = p_0 + \alpha\, t.$$

a Discuss the trajectories in the phase space when α is a rational and irrational number.

b Show that the system is ergodic when α is irrational, i.e. the time averages of all functions $f(q, p)$ coincide with their average on the phase space.

Hint. Use the fact that the volume of the phase space is finite to expand any function of the coordinate and momentum in Fourier series.

10. Density of states

Determine the number of quantum states with energy less than E for a free particle in a cubic box of length L. Compare this quantity with the volume of the classical phase space and find the corresponding density of states of the system.

11. Quantum harmonic oscillator

The one-dimensional quantum oscillator has an energy spectrum given by $E_n = \hbar\omega(n + 1/2)$, $n = 0, 1, 2, \ldots$

a Compute the partition function in the canonical ensemble.

b Compute the specific heat as a function of the temperature and discuss how this quantity differs from the analogous classical expression.

12. Riemann function

The Riemann function $\zeta(\beta)$ is defined by

$$\zeta(\beta) = \sum_{n=1}^{\infty} \frac{1}{n^{\beta}}.$$

a Interpret this expression as the partition function in the canonical ensemble of a quantum system and identity the discrete spectrum of the energies.

b Compute the density of states and the entropy of the quantum system. Interpret the singularity of $\zeta(\beta)$ at $\beta = 1$ as a phase transition.

13. Bose–Einstein condensation

In Appendix B we saw that, in three dimensions, an ideal gas with bosonic statistics presents a Bose–Einstein condensation for sufficiently low temperature. Discuss if the same phenomenon can take place in one and two dimensions. Study if a Bose–Einstein condensation can happen for a harmonic oscillator in dimension $d = 1, 2, 3$.

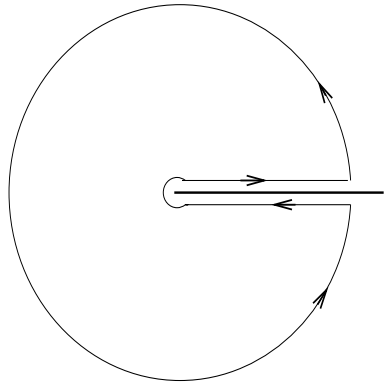


Fig. 1.10 *Integration contour C.*

14. Dimensional regularization

Let d the dimension of the space. Discuss the convergence of the integral

$$I(d) = \int_0^\infty \frac{r^{d-1}}{r^2+1}\, dr$$

by varying d.

a Determine, in its convergent domain, the exact expression of the integral as a function of d and identify the position of its poles.

b Analytically continue the definition of the integral in any other domain.

c Compute its value for $d = \frac{1}{3}$ and $d = \pi$.

Hint. Consider the integral in the complex plane

$$\oint_C \frac{z^{d-1}}{z^2+1}\, dz$$

where C is the contour shown in Fig. 1.10.

15. Braid group

The braid group on n strands, denoted by B_n, is a set of operations which has an intuitive geometrical representation, and in a sense generalizes the symmetric group S_n. Here, n is a natural number. Braid groups find applications in knot theory, since any knot may be represented as the closure of certain braids. From the algebraic point of view, the braid group is represented in terms of generators β_i, with $1 \le i \le (n-1)$; β_i is a counterclockwise exchange of the i-th and $(i+1)$-th strands. β_i^{-1} is therefore a clockwise exchange of the i-th and $(i+1)$-th strands. The generators β_i satisfy the defining relations, called Artin relations (see Fig. 1.11):

$$\begin{aligned} \beta_i\, \beta_j &= \beta_j\, \beta_i && \text{for } \mid i-j \mid \geq 2 \\ \beta_i\, \beta_{i+1}\, \beta_i &= \beta_{i+1}\, \beta_i\, \beta_{i+1} && \text{for } \quad 1 \le i \le n-1. \end{aligned}$$

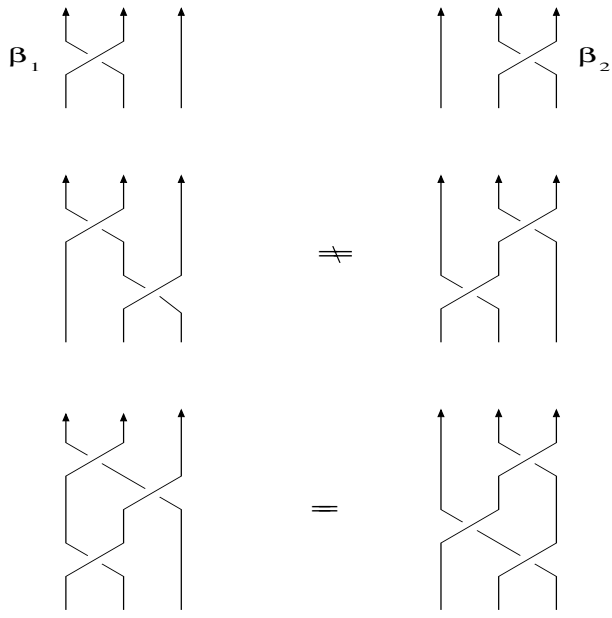


Fig. 1.11 Top: *The two elementary braid operations* β_1 *and* β_2. **Middle**: *Graphical proof that* $\beta_2\beta_1 \neq \beta_1\beta_2$, *hence the braid group is not abelian.* **Bottom**: *the Yang–Baxter relation of the braid group.*

The second is also called the Yang–Baxter equation. The only difference from the permutation group is that $\beta_i^2 \neq 1$, but this is an enormous difference: while the permutation group is finite (the dimension is $n!$), the braid group is infinite, even for just two strands. The irreducible representation of the braid group can be given in terms of $g \times g$ dimensional unitary matrices, $\beta_i \to \gamma_i$, where the matrices γ_i satisfy the Artin relations.

a Consider the group B_3. Prove that

$$\gamma_1 = \begin{pmatrix} e^{-7i\pi/10} & 0 \\ 0 & -e^{-3i\pi/10} \end{pmatrix}, \qquad \gamma_2 = \begin{pmatrix} -\tau e^{-i\pi/10} & -i\sqrt{\tau} \\ -i\sqrt{\tau} & -\tau e^{i\pi/10} \end{pmatrix}$$

provide a representation of the Artin relations. Here $\tau = (\sqrt{5}-1)/2$, which satisfies $\tau^2 + \tau = 1$.

b Both matrices γ_i $(i = 1, 2)$ are matrices of $SU(2)$ and can be written as

$$\gamma_i = \exp\left[i\frac{\theta_i \vec{n}_i}{2} \cdot \vec{\sigma}\right]$$

where σ_j $(j = 1, 2, 3)$ are the Pauli matrices and θ_i is the angle of rotation around the axis $\vec{n_i}$. Identify the angles and the axes of rotation that correspond to γ_1 and γ_2.

c By multiplying the γ_i (and their inverse) in a sequence of L steps, as in the example below

$$A_L = \underbrace{\gamma_1\,\gamma_2\gamma_1^{-1}\gamma_2\ldots\gamma_1}_{L}\,.$$

one generates another matrix A_L of SU(2), identified by the angle α of rotation around an axis $\vec{n}$, $A_L = \exp\left[i\frac{\alpha\vec{n}}{2}\cdot\vec{\sigma}\right]$. Argue that making L sufficiently large, one can always find a string of γ_i and its inverse that approximates with an arbitrary precision any matrix of $SU(2)$.

2
One-dimensional Systems

If our highly pointed Triangles of the Soldier class are formidable, it may be readily inferred that far more formidable are our Women. For, if a Soldier is a wedge, a Woman is a needle.

Edwin A. Abbott, *Flatland*

In this chapter we present several approaches to get the exact solution of the one-dimensional Ising model. As already mentioned, the one-dimensional case does not present a phase transition at a finite value of the temperature. However we will show that the origin $T = B = 0$ of the phase diagram may nevertheless be regarded as a critical point: by using appropriate variables, one can define the set of critical exponents and verify that the scaling relations are indeed satisfied.

In this chapter we also discuss three different generalizations of the Ising model: the first is given by the q-state Potts model, a system that is invariant under the permutation group $\mathcal{S}_q$ of q objects; the second is provided by a system of spins with n components, invariant under the continuum group of transformations $O(n)$; the third one is the so-called $Z(n)$ model, i.e. a spin system that is invariant under the set of the discrete rotations associated to the n-th roots of unity. We compute the partition function of all these models, pointing out their interesting properties. Finally, we analyze the thermodynamics of the so-called *Feynman gas*, i.e. a one-dimensional gas of particles with a short-range potential $V(\mid x_i - x_j \mid)$: the results of this analysis will be useful when we face in later chapters the study of the correlation functions of the two-dimensional models.

2.1 Recursive Approach

The first method we are going to introduce is based on a recursive approach: it permits us to obtain the exact solution of the one-dimensional Ising model in the absence of an external magnetic field.

Consider a linear chain of N Ising spins (see Fig. 2.1) in the absence of an external magnetic field, with free boundary conditions on the first and the last spin of the chain. The more general hamiltonian of such a system is given by

$$\mathcal{H} = -\sum_{i=1}^{N-1} J_i \sigma_i \, \sigma_{i+1},$$

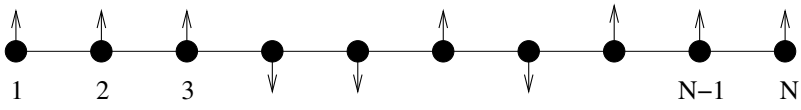


Fig. 2.1 *Linear chain of N Ising spins.*

with an interaction J_i that may change from site to site. The partition function is expressed by

$$Z_N = \sum_{\sigma_1=-1}^{1} \sum_{\sigma_2=-1}^{1} \cdots \sum_{\sigma_N=-1}^{1} \exp\left(\sum_{i=1}^{N-1} \mathcal{J}_i \sigma_i \sigma_{i+1} \right), \tag{2.1.1}$$

where we have introduced the notation $\mathcal{J}_i = \beta J_i$. The recursive method consists of adding an extra spin to the chain and expressing the resulting partition function Z_{N+1} in terms of the previous Z_N. By adding another spin, we have

$$Z_{N+1} = \sum_{\sigma_1=-1}^{1} \sum_{\sigma_2=-1}^{1} \cdots \sum_{\sigma_N=-1}^{1} \exp\left(\sum_{i=1}^{N-1} \mathcal{J}_i \sigma_i \sigma_{i+1}) \right) \sum_{\sigma_{N+1}=-1}^{1} \exp\left(\mathcal{J}_N \sigma_N \sigma_{N+1} \right). \tag{2.1.2}$$

The last sum can be easily computed

$$\sum_{\sigma_{N+1}=-1}^{1} \exp\left(\mathcal{J}_N \sigma_N \sigma_{N+1} \right) = e^{\mathcal{J}_N \sigma_N} + e^{-\mathcal{J}_N \sigma_N} = 2\cosh(\mathcal{J}_N \sigma_N) = 2\cosh \mathcal{J}_N,$$

and the result is *independent* of σ_N, a particularly important circumstance. This permits us to rewrite eqn (2.1.2) as

$$Z_{N+1} \;=\; (2\cosh \mathcal{J}_N)\, Z_N,$$

and the iteration of this relation leads to

$$Z_{N+1} \;=\; \left(2^N \prod_{i=1}^{N} \cosh \mathcal{J}_i \right) Z_1.$$

Since the partition function Z_1 of an isolated spin is equal to the number of its states, i.e. $Z_1 = 2$, the exact expression of the partition function of N spins is given by

$$Z_N \;=\; 2^N \prod_{i=1}^{N-1} \cosh \mathcal{J}_i. \tag{2.1.3}$$

To see whether there is a critical value T_c of the temperature (below which the system presents a magnetized phase), it is useful to compute the two-spin correlation function

$$G^{(2)}(r) \;=\; \langle \sigma_k \sigma_{k+r} \rangle \;=\; Z_N^{-1} \sum_{\{\sigma\}} \sigma_k \sigma_{k+r} \, \exp\left(\sum_{i=1}^{N-1} \mathcal{J}_i \sigma_i \sigma_{i+1} \right), \tag{2.1.4}$$

where the first sum stands for a concise way of expressing the sum on the ± 1 values of all the N spins. If $r = 1$, the correlation function is obtained by taking the derivative

$$G^{(2)}(1) \;=\; \langle \sigma_k \sigma_{k+1} \rangle \;=\; Z_N^{-1} \frac{\partial}{\partial \mathcal{J}_k} \sum_{\{\sigma\}} \exp\left(\sum_{i=1}^{N-1} \mathcal{J}_i \sigma_i \sigma_{i+1} \right).$$

Thanks to the identity $\sigma_i^2 = 1$, valid for the Ising spins, the formula can be easily generalized to arbitrary r

$$Z_N \, G^{(2)}(r) \;=\; \frac{\partial}{\partial \mathcal{J}_k} \frac{\partial}{\partial \mathcal{J}_{k+1}} \cdots \frac{\partial}{\partial \mathcal{J}_{k+r-1}} Z_N. \tag{2.1.5}$$

Substituting in this formula eqn (2.1.3), one has

$$G^{(2)}(r) \;=\; \prod_{i=1}^{r} \tanh \mathcal{J}_{k+i-1}. \tag{2.1.6}$$

This expression makes it possible to check in an easy way the validity of simple physical intuition. It correctly predicts that, by taking the limit $\mathcal{J}_i \to 0$ that breaks the chain into two separate blocks, if the site i is placed between k and $k + r$, the correlation function vanishes; vice versa, if the site i is external to the interval $(k, k + r)$, the correlation function is unaffected by the limit $\mathcal{J}_i \to 0$.

If the system is homogeneous, with the same coupling constant J for all spins, we have the simpler expression

$$G^{(2)}(r) \;=\; (\tanh \mathcal{J})^r \tag{2.1.7}$$

that can be written in a scaling form as

$$G^{(2)}(r) \;=\; \exp\left[-r/\xi \right].$$

The correlation length ξ, in units of the lattice space a, is given by

$$\xi(\mathcal{J}) \;=\; -\frac{1}{\log \tanh \mathcal{J}}. \tag{2.1.8}$$

We can use this expression for ξ to identify the possible critical points of the system, since ξ diverges at a phase transition. It is easy to see that ξ has only one singular point, given by

$$\mathcal{J} \;=\; \beta J = \frac{J}{kT} \to \infty,$$

i.e. $T = 0$ (if J is a finite quantity). One arrives at the same conclusion by analyzing the possibility of having a non-zero expectation value of the spin, i.e. a non-vanishing limit

$$|\, \langle \sigma \rangle \,|^2 = \lim_{r \to \infty} G^{(2)}(r). \tag{2.1.9}$$

Since for finite βJ the hyperbolic tangent entering $G^{(2)}(r)$ is always less than 1, the spontaneous magnetization always vanishes, except for the limiting case $\beta J = \infty$, i.e. $T = 0$.

The absence of an ordered phase in a finite interval of the temperature T of the one-dimensional Ising model can be readily explained by some simple thermodynamic considerations. In fact, let's assume that at a sufficiently low temperature the system is a complete ordered state, i.e. with all spins aligned, for istance, $\sigma_i = 1$. The energy of this configuration is $E_0 = -(N-1)J$. The configurations of the system with the next higher energy are those in which an entire spin block is inverted at an arbitrary point of the chain (see Fig. 2.2). Their number is $N-1$ (it is equal to the number of sites where this inversion of the spins can take place) and their energy is $E = E_0 + 2J$. At a temperature T, the variation of the free energy induced by these excitations is expressed by

$$\Delta F \;=\; \Delta E - T\,\Delta S \;=\; 2J - kT\ln(N-1), \tag{2.1.10}$$

and, for N sufficiently large, it is always negative for all value of $T \neq 0$. Hence, the ordered state of the system is not the configuration that minimizes the free energy. Since the configurations with inverted spin blocks disorder the system, the ordered phase of the one-dimensional Ising model is always unstable for $T \neq 0$.

The absence of a spontaneous magnetization at a finite T does not imply, however, the absence of a singularity at $T = 0$. Let's compute, for instance, the magnetic susceptibility at $B = 0$ by using the fluctuation-dissipation theorem

$$\chi(T, B=0) \;=\; \frac{\beta}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\langle\sigma_i\,\sigma_j\rangle. \tag{2.1.11}$$

For simplicity, consider the homogeneous case $\langle\sigma_i\,\sigma_j\rangle = v^{|i-j|}$, with $v = \tanh\beta J$. In the sum above, there are

- N terms, for which $\mid i-j \mid = 0$. Each of them gives rise to a factor $v^0 = 1$.
- $2(N-1)$ terms, for which $\mid i-j \mid = 1$. They correspond to the $N-1$ next neighboring pairs of spins of the open chain and each of them brings a term v^1.

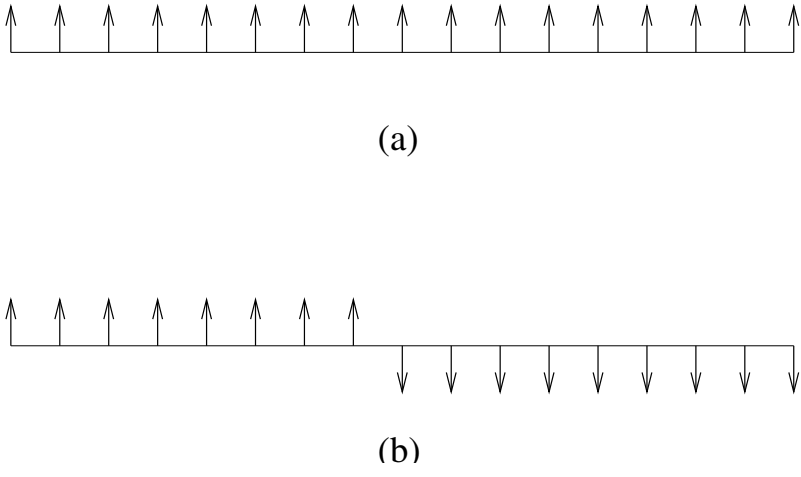


Fig. 2.2 *(a) Ordered low-energy state; (b) excited state.*

- $2(N-2)$ terms, for which $\mid i-j \mid = 2$ and a term v^2, and so on, till we arrive at the last two terms for which $\mid i-j \mid = N-1$, each of them bringing a factor v^{N-1}.

Hence, the double sum (2.1.11) can be expressed as

$$\chi(T, B=0) \;=\; \frac{\beta}{N}\left(N + 2\sum_{k=1}^{N-1}(N-k)\,v^k\right).$$

By using

$$\sum_{k=1}^{N-1} v^k \;=\; \frac{1-v^N}{1-v},$$

$$\sum_{k=1}^{N-1} k v^k \;=\; v\frac{\partial}{\partial v}\sum_{k=1}^{N-1} v^k,$$

we arrive at

$$\chi(T, B=0) \;=\; \frac{\beta}{N}\left[N\left(1+\frac{2v}{1-v}\right) - \frac{2v(1-v^N)}{(1-v)^2}\right].$$

This expression can be simplified by taking the thermodynamic limit $N \to \infty$

$$\chi(T, B=0) \;=\; \beta\,\frac{1+v}{1-v} = \beta\, e^{2J/kT},$$

and this expression presents an essential singularity for $T \to 0$.

It is also interesting to study the case $J < 0$ that corresponds to the *antiferromagnetic* situation. In such a case, the minimum of the energy of the system is realized by those configurations where the spins alternate their values by moving from one site to the next one. The two-point correlation function of the spins is given by eqn (2.1.7) also in the antiferromagnetic case. However, for negative values of J, it changes its sign by changing the lattice sites, as shown in Fig. 2.3. The oscillating behavior of this function is responsable for a partial cancellation of the terms entering the series (2.1.11) of the magnetic susceptibility that indeed remains finite for all values of temperature.

Using the previous formulas, we can explicitly compute the mean energy U and the specific heat C at $B = 0$. For the mean energy we have

$$\langle U \rangle \;=\; -\frac{\partial}{\partial\beta}\left(\ln\, Z_N(T, B=0)\right) = -\sum_{i=1}^{N-1} J_i \tanh \mathcal{J}_i = -J(N-1)\tanh \mathcal{J},$$

where the last identity holds in the homogeneous case, while for the specific heat we get

$$C(T, B=0) \;=\; \frac{\partial\langle U\rangle}{\partial T} = k(N-1)\left(\frac{\mathcal{J}}{\cosh \mathcal{J}}\right)^2. \tag{2.1.12}$$

The plot of this function is shown in Fig. 2.4. Similar functions, with a pronounced maximum, are obtained for the specific heat of all those substances which have only one energy gap ΔE and, in the literature, are known as Schottky curves. The reason why the one-dimensional Ising model is equivalent to a system with only one energy

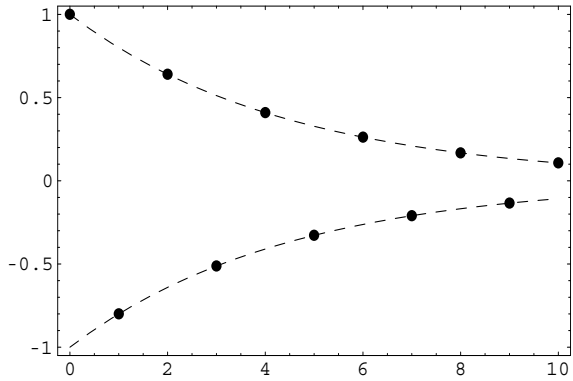


Fig. 2.3 *Two-point correlation function of the spins in the ferromagnetic case (upper curve) and in the antiferromagnetic case (lower curve).*

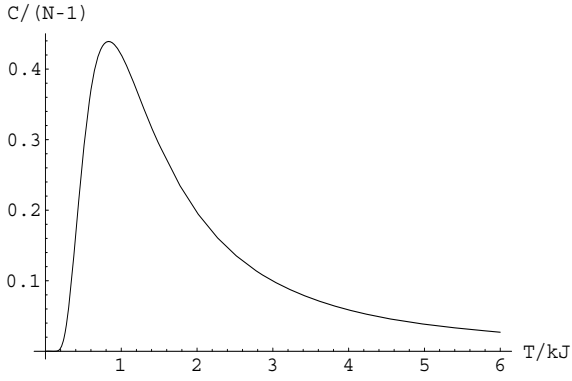


Fig. 2.4 *Specific heat of the one-dimensional Ising model versus temperature.*

gap ΔE will become clear after the discussion in the next section on the transfer matrix of the model.

By using eqn (2.1.3), we can also compute the entropy of the system

$$S(T, B=0) = \frac{\partial}{\partial T}\left[\frac{1}{\beta}\ln Z_N\right] = \\ = k\left[N\ln 2 + (N-1)\ln\cosh\mathcal{J} - (N-1)\mathcal{J}\tanh\mathcal{J}\right]. \tag{2.1.13}$$

The plot of the entropy is in Fig. 2.5. For $T \to 0$, the entropy goes correctly to the value $k\ln 2$: at $T = 0$, there are in fact only two effective states of the system, the one in which all spins are up and the other one in which all spins are down. For $T \to \infty$, we have instead $S \to Nk\ln 2$: in this limit all spins are free to fluctuate in an independent way and, correspondingly, the available number of states of the systems is given by 2^N.

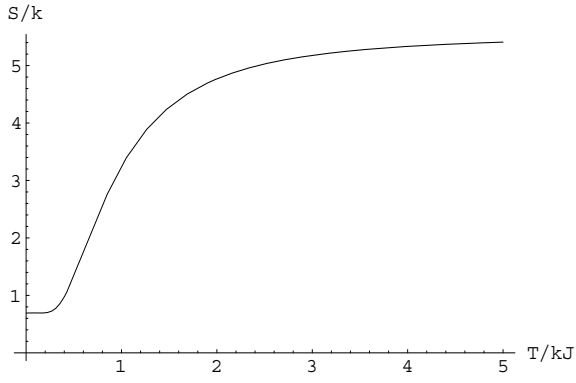


Fig. 2.5 *Entropy versus temperature.*

2.2 Transfer Matrix

The exact solution of the one-dimensional Ising model can be obtained by using the alternative method of the transfer matrix. This method presents a series of advantages: unlike the recursive method, it also can be applied when there is an external magnetic field. Moreover, it has many points in common with a discrete formulation of quantum mechanics, in particular the Feynman formulation in terms of a path integral. The transfer matrix method relies on a set of ideas that go beyond the application to the one-dimensional case and permits us to show the remarkable relationship that links classical systems of statistical mechanics in d dimensions with quantum systems in $(d-1)$, as will be discussed in more detail in Chapter 7. In the two-dimensional case, for instance, it permits us to obtain the exact solution of the Ising model in the absence of an external magnetic field (see Chapter 6).

To study the one-dimensional case, let us consider once again a chain of N spins. For simplicity, we consider here the homogeneous case, in which there is only one coupling constant J, with hamiltonian

$$\mathcal{H} = -J \sum_{i=1}^{N-1} \sigma_i \, \sigma_{i+1} - B \sum_{i=1}^{N} \sigma_i. \tag{2.2.1}$$

We firstly analyze the periodic boundary condition case while more general boundary conditions will be considered later.

2.2.1 Periodic Boundary Conditions

Assuming periodic boundary conditions, the chain has a ring geometry, implemented by the condition

$$\sigma_i \equiv \sigma_{N+i}.$$

The transfer matrix method is based on the observation that the sum on the spin configurations can be equivalently expressed in terms of a product of 2×2 matrices, as follows

$$Z_N = \sum_{\{\sigma\}} V(\sigma_1, \sigma_2)\, V(\sigma_2, \sigma_3) \cdots V(\sigma_N, \sigma_1), \tag{2.2.2}$$

where the matrix elements of $V(\sigma, \sigma')$ are defined by

$$V(\sigma, \sigma') = \exp\left[\mathcal{J}\sigma\sigma' + \frac{1}{2}\mathcal{B}(\sigma + \sigma')\right], \tag{2.2.3}$$

with $\mathcal{J} = \beta J$ and $\mathcal{B} = \beta B$. Explicitly

$$\begin{aligned} \langle +1 \mid V \mid +1 \rangle &= e^{\mathcal{J}+\mathcal{B}}; \\ \langle -1 \mid V \mid +1 \rangle &= e^{-\mathcal{J}}; \\ \langle +1 \mid V \mid -1 \rangle &= e^{-\mathcal{J}}; \\ \langle -1 \mid V \mid -1 \rangle &= e^{\mathcal{J}-\mathcal{B}}, \end{aligned}$$

and therefore V can be written as

$$V = \begin{pmatrix} e^{\mathcal{J}+\mathcal{B}} & e^{-\mathcal{J}} \\ e^{-\mathcal{J}} & e^{\mathcal{J}-\mathcal{B}} \end{pmatrix}. \tag{2.2.4}$$

It is easy to see that the product of the matrix V correctly reproduces the Boltzmann weights of the Ising model configurations. In this approach, the configuration space of a single spin may be regarded as the Hilbert space of a *two-state quantum system*: the states will be denoted by $\mid +1\rangle$ and $\mid -1\rangle$, and the completeness relation is expressed by the formula

$$\sum_{\sigma=\pm 1} \mid \sigma\rangle\langle\, \sigma \mid = 1. \tag{2.2.5}$$

The original one-dimensional lattice can be seen as the temporal axis, along which the quantum dynamics of the two-state system takes place. In more detail, the transfer matrix V plays the role of the quantum time evolution operator for the time interval $\Delta t = a$ (see Fig. 2.6)

$$\mid \sigma_{i+1}\rangle = V \mid \sigma_i\rangle \equiv e^{-a\mathbf{H}} \mid \sigma_i\rangle. \tag{2.2.6}$$

In this formula $\mathbf{H}$ expresses the quantum hamiltonian which must not be confused with the original classical hamiltonian $\mathcal{H}$ given in eqn (2.2.1). By adopting this scheme based on a two-state Hilbert space, it becomes evident that the one-dimensional Ising model presents only one energy gap $\Delta\mathcal{E}$: one has, then, a natural explanation of the Schottky form of the specific heat, discussed in the previous section.

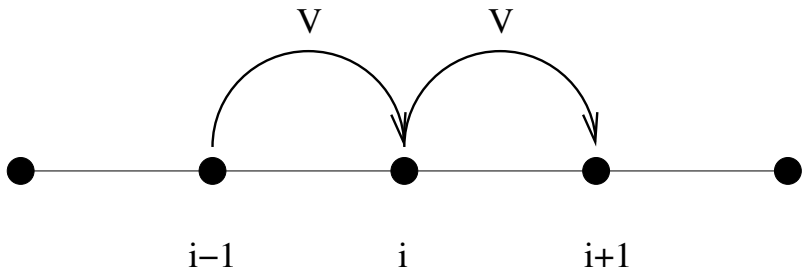


Fig. 2.6 *Transfer matrix as quantum time evolution operator.*

Quantum hamiltonian. It is an interesting exercise to find an explicit expression for the quantum hamiltonian **H**. Let us recall that, in the linear space of 2×2 matrices, a basis is provided by the identity matrix $\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and by the Pauli matrices $\hat{\sigma}_i$

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.2.7}$$

They satisfy

$$\{\hat{\sigma}_k, \hat{\sigma}_l\} = 2\delta_{kl}, \qquad [\hat{\sigma}_k, \hat{\sigma}_l] = 2\, i\, \epsilon_{klm}\, \hat{\sigma}_m \tag{2.2.8}$$

where $\{a, b\} = ab + ba$, $[a, b] = ab - ba$ and ϵ_{klm} is the antisymmetric tensor in all three indices, with $\epsilon_{123} = 1$.

In terms of these matrices, V can be written as

$$V = \left(e^{\mathcal{J}} \cosh \mathcal{B}\right)\ \mathbf{1} + e^{-\mathcal{J}}\ \hat{\sigma}_1 + \left(e^{\mathcal{J}} \sinh \mathcal{B}\right)\ \hat{\sigma}_3. \tag{2.2.9}$$

Let us determine the constants C, c_1, c_2, c_3 so that V is expressed as

$$V = C \exp\left[\, c_1\hat{\sigma}_1 + c_2\hat{\sigma}_2 + c_3\hat{\sigma}_3 \,\right]. \tag{2.2.10}$$

By making a series expansion of the exponential

$$\exp\left[c_1\hat{\sigma}_1 + c_2\hat{\sigma}_2 + c_3\hat{\sigma}_3\right] = \sum_{k=0}^{\infty} \frac{\left(c_1\hat{\sigma}_1 + c_2\hat{\sigma}_2 + c_3\hat{\sigma}_3\right)^k}{k!}, \tag{2.2.11}$$

and using the anticommutation rule (2.2.8), it is easy to see that we arrive at

$$\begin{aligned} \left(c_1\hat{\sigma}_1 + c_2\hat{\sigma}_2 + c_3\hat{\sigma}_3\right)^{2n} &= r^{2n+1}, \\ \left(c_1\hat{\sigma}_1 + c_2\hat{\sigma}_2 + c_3\hat{\sigma}_3\right)^{2n+1} &= \left(c_1\hat{\sigma}_1 + c_2\hat{\sigma}_2 + c_3\hat{\sigma}_3\right)\, r^{2n}, \end{aligned}$$

where $r = \sqrt{c_1^2 + c_2^2 + c_3^2}$. By summing the series (2.2.11), eqn (2.2.10) becomes

$$V \;=\; C\left[\cosh r\;\mathbf{1} \;+\; \frac{\sinh r}{r}\;(c_1\,\hat{\sigma}_1 + c_2\,\hat{\sigma}_2 + c_3\,\hat{\sigma}_3)\right].$$

Comparing this expression with eqn (2.2.9), we have

$$\begin{aligned}
C\,\cosh r \;&=\; e^{\mathcal{J}}\,\cosh\mathcal{B},\\
C\,\frac{\sinh r}{r}\,c_1 \;&=\; e^{-\mathcal{J}},\\
C\,\frac{\sinh r}{r}\,c_2 \;&=\; 0,\\
C\,\frac{\sinh r}{r}\,c_3 \;&=\; e^{\mathcal{J}}\,\cosh\mathcal{B},
\end{aligned}$$

from which it immediately follows that $c_2 = 0$. From the ratio between the fourth and the second equation, we have

$$c_3 \;=\; c_1\,e^{2\mathcal{J}}\,\sinh\mathcal{B}.$$

Summing the square of the second and the fourth equations and subtracting the square of the first equation, we get

$$C^2 \;=\; 2\,\sinh 2\mathcal{J},$$

i.e. $C = \sqrt{2\sinh 2\mathcal{J}}$. Finally, by taking the ratio of the square of the first and the second equations and using eqn (2.2.1), c_1 is given by the solution of the trascendental equation

$$\tanh\left[c_1\,\sqrt{1 + e^{4\mathcal{J}}\,\sinh^2\mathcal{B}}\right] \;=\; \frac{\sqrt{1 + e^{4\mathcal{J}}\,\sinh^2\mathcal{B}}}{e^{2\mathcal{J}}\,\cosh\mathcal{B}}. \tag{2.2.12}$$

Hence, the quantum hamiltonian $\mathbf{H}$ is given by

$$\mathbf{H} \;=\; -\frac{1}{a}\left[\left(\frac{1}{2}\log(\sinh 2\mathcal{J})\right) \;+\; c_1\left(\hat{\sigma}_1 + e^{2\mathcal{J}}\,\sinh\mathcal{B}\,\hat{\sigma}_3\right)\right], \tag{2.2.13}$$

where c_1 is the solution of (2.2.12). This expression simplifies when $B = 0$

$$\mathbf{H} \;=\; -\frac{1}{a}\left[\left(\frac{1}{2}\log(\sinh 2\mathcal{J})\right) \;+\; c_1\;\hat{\sigma}_1\right], \tag{2.2.14}$$

with $\tanh c_1 \;=\; e^{-2\mathcal{J}}$. It is interesting to study the limit $a \to 0$ of this expression, the so-called *hamiltonian limit*. To do that, it is convenient to subtract the first term of the hamiltonian (2.2.14), which corresponds anyhow to an additive constant. One can get a finite expression for $\mathbf{H}$ in the limit $a \to 0$ only by taking the simultaneous limit $\mathcal{J} \to \infty$, with the combination $y \equiv a\,e^{2\mathcal{J}}$ kept fixed. This relationship between

the coupling constant $\mathcal{J}$ and the lattice space a is perhaps the simplest equation of the renormalization group: it is the one that guarantees that the physical properties of the system remain the same even in the limit $a \to 0$. Consider, for instance, the correlation length ξ

$$\xi = -\frac{a}{\log(\tanh \mathcal{J})};$$

ξ remains finite in the limit $a \to 0$ only by increasing correspondingly the coupling constant among the spin, keeping fixed their combination y.

Let's come back to the computation of the partition function. By using eqn (2.2.2) and the completeness (2.2.5), one has

$$\begin{aligned} Z_N &= \sum_{\sigma_1=\pm 1} \sum_{\sigma_2=\pm 1} \cdots \sum_{\sigma_N=\pm 1} \langle \sigma_1 \mid V \mid \sigma_2 \rangle \langle \sigma_2 \mid V \mid \sigma_3 \rangle \cdots \langle \sigma_N \mid V \mid \sigma_1 \rangle \\ &= \sum_{\sigma_1=\pm 1} \langle \sigma_1 \mid V^N \mid \sigma_1 \rangle = \operatorname{Tr} V^N. \end{aligned} \tag{2.2.15}$$

The fact that Z_N is expressed in terms of the trace of the N-th power of the operator V is clearly due to the periodic boundary conditions we adopted. The simplest way to compute the trace of V^N consists of bringing V into a diagonal form. Being an hermitian matrix, it can be diagonalized by means of a unitary matrix U

$$U^{-1} V U = \mathcal{D} = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix},$$

with $\lambda_+ \geq \lambda_-$. If we define the quantity ϕ by the relation

$$\cot 2\phi = e^{2\mathcal{J}} \sinh \mathcal{B}, \tag{2.2.16}$$

the explicit expression for U is given by

$$U = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}. \tag{2.2.17}$$

Since the trace of a product of matrices is cyclic, by inserting in (2.2.15) the identity matrix $\mathbf{1}$ in the form $UU^{-1} = \mathbf{1}$ we have

$$\operatorname{Tr} V^N = \operatorname{Tr} U U^{-1} V^N = \operatorname{Tr} U^{-1} V^N U = \operatorname{Tr} \mathcal{D}^N = \lambda_+^N + \lambda_-^N. \tag{2.2.18}$$

We need now to determine explicitly the two eigenvalues: by an elementary computation, they are given by

$$\lambda_\pm = e^{\mathcal{J}} \cosh \mathcal{B} \pm \sqrt{e^{2\mathcal{J}} \cosh^2 \mathcal{B} - 2\sinh(2\mathcal{J})}. \tag{2.2.19}$$

The free energy per unit spin is then expressed by

$$F(\beta, B) = -\frac{1}{\beta N} \ln Z_N = -\frac{1}{\beta} \left\{ \ln \lambda_+ + \frac{1}{N} \ln \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \right\}. \tag{2.2.20}$$

In the thermodynamic limit $N \to \infty$, taking into account that $\lambda_+ > \lambda_-$ for any value of B, the free energy is determined only by the larger eigenvalue λ_+:

$$F(\beta, B) \;=\; -\frac{1}{\beta}\,\ln\left[e^{\mathcal{J}}\,\cosh\mathcal{B} + \sqrt{e^{2\mathcal{J}}\,\cosh^2\mathcal{B} - 2\sinh(2\mathcal{J})}\right]. \tag{2.2.21}$$

Taking the derivative with respect to B of this expression, we obtain the mean value of the magnetization

$$\langle\sigma\rangle \;=\; \frac{e^{\mathcal{J}}\,\sinh\mathcal{B}}{\sqrt{e^{2\mathcal{J}}\cosh^2\mathcal{B} - 2\sinh 2\mathcal{J}}}. \tag{2.2.22}$$

The graph of this function, for different values of the temperature, is given in Fig. 2.7.

The free energy (2.2.21) is an analytic function of B and T for all real values of B and for positive values of T. The magnetization is an analytic function of B that vanishes if $B = 0$. The system does not then present any phase transition at finite values of T, as we have previously seen. However, in the limit $T \to 0$ at B finite, the magnetization presents a discontinuity, expressed by

$$\langle\sigma\rangle \;=\; \epsilon(B), \tag{2.2.23}$$

where the function $\epsilon(x)$ is defined by

$$\epsilon(x) \;=\; \begin{cases} 1 & \text{if } x > 0; \\ 0 & \text{if } x = 0; \\ -1 & \text{if } x < 0. \end{cases}$$

Correlation function. The transfer matrix method can also be applied to compute the correlation functions of the spins. To this aim, it is convenient to write the correlator as

$$\langle\sigma_1\sigma_{r+1}\rangle \;=\; Z_N^{-1}\sum_{\{\sigma\}}\sigma_1 V(\sigma_1,\sigma_2)\cdots\sigma_{r+1}V(\sigma_{r+1},\sigma_{r+2})\cdots V(\sigma_N,\sigma_1). \tag{2.2.24}$$

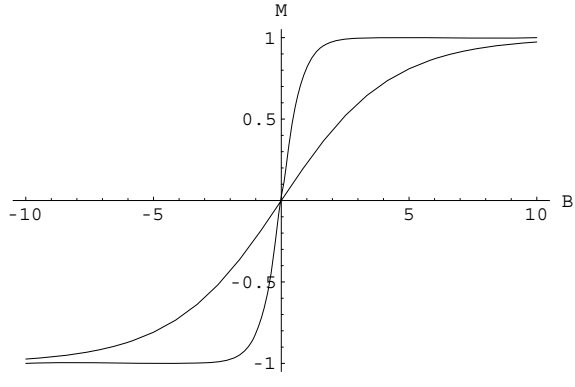


Fig. 2.7 *Magnetization versus the magnetic field B, for different values of the temperature.*

Introducing the diagonal matrix $\mathcal{S}$, with matrix elements

$$\mathcal{S}_{\sigma,\sigma'} \;=\; \sigma\,\delta_{\sigma,\sigma'},$$

i.e.

$$\mathcal{S} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

eqn (2.2.24) can be written as

$$\langle \sigma_1 \sigma_{r+1} \rangle \;=\; Z_N^{-1} \mathrm{Tr}\, \left(\mathcal{S}\, V^r\, \mathcal{S}\, V^{N-r} \right). \tag{2.2.25}$$

For the expectation value of σ, we have

$$\langle \sigma \rangle \;=\; Z_N^{-1}\, \mathrm{Tr}\; \mathcal{S}\, V^N. \tag{2.2.26}$$

Using the unitary matrix U that diagonalizes V, we get

$$U^{-1}\, \mathcal{S}\, U \;=\; \begin{pmatrix} \cos 2\phi & -\sin 2\phi \\ -\sin 2\phi & -\cos 2\phi \end{pmatrix}.$$

Substituting this expression and the diagonal form of V in eqns (2.2.25) and (2.2.26), in the limit $N \to \infty$ we have

$$\langle \sigma_i \sigma_{i+r} \rangle \;=\; \cos^2 2\phi + \sin^2 2\phi \left(\frac{\lambda_-}{\lambda_+} \right)^r,$$

$$\langle \sigma_i \rangle \;=\; \cos 2\phi.$$

Hence, the connected two-point correlation function is given by

$$G_c^{(2)}(r) \;=\; \langle \sigma_i \sigma_{i+r} \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle = \sin^2 2\phi \left(\frac{\lambda_-}{\lambda_+} \right)^r. \tag{2.2.27}$$

Besides its elegance, this formula points out an important conceptual aspect of general validity, namely that the correlation length of a statistical system is determined by the ratio of the two largest eigenvalues of the transfer matrix

$$\xi \;=\; \frac{1}{\ln \lambda_+ / \lambda_-}. \tag{2.2.28}$$

2.2.2 Other Boundary Conditions: Boundary States

Let's now proceed to the computation of the partition function of the one-dimensional Ising model with N spins but with boundary conditions of type (a, b) relative to the two spins at the end of the chain. The quantum mechanical interpretation given for the transfer matrix is particularly useful to solve this problem. In fact, the boundary condition of type (a) for the first spin of the chain can be implemented by associating to this spin a special state $\mid a \,\rangle$ of the Hilbert space. Analogously, the boundary condition of type (b) for the last spin of the chain can be put in relation with another vector $\mid b \,\rangle$. These two vectors play the role of the initial and final states respectively of the

time evolution of the corresponding quantum system and, for that reason, they are called *boundary states*. Hence, in order to compute the partition function $Z^{(a,b)}$, we have simply to evaluate the matrix element of the quantum time evolution operator between the initial $\langle a \mid$ and the final state $\mid b\rangle$

$$\begin{aligned} Z_N^{(a,b)} &= \sum_{\sigma_2=\pm 1} \cdots \sum_{\sigma_{N-1}=\pm 1} \langle a \mid V \mid \sigma_2\rangle\langle\sigma_2 \mid V \mid \sigma_3\rangle \cdots \langle\sigma_{N-1} \mid V \mid b\rangle \\ &= \langle a \mid V^{N-1} \mid b\rangle. \end{aligned} \tag{2.2.29}$$

This expression can be made explicit by using the unitary matrix U that diagonalizes V. By inserting in (2.2.29) both on the right and left sides of the operator V the identity operator as $UU^{-1} = \mathbf{1}$, we have

$$Z^{(a,b)} = \langle a \mid U\, U^{-1}\, V^{N-1}\, U\, U^{-1} \mid b\rangle = \langle a \mid U\, \mathcal{D}^{N-1}\, U^{-1} \mid b\rangle. \tag{2.2.30}$$

It is interesting to consider some explicit examples. Consider, for instance, the partition function with boundary conditions $\sigma_1 = \sigma_N = 1$. In this case we have

$$\mid a\rangle = \mid b\rangle = \mid +\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Using the expressions for U, $\mathcal{D}$, and $\mid +\rangle$ to compute the matrix element (2.2.30), we have

$$\begin{aligned} Z_N^{++} &= \langle + \mid U\, \mathcal{D}^{N-1}\, U^{-1} \mid +\rangle \qquad (2.2.31) \\ &= (1,0) \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \lambda_+^{N-1} & 0 \\ 0 & \lambda_-^{N-1} \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \lambda_+^{N-1} \cos^2\phi + \lambda_-^{N-1} \sin^2\phi. \end{aligned}$$

It is easy to obtain the partition functions also in other cases: for instance, with an obvious choice of the notation, we have

$$\begin{aligned} Z_N^{--} &= \lambda_+^{N-1} \sin^2\phi + \lambda_-^{N-1} \cos^2\phi \ ; \\ Z_N^{+-} &= Z_N^{-+} = \sin\phi\, \cos\phi\, (\lambda_+^{N-1} - \lambda_-^{N-1}), \end{aligned} \tag{2.2.32}$$

where the boundary condition $\sigma = -1$ is expressed by the vector

$$\mid -\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

For free boundary conditions, the corresponding vector is given by

$$\mid \mathrm{f}\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

and the corresponding partition function is

$$Z_N^{ff} = Z_N^{++} + Z_N^{--} + 2Z_N^{+-} = \lambda_+^{N-1} + \lambda_-^{N-1} + \sin 2\phi \left(\lambda_+^{N-1} - \lambda_-^{N-1}\right). \tag{2.2.33}$$

When $B = 0$ (which corresponds to $\phi = \pi/4$), this expression coincides with (2.1.3), obtained by the recursive method.

The boundary conditions should not affect the bulk properties of the system when N is very large. Indeed, in the thermodynamic limit $N \to \infty$, they only enter a correction of order $\mathcal{O}(1/N)$ to the free energy. In the case of fixed boundary conditions, for instance, in the large N limit we have

$$F^{(++)} = -\frac{1}{\beta N} \ln Z_N^{(++)} = -\frac{1}{\beta} \ln \lambda_+ - \frac{1}{\beta N} \ln \cos^2 \phi.$$

The first term is the same for all boundary conditions and coincides with the free energy per unit volume of the system, whereas the second term is associated to the free boundary condition.

2.3 Series Expansions

In this section we discuss another method to compute the partition function of the one-dimensional Ising model. It is worth mentioning that the nature of this method is quite general: it can be applied to higher dimensional lattices and, as a matter of fact, it is presently one of the most powerful approaches to analyze the three-dimensional case. The proposal consists of identifying a perturbative parameter in the high-temperature region and expressing the partition function as a series expansion in this small parameter. In the one-dimensional case the application of this method is particularly simple.

Let us consider once again the partition function in the absence of a magnetic field and, initially, with periodic boundary conditions. It can be written as

$$Z_N(T) = \sum_{\{\sigma\}} e^{-\beta \mathcal{H}} = \sum_{\{\sigma\}} \prod_{i=1}^{N} e^{\mathcal{J}\sigma_i\sigma_{i+1}}. \tag{2.3.1}$$

For any pair of Ising spins, there is the identity

$$e^{\mathcal{J}\sigma_i\sigma_j} = \cosh \mathcal{J} + \sigma_i\sigma_j \sinh \mathcal{J} = \cosh \mathcal{J} \, (1 + \sigma_i\sigma_j \, \tanh \mathcal{J}) \tag{2.3.2}$$

that permits us to express eqn (2.3.1) as

$$Z_N(T) = \cosh^N \mathcal{J} \sum_{\{\sigma\}} \prod_{i=1}^{N} (1 + \sigma_i\sigma_{i+1} \, v), \tag{2.3.3}$$

where $v \equiv \tanh \mathcal{J}$. The parameter v is always less than 1 for all temperatures (except for $T = 0$) and, in particular, it is quite small in the high-temperature phase. Once the product in (2.3.3) is developed, one gets a polynomial of order N in the variable v, whose coefficients are expressed in terms of combinations of the spins σ_i. Consider, for example, a lattice made of three spins. In this case we have

$$\begin{aligned}\prod_{i=1}^{3} (1 + \sigma_i\sigma_{i+1} \, v) &= (1 + v\sigma_1\sigma_2)(1 + v\sigma_2\sigma_3)(1 + v\sigma_3\sigma_1) \\ &= 1 + v(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_1) + v^2(\sigma_1\sigma_2\sigma_2\sigma_3 + \sigma_1\sigma_2\sigma_3\sigma_1 + \sigma_2\sigma_3\sigma_3\sigma_1) \\ &\quad + v^3(\sigma_1\sigma_2\sigma_2\sigma_3\sigma_3\sigma_1).\end{aligned}$$

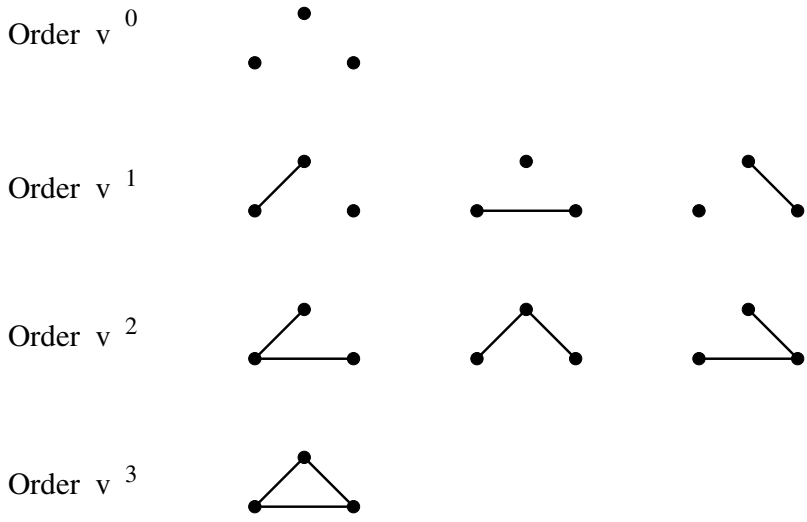


Fig. 2.8 *Graphs relative to a lattice with* 3 *spins.*

We can associate a graph to each of the eight terms of the expression above, by simply drawing a line for each pair of spins entering the product. The whole set of such graphs is shown in Fig. 2.8. Since v appears each time that a term $\sigma_i\sigma_{i+1}$ is involved, it follows that all graphs of order v^l contain exactly l lines. In order to compute the partition function we need, however, to sum over all values ± 1. Thanks to the following properties of the spins of the Ising model

$$\sum_{\sigma_j=-1}^{1} \sigma_j^l = \begin{cases} 2 & \text{if } l \text{ is even} \\ 0 & \text{if } l \text{ is odd} \end{cases}$$

the only non-vanishing contributions come from those graphs where all vertices are of even order (i.e. with an even number of lines). These are the *closed graphs*.

The observation made above is completely general and applies to lattices of arbitrary dimension. In the one-dimensional case, it leads to a particularly simple result: in fact, among the 2^N initial graphs, the only ones that give rise to a non-vanishing result are the graph of order v^0 (i.e. the one without any line) and the graph of order v^N (i.e. the one in which the lines link all sites and give rise to a ring). Hence, in the one-dimensional case of a lattice with N sites and periodic boundary conditions, we have

$$Z_N(T) = \cosh^N \mathcal{J}\,(2^N + 2^N v^N) = 2^N\,(\cosh^N \mathcal{J} + \sinh^N \mathcal{J}) \tag{2.3.4}$$

which coincides with the one obtained by the transfer matrix method, eqn (2.2.15).

It is easy to see the difference between the case in which the chain is closed (periodic boundary conditions) and the case in which the chain is open (free boundary conditions). In the absence of periodic boundary conditions, the only graph that has

all even vertices is the one without lines, i.e. the graph of order v^0. Hence, for the free boundary conditions, this method leads directly to the result that was previously obtained by the recursive method

$$Z_N(T) \,=\, 2^N \cosh^{N-1} \mathcal{J}. \tag{2.3.5}$$

For an arbitrary lattice in which the interaction is restricted to the next neighbor spins, the series expansion approach permits us to express the partition function in the following form

$$Z_N(T) \,=\, 2^N \,(\cosh \mathcal{J})^P \,\sum_{l=0}^{P} h(l)\, v^l, \tag{2.3.6}$$

where P is the total number of segments of the lattice and $h(l)$ is the number of graphs that can be drawn on it by using l lines, with the condition that each vertex is of even order. Hence, in the series expansion approach, the solution of the Ising model on an arbitrary lattice reduces to solving the geometrical problem of the counting of the close graphs on the lattice under investigation.

2.4 Critical Exponents and Scaling Laws

The one-dimensional Ising model does not have a phase transition at a finite value of the temperature. However, the point $B = T = 0$ of the phase diagram can be considered a critical point of the system, for the correlation length ξ diverges in correspondence with these values. This leads to a definition of critical exponents that verify the scaling relations (1.1.26). In Chapter 1, we adopted the variables $t = (T - T_c)/T_c$ and B in order to characterize the displacement from the critical point. In this case, in view of the condition $T_c = 0$, it is more convenient to use the variables $\mathcal{B} = B/kT$ and

$$t \,=\, \exp(-2\mathcal{J}) = \exp(-2J/kT). \tag{2.4.1}$$

Looking at the divergence ξ with respect to the new variable t, $\xi \sim (2t)^{-1}$, we have

$$\nu \,=\, 1.$$

Analogously, the divergence of the magnetic susceptibility, given by $\chi \sim t^{-1}$, fixes the value of the critical exponent γ

$$\gamma \,=\, 1.$$

At the critical point the correlation function of the spins is constant, hence

$$\eta \,=\, 1.$$

Since the spontaneous magnetization always vanishes for $B = 0$, the exponent β is identically null:

$$\beta \,=\, 0.$$

In an external magnetic field, the magnetization at $T = 0$ is a discontinous function and therefore the critical exponent δ is infinite:

$$\delta = \infty.$$

Finally, in the vicinity of the critical point the singular part of the free energy can be written as

$$F_{sing} \sim t\sqrt{1 + \frac{\mathcal{B}^2}{t^2}}.$$

Comparing with the scaling law (1.1.30) of the free energy, we obtain the two relations

$$\alpha = 1, \quad \beta\delta = 1.$$

It is an easy exercise to check that the critical exponents derived above satisfy the scaling laws (1.1.26).

2.5 The Potts Model

The Ising model can be generalized in several ways. One possibility is provided by the Potts model. It consists of a statistical model in which, at each site of a lattice, there is a variable σ_i that takes q discrete values, $\sigma_i = 1, 2, \ldots, q$. In this model, two adjacent spins have an interaction energy given by $-J\,\delta(\sigma_i.\sigma_j)$, where

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

and the hamiltonian reads

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \delta(\sigma_i, \sigma_j). \tag{2.5.1}$$

This expression is invariant under the group $\mathcal{S}_q$ of the permutations of q objects. This is a non-abelian group if $q \geq 3$. For the type of interaction, it is clear that the nature of the values taken by the spins is completely inessential: instead of the q values listed above, one can consider other q distinct numbers or variables of other nature. One can conceive, for instance, that the q values stand for q different colors. When $q = 2$, as the two distinct values we can take ± 1: thanks to the identity $\delta(\sigma, \sigma') = \frac{1}{2}(1 + \sigma\sigma')$, making the change $J \to 2J$, the Potts model is equivalent to the original Ising model.

The partition function of the Potts model defined on a lattice of N sites is expressed by a sum of q^N terms ($\mathcal{J} = \beta J$)

$$Z_N = \sum_{\{\sigma\}} \exp\left[\mathcal{J} \sum_{\langle ij \rangle} \delta(\sigma_i, \sigma_j)\right]. \tag{2.5.2}$$

In the one-dimensional case, it can be exactly computed by using either the recursive method or the transfer matrix approach.

Recursive method. Consider a chain of N spins with free boundary conditions at the last spins of the chain. Adding an extra spin, the partition function becomes

$$Z_{N+1} \;=\; \left(\sum_{\sigma_{N+1}=1}^{q} e^{\mathcal{J}\,\delta(\sigma_N,\sigma_{N+1})} \right) Z_N. \tag{2.5.3}$$

Making use of the identity

$$e^{x\delta(a,b)} \;=\; 1 + (e^x - 1)\,\delta(a,b), \tag{2.5.4}$$

the sum in (2.5.3) can be expressed as

$$\begin{aligned} \sum_{\sigma_{N+1}=1}^{q} e^{[\mathcal{J}\,\delta(\sigma_N,\sigma_{N+1})]} &= \sum_{\sigma_{N+1}=1}^{q} \left[1 + (e^{\mathcal{J}} - 1)\,\delta(\sigma_N,\sigma_{N+1})\right] \\ &= q + (e^{\mathcal{J}} - 1). \end{aligned}$$

The recursive equation is expressed by

$$Z_{N+1} \;=\; \left(q - 1 + e^{\mathcal{J}}\right)\, Z_N.$$

Since $Z_1 = q$, the iteration of the formula leads to the exact result

$$Z_N \;=\; q\,\left(q - 1 + e^{\mathcal{J}}\right)^{N-1}. \tag{2.5.5}$$

In the thermodynamic limit, the free energy per unit of spin is given by

$$F(T) \;=\; -\lim_{N\to\infty} \frac{1}{\beta N} \ln Z_N \;=\; -\frac{1}{\beta} \ln\left(e^{\mathcal{J}} + q - 1\right). \tag{2.5.6}$$

Transfer matrix. Equally instructive is the computation of the partition function done with the transfer matrix method. For simplicity, let's assume periodic boundary conditions, i.e. $\sigma_{N+1} \equiv \sigma_1$. In the transfer matrix formalism, the spins are associated to a vector of a q-dimensional Hilbert space, with the completeness relation given by

$$\sum_{\sigma=1}^{q} |\,\sigma\rangle\,\langle\sigma\,| = 1.$$

Analogously to the Ising model, the partition function can be expressed as

$$Z_N \;=\; \mathrm{Tr}\; V^N, \tag{2.5.7}$$

where the transfer matrix V is a $q \times q$ matrix, whose elements are

$$\langle\sigma \mid V \mid \sigma'\rangle \;=\; \exp\left[\mathcal{J}\delta(\sigma,\sigma')\right]. \tag{2.5.8}$$

Hence, V has diagonal elements equal to $e^{\mathcal{J}}$ whereas all the other off-diagonal elements are equal to 1:

$$V = \begin{pmatrix} e^{\mathcal{J}} & 1 & 1 & \cdots & 1 & 1 \\ 1 & e^{\mathcal{J}} & 1 & \cdots & 1 & 1 \\ 1 & 1 & e^{\mathcal{J}} & \cdots & 1 & 1 \\ \cdots & \cdots & \cdots & e^{\mathcal{J}} & \cdots & 1 \\ 1 & 1 & \cdots & \cdots & e^{\mathcal{J}} & 1 \\ 1 & 1 & \cdots & \cdots & 1 & e^{\mathcal{J}} \end{pmatrix}. \tag{2.5.9}$$

To compute the trace of V^N it is useful to determine the eigenvalues of V, which are solutions of the equation

$$\mathcal{D} = || \, V - \lambda \mathbf{1} \, || = 0. \tag{2.5.10}$$

Denote $x \equiv e^{\mathcal{J}} - \lambda$. The determinant (2.5.10) can be computed by using the well-known property that a determinant does not change by summing or subtracting rows and columns. Subtracting the second column from the first one, the third column from the second one, and so on, we have

$$\mathcal{D} = \begin{Vmatrix} x-1 & 0 & 0 & \cdots & 0 & 1 \\ 1-x & x-1 & 0 & \cdots & 0 & 1 \\ 0 & 1-x & x-1 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & x-1 & 0 & 1 \\ 0 & 0 & \cdots & \cdots & x-1 & 1 \\ 0 & 0 & \cdots & \cdots & 1-x & x \end{Vmatrix}.$$

Summing the first row and the second one, we get

$$\mathcal{D} = \begin{Vmatrix} x-1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & x-1 & 0 & \cdots & 0 & 2 \\ 0 & 1-x & x-1 & 0 & \cdots & 1 \\ 0 & \cdots & \cdots & x-1 & \cdots & 1 \\ 0 & 0 & \cdots & \cdots & x-1 & 1 \\ 0 & 0 & \cdots & \cdots & 1-x & x \end{Vmatrix}.$$

If we now sum the second row and the third one, the third row and the fourth one, and so on, we have the final expression

$$\mathcal{D} = \begin{Vmatrix} x-1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & x-1 & 0 & \cdots & 0 & 2 \\ 0 & 0 & x-1 & 0 & \cdots & 3 \\ 0 & \cdots & 0 & x-1 & \cdots & 4 \\ 0 & 0 & \cdots & \cdots & x-1 & q-1 \\ 0 & 0 & 0 & \cdots & 0 & x+q-1 \end{Vmatrix}.$$

So the determinant of the secular equation is given by

$$|| \, V - \lambda \mathbf{1} \, || = \left(e^{\mathcal{J}} - 1 - \lambda\right)^{q-1} \left(e^{\mathcal{J}} - q + 1 - \lambda\right) = 0, \tag{2.5.11}$$

and the roots are expressed by

$$\lambda_+ \,=\, e^{\mathcal{J}} + q - 1, \qquad \lambda_- \,=\, e^{\mathcal{J}} - 1. \tag{2.5.12}$$

For $q \geq 0$, we have $\lambda_+ \geq \lambda_-$. The eigenvalue λ_+ is not degenerate, while λ_- is $(q-1)$ times degenerate. The physical origin of this degeneration is obvious, since the interaction of the Potts model only distinguishes if two sites are in the same state or not: there is only one way in which they can be equal but $(q-1)$ ways in which they can be different. Once the eigenvalues of V are known, the partition function (2.5.7) can be expressed as

$$Z_N \,=\, \mathrm{Tr}\, V^N \,=\, \lambda_+^N + (q-1)\,\lambda_-^N. \tag{2.5.13}$$

In the thermodynamic limit, the free energy per unit spin depends only on the largest eigenvalue λ_+:

$$\begin{aligned} F(T) &= -\lim_{N\to\infty} \frac{1}{\beta N} \ln Z_N \,=\, -\lim_{N\to\infty} \frac{1}{\beta N} \left[N \ln \lambda_+ + \ln \left[1 + (q-1) \left(\frac{\lambda_-}{\lambda_+} \right)^N \right] \right] \\ &= -\frac{1}{\beta} \ln \left(e^{\mathcal{J}} + q - 1 \right). \end{aligned} \tag{2.5.14}$$

This result coincides with (2.5.6).

Series expansion. Let us now consider the solution of the Potts model obtained in terms of the high-temperature series expansion. Since this method points out some interesting geometrical properties, it is convenient to study the general case of a Potts model defined on an arbitrary lattice $\mathcal{L}$ as, for instance, the one shown in Fig. 2.9. Putting

$$v \,\equiv\, e^{\mathcal{J}} - 1,$$

and using the identity (2.5.4), the partition function (2.5.2) can be written as

$$Z_N \,=\, \sum_{\{\sigma\}} \prod_{\langle ij \rangle} \left[1 + v\, \delta(\sigma_i, \sigma_j) \right]. \tag{2.5.15}$$

Note that v is a small parameter when the temperature T is very high. Let E be the total number of links of the graph $\mathcal{L}$. Inside the sum (2.5.15) there is a product of E factors, each of them being either 1 or $v\,\delta(\sigma_i, \sigma_j)$. Expanding the product above, there are 2^E terms: their graphical representation is obtained by drawing a line on the link between the sites i and j when the factor $v\,\delta(\sigma_i, \sigma_j)$ is present. In such a way, there is a one-to-one correspondance between the terms in (2.5.15) and the graphs that can be drawn on the lattice $\mathcal{L}$. Let us now consider one of these graphs $\mathcal{G}$, made of l links and C connected components (an isolated site is considered as a single component). The corresponding term in Z_N contains a factor v^l and, thanks to the factor $\delta(\sigma, \sigma')$ that accompanies v, all spins that belong to the same component have the same value. Summing over all possible values of σ_i, the contribution of this graph to the partition

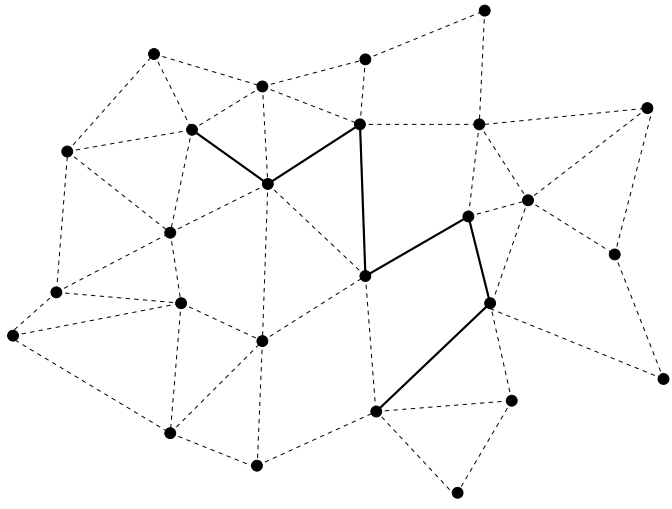


Fig. 2.9 *Lattice $\mathcal{L}$ and graph $\mathcal{G}$.*

function amounts to $q^C\, v^l$. Considering all graphs $\mathcal{G}$ of the lattice $\mathcal{L}$, the partition function can thus be expressed in terms of a sum over graphs:

$$Z_N \;=\; \sum_{\mathcal{G}} q^C\, v^l. \tag{2.5.16}$$

For the analytic form of this expression, q does not necessarily have to be an integer and therefore this formula can be used to define the Potts model for arbitrary values of q. This observation is useful, for instance, in the study of percolation[1] (associated to the limit $q \to 1$ of the Potts model) or in the analysis of the effective resistance between two nodes of an electric circuit made of linear resistances (expressed in terms of the limit $q \to 0$ of the model).

Chromatic polynomial. It is interesting to study the Potts model in the limit $\mathcal{J} \to -\infty$, i.e. when the temperature T goes to zero and the model is antiferromagnetic. In such a limit, neighbor sites should necessarily take different values in order to contribute to the partition function Z_N: hence this quantity provides in this case the number of ways in which it is possible to color the sites of $\mathcal{L}$ with q colors, with the constraint that two neighbor sites do not have the same color. The expression obtained by substituting $v = -1$ in Z_N is a polynomial $\mathcal{P}_N(q)$ in the variable in q, called the *chromatic polynomial* of the graph $\mathcal{L}$.

In the one-dimensional case, taking the limit $\mathcal{J} \to -\infty$ in the partition function (2.5.5) associated to the free boundary condition of the chain, we get

$$\mathcal{P}_N^a(q) \;=\; q\,(q-1)^{N-1}. \tag{2.5.17}$$

[1]For the elaboration of this topic, see the suggested texts at the end of the chapter.

The zeros $q = 0$ and $q = 1$ of this polynomial clearly show that, if we wish to distinguish neighbor sites by means of different colors, it is impossible to color a one-dimensional lattice by having only one color or none. The combinatoric origin of (2.5.17) is simple: in fact, the first site can be colored in q different ways but, once a color is chosen, the next site can be distinguished by employing one of the $(q-1)$ remaining colors, and this argument repeats for the other sites.

For periodic boundary conditions, taking the limit $\mathcal{J} \to -\infty$ in the corresponding expression (2.5.13) of the partition function, we have

$$\mathcal{P}_N^c(q) \;=\; (q-1)^N + (-1)^N\,(q-1) = (q-1)\,\left[(q-1)^{N-1} + (-1)^N\right]. \qquad (2.5.18)$$

Although this expression differs from (2.5.17), it is easy to see that it has the same real roots $q = 0$ and $q = 1$. It is an exercise left to the reader to derive it by using a combinatoric argument.

For planar two-dimensional lattices, the limit $\mathcal{J} \to -\infty$ of the Potts model is deeply related to a famous problem of topology, i.e. the *four-color problem*. It consists of proving the conjecture that any geographical planar map, in which different neighbor nations are distinguished by different colors, can be drawn using only four colors. If one assumes the validity of this result, the conclusion is that the partition function of the Potts model for any planar graph, in the limit $\mathcal{J} \to -\infty$, does not ever have $q = 4$ among the set of its zeros. A brief discussion of the four-color problem is reported in Appendix 2C.

2.6 Models with $O(n)$ Symmetry

Another interesting generalization of the Ising model is provided by the $O(n)$ model, in which each spin $\vec{S}_i$ is a n-component vector associated to a point of the n-dimensional sphere

$$|\,\vec{S}_i\,|^2 = \sum_{k=1}^{n} (S_i)_k^2 = 1.$$

In the one-dimensional case, the hamiltonian of the model is given by

$$\mathcal{H} \;=\; -\sum_{i=1}^{N-1} J_i\,\vec{S}_i \cdot \vec{S}_{i+1}, \qquad (2.6.1)$$

and this expression is clearly invariant under the rotations of the vectors $\vec{S}_i$ associated to the $O(n)$ group. In this formulation, the Ising model is obtained in the limit $n \to 1$. The sum the configurations of the $O(n)$ model consists of the integrals of the solid angles of the n-dimensional spins

$$Z_N(T) \;=\; \int d\Omega_1^{(n)} \int d\Omega_2^{(n)} \cdots \int d\Omega_N^{(n)}\; \exp\left[\sum_{i=1}^{N-1} \mathcal{J}_i\,\vec{S}_i \cdot \vec{S}_{i+1}\right], \qquad (2.6.2)$$

where

$$d\Omega^{(n)} = \sin^{n-2}\theta_{n-1} d\theta_{n-1} \sin^{n-3}\theta_{n-2} d\theta_{n-2} \cdots d\theta_1,$$
$$0 \le \theta_1 \le 2\pi,$$
$$0 \le \theta_k \le \pi.$$

The solid angle is given by

$$\Omega(n) = \int d\Omega^{(n)} = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}, \tag{2.6.3}$$

where $\Gamma(x)$ is the function that generalizes the factorial to arbitrary real and complex numbers.[2]

To prove (2.6.3), let's consider the well-known identity of the gaussian integral

$$I = \int_{-\infty}^{+\infty} dx\, e^{-x^2} = \sqrt{\pi}.$$

By taking the product of n such integrals, we have ($r^2 = x_1^2 + x_2^2 + \cdots + x_n^2$)

$$I^n = \left[\int_{-\infty}^{+\infty} dx\, e^{-x^2}\right]^n = \int d^n x\, e^{-r^2} = \frac{\Omega(n)}{2} \int_0^\infty dt\, t^{\frac{n}{2}-1} e^{-t} = \frac{\Omega(n)}{2}\, \Gamma\left(\frac{n}{2}\right).$$

On the other hand, $I^n = \pi^{n/2}$ and therefore we arrive at (2.6.3).
Using

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi},$$

it is easy to check that we obtain the known values of planar and three-dimensional solid angle when $n = 2$ and $n = 3$. For $n = 1$ it correctly reproduces the sum of the states of the Ising model, i.e. $\Omega(1) = 2$, since a *one-dimensional sphere* consists of two points. Other interesting properties of the n-dimensional solid angle are discussed in Appendix 2B.

To compute (2.6.2) we can use the recursive method. Let's add an extra spin to the system, so that

$$Z_{N+1}(T) = \left(\int d\Omega_{N+1}^{(n)} \exp\left[\mathcal{J}_N \vec{S}_N \cdot \vec{S}_{N+1}\right]\right) Z_N(T).$$

Since the n-th axis can always be chosen along the direction of the spin $\vec{S}_n$, we have $\vec{S}_N \cdot \vec{S}_{N+1} = \cos\theta_{n-1}$. Integrating over the remaining angles $\theta_1, \theta_2, \ldots, \theta_{n-2}$ we get

$$Z_{N+1}(T) = \left(\Omega(n-1) \int_0^\pi d\theta_{n-1} \sin^{n-2}\theta_{n-1} e^{\mathcal{J}_N \cos\theta_{n-1}}\right) Z_N(T). \tag{2.6.4}$$

[2]The properties of the function $\Gamma(x)$ and the Bessel functions $I_\nu(x)$ that enter the discussion of this model are reported in Appendix 2A.

Although this is not an elementary integral, it can nevertheless be expressed as a closed formula in terms of the $\Gamma(x)$ function and the Bessel functions $I_\nu(z)$

$$\int_0^\pi d\theta_{n-1}\, \sin^{n-2}\theta_{n-1} e^{\mathcal{J}_N \cos\theta_{n-1}} \;=\; \frac{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}{\left(\frac{\mathcal{J}_N}{2}\right)^{\frac{n-2}{2}}}\, I_{\frac{n-2}{2}}(\mathcal{J}_N).$$

Substituting $\Omega(n-1)$ in (2.6.4) and simplifying the resulting expression, we get

$$\int d\Omega_{N+1}^{(n)} \exp\left[\mathcal{J}_N \vec{S}_N \cdot \vec{S}_{N+1}\right] \;=\; \left[(2\pi)^{n/2} \frac{I_{\frac{n-2}{2}}(\mathcal{J}_N)}{\mathcal{J}_N^{\frac{n-2}{2}}}\right] \equiv \lambda_1(\mathcal{J}_N). \tag{2.6.5}$$

The recursive equation is then given by

$$Z_{N+1} \;=\; \lambda_1(\mathcal{J}_N)\; Z_N.$$

Let us consider, for simplicity, the case of equal couplings. By iterating (2.6), we obtain

$$Z_N(T) \;=\; [\lambda_1(\mathcal{J})]^{N-1}\; Z_1,$$

where Z_1 is the partition function of a single spin. This is simply expressed by the phase space of the configuration of a single spin, i.e. by the n-dimensional solid angle (2.6.3), so that the final expression is

$$Z_N(T) \;=\; \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}\, [\lambda_1(\mathcal{J})]^{N-1} \;=\; \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)} \left[(2\pi)^{n/2}\, \frac{I_{\frac{n-2}{2}}(\mathcal{J})}{\mathcal{J}^{\frac{n-2}{2}}}\right]^{N-1}. \tag{2.6.6}$$

The free energy, per unit spin, of the $O(n)$ model is

$$\beta F(\beta) \;=\; -\frac{1}{N}\, \log Z_N \;=\; -\frac{N-1}{N}\, \log\left[\frac{I_{\frac{n-2}{2}}(\mathcal{J})}{\mathcal{J}^{\frac{n-2}{2}}}\right] - \frac{1}{N} \log \Omega(n),$$

and in the thermodynamic limit $N \to \infty$

$$\beta F(\beta) \;=\; -\log\left[\frac{I_{\frac{n-2}{2}}(\mathcal{J})}{\mathcal{J}^{\frac{n-2}{2}}}\right]. \tag{2.6.7}$$

As for the Ising model, also for the $O(n)$ model it is possible to obtain the exact expression of the two-point correlation function (see Fig. 2.10)

$$\langle \vec{S}_i \cdot \vec{S}_{i+r} \rangle \;=\; \left[\frac{I_{\frac{n}{2}}(\mathcal{J})}{I_{\frac{n-2}{2}}(\mathcal{J})}\right]^r. \tag{2.6.8}$$

Expressed as

$$\langle \vec{S}_i \cdot \vec{S}_{i+r} \rangle \;\equiv\; e^{-r/\xi},$$

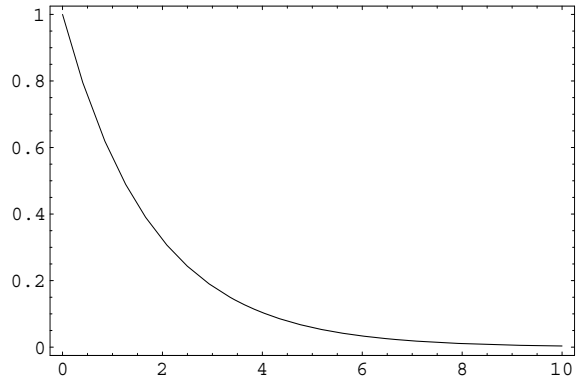


Fig. 2.10 *Typical behavior of the two-point correlation function of the spins, as a function of the distance r between the spin, for $n \geq 1$.*

we can determine the correlation length of the model that, in units of the lattice space a, is given by

$$\xi(\mathcal{J}) \;=\; -\frac{1}{\log\left[\frac{I_{\frac{n}{2}}(\mathcal{J})}{I_{\frac{n-2}{2}}(\mathcal{J})}\right]}. \tag{2.6.9}$$

The proof of (2.6.8) comes from the following identity of the Bessel functions

$$\frac{d}{dx}\left[x^{-\mu}\, I_\mu(x)\right] \;=\; x^{-\mu}\, I_{\mu+1}(x).$$

Taking the derivative with respect to $\mathcal{J}$ of eqn (2.6.5), this identity permits us to compute the integral

$$\int d\Omega^{(n)}\, \vec{S}\, \exp[\mathcal{J}\vec{S}\cdot\vec{S}'] \;=\; \left[(2\pi)^{n/2}\frac{I_{\frac{n}{2}}(\mathcal{J})}{\mathcal{J}^{\frac{n-2}{2}}}\right] \vec{S}' \;\equiv \lambda_2(\mathcal{J})\,\vec{S}'.$$

We then have

$$\int d\Omega_1^{(n)} \int d\Omega_2^{(n)} \cdots \int d\Omega_N^{(n)}\; \vec{S}_i\cdot\vec{S}_{i+r}\; \exp\left[\sum_{i=1}^{N-1} \mathcal{J}\,\vec{S}_i\cdot\vec{S}_{i+1}\right]$$
$$= [\lambda_1(\mathcal{J})]^{i-1}\; [\lambda_2(\mathcal{J})]^r\; [\lambda_1(\mathcal{J})]^{N-i}\; \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}.$$

Dividing this expression by the partition function Z_N, given by (2.6.6), we arrive at the final result (2.6.8) of the correlators.

It is interesting to observe that all expressions considered so far are *analytic functions* of the parameter n and, for that reason, they can be used to study the behavior of the $O(n)$ model for arbitrary values of n, not necessarily integers. This is a useful observation: it extends to higher dimensions and permits us to study, for instance, the statistical properties of polymers,[3] whose dilute phase is described by the limit $n \to 0$.

It is important to underline that in the range $n < 1$ there could be surprising behaviors that need further considerations for their correct physical interpretation. The following analysis aims to study the nature of the model by varying the parameter n. It is convenient to define the quantity

$$\Lambda(\mathcal{J}) \equiv \frac{\lambda_2}{\lambda_1} = \left[\frac{I_{\frac{n}{2}}(\mathcal{J})}{I_{\frac{n-2}{2}}(\mathcal{J})}\right],$$

and to distinguish the cases: (i) $n \geq 1$; (ii) $0 \leq n \leq 1$; and (iii) $n \leq 0$.

- In the first interval, $n \geq 1$, using eqns (2.A.9) and (2.A.13) given in Appendix 2A, it is easy to check that for all values of $\mathcal{J}$, i.e. of the temperature, we have

  $$\lambda_1(\mathcal{J}) > 0, \qquad \Lambda(\mathcal{J}) < 1.$$

  The first condition, as can be seen in (2.6.6), implies that the partition function of the model is a positive quantity and, consequently, that the free energy is a real function. The second condition, using eqn (2.6.8), implies that the correlator has the usual behavior of a decreasing exponential, as a function of the distance r between the spins.

  Both results agree with what is expected on the basis of physical considerations. When $n = 1$, using the identity

  $$\frac{I_{\frac{1}{2}}(\mathcal{J})}{\mathcal{J}^{\frac{1}{2}}} = \sqrt{\frac{2}{\pi}}\cosh\mathcal{J},$$

  one recovers the previous expressions of the partition function and correlator of the one-dimensional Ising model.

  To study the limit $n \to \infty$, we need to use the asymptotic expressions of the Bessel functions

  $$I_\nu(\nu x) \simeq \frac{1}{\sqrt{2\pi\nu}} \frac{e^{\nu(\sqrt{1+x^2}-\xi^{-1}}}{(1+x^2)^{1/4}}, \qquad \nu \to \infty,$$

  with

  $$\xi^{-1} = \ln\left(\frac{1+\sqrt{1+x^2}}{x}\right).$$

  When $n \to \infty$ an interesting result is obtained by taking, simultaneously, the limit $\mathcal{J} \to \infty$. It is convenient to introduce $x \equiv 2\mathcal{J}/(n-2)$ and express all

[3]The relation between the $O(n)$ model and the statistics of polymers is due to De Gennes. Those who are interested in further development of this issue can consult the bibliographic references given at the end of the chapter.

thermodynamic quantities in terms of this variable. Consider, for instance, the ratio of the two eigenvalues λ_2 and λ_1 in this double limit:

$$\frac{\lambda_2(x)}{\lambda_1(x)} = e^{-\xi^{-1}}.$$

This allows us to identify the parameter ξ with the correlation length of the model. This quantity diverges for $T \to 0$, whereas it vanishes for $T \to \infty$. The last limit corresponds to the full disordered state of the system, where each spin is independent and completely uncorrelated with the others. The internal energy is given by

$$U = -\frac{\partial}{\partial x} \ln \lambda_1(x),$$

and, using the asymptotic expression of the Bessel functions, it can be expressed as

$$\frac{U(x)}{n} = \frac{1 - \sqrt{1 + x^2}}{x}. \tag{2.6.10}$$

This formula shows that the internal energy, relative to each component of the spin, remains finite in the double limit $n \to \infty$, $\mathcal{J} \to \infty$, with x finite.

- In the second interval, $0 \leq n < 1$, using (2.A.9) and (2.A.13), it is easy to see that, for all values of $\mathcal{J}$, we have

$$\lambda_1(\mathcal{J}) > 0.$$

However, the inequality

$$\Lambda(\mathcal{J}) < 1,$$

is not always true: in this interval of values of n, it is always possible to find a value $\mathcal{J}_c$ such that, for $\mathcal{J} > \mathcal{J}_c$, we have $\Lambda(\mathcal{J}) > 1$, as shown in Fig. 2.11.

From eqn (2.6.9), the correlation length $\xi(\mathcal{J})$ is positive for $\mathcal{J} < \mathcal{J}_c$ while, for $\mathcal{J} > \mathcal{J}_c$, it becomes negative! Moreover, it diverges at $\mathcal{J}_c$, as shown in Fig. 2.12. The critical value $\mathcal{J}_c$ moves toward the origin by decreasing n and, when $n \to 0$, we have $\mathcal{J}_c = 0$. In such a limit, taking into account the factor

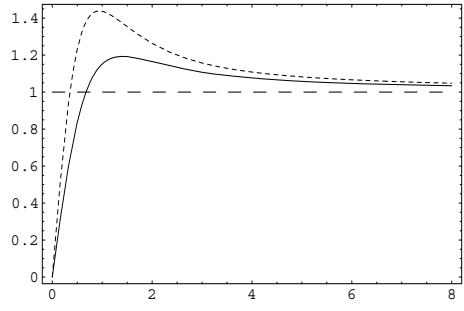


Fig. 2.11 Λ *as a function of* $\mathcal{J}$*. The dashed line corresponds to* $n = 0.3$*, the other curve to* $n = 0.6$*.*

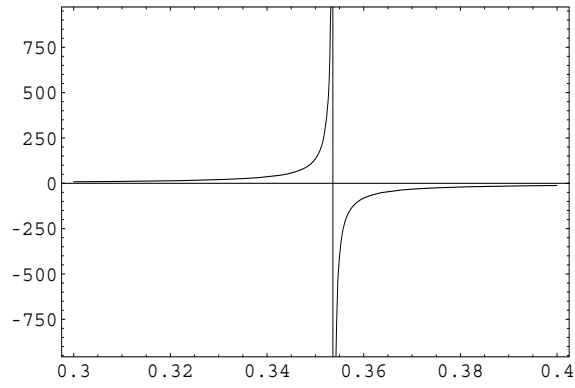


Fig. 2.12 *Plot of the correlation length in the vicinity of $\mathcal{J} = \mathcal{J}_c$.*

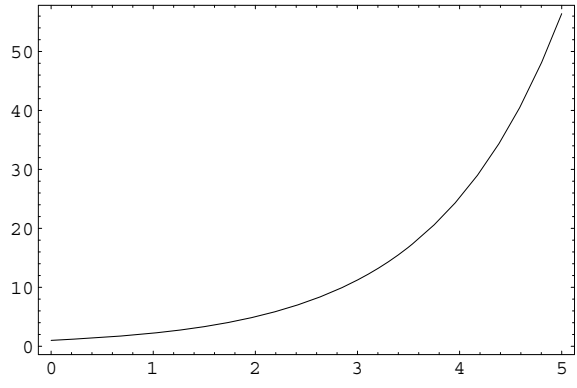


Fig. 2.13 *Correlation function $\langle \vec{S}_i \cdot \vec{S}_{i+r} \rangle$ as a function of the separation r, for $n = 0$.*

$\Gamma\left(\frac{n}{2}\right)$ in the denominator of (2.6.6), the partition function vanishes linearly[4] in the variable n, while the correlation function is finite and takes the form

$$\lim_{n\to 0} \langle \vec{S}_i \cdot \vec{S}_{i+r} \rangle \;=\; \left[\frac{I_0(\mathcal{J})}{I_{-1}(\mathcal{J})} \right]^r . \tag{2.6.11}$$

Note that, for all the values of temperature, this is an exponential increasing function of the distance of the spins! Namely, increasing the separation between the spins, their correlation increases exponentially, instead of decreasing – a behavior that is quite anti-intuitive from a physical point of view (see Fig. 2.13).

- Let us consider the last interval, $n < 0$. Using eqns (2.A.9) and (2.A.13), the Bessel function $I_{\frac{n-2}{2}}(\mathcal{J})$ is always positive (as a function of $\mathcal{J}$), in the following ranges of n

$$-4k \;<\; n \;<\; -4k+2, \quad k = 1, 2, 3, \ldots \tag{2.6.12}$$

[4]This implies that there exists the finite limit $\lim_{n\to 0} \frac{\partial Z}{\partial n}$.

Fig. 2.14 *The continuous intervals and the points identified by the circles are those in which the free energy is real.*

In the other intervals

$$-4k-2 < n < -4k, \quad k = 0, 1, 2, 3, \ldots \tag{2.6.13}$$

there are instead values of $\mathcal{J}$ where $I_{\frac{n-2}{2}}(\mathcal{J})$ assumes negative values. Correspondingly, the free energy per unit of spin, given by (2.6.7), is a real function of $\mathcal{J}$ only in the intervals (2.6.12), whereas in the other intervals it develops an imaginary part that signals the thermodynamic instability of the system. Finally, for $n = -2k$, where $k = 0, 1, 2, \ldots$, $I_{\frac{n-2}{2}}(\mathcal{J})$ is always positive and therefore the free energy is real for those values. The behavior of the free energy of the model is given in Fig. 2.14.

Let's now analyze the ratio $\Lambda(\mathcal{J})$, by starting with the study of the positivity of such a quantity. This is determined by the positivity of the functions $I_{\frac{n-2}{2}}(\mathcal{J})$ and $I_{\frac{n}{2}}(\mathcal{J})$. The investigation of the first function coincides with what has been done previously with the free energy. Concerning the second function, this is positive for all values of $\mathcal{J}$ in the intervals

$$-4k-2 < n < -4k, \quad k = 1, 2, 3, \ldots \tag{2.6.14}$$

In the other intervals of n, there are instead values of $\mathcal{J}$ where this function takes negative values. In conclusion, there is no interval of n where the two functions are both positive. This implies that, for any negative n with $n \neq -2k$ ($k = 0, 1, 2, \ldots$), there is always a value $\mathcal{J}_c$ in which the correlation length diverges, assuming complex values in an interval $\mathcal{J} < \mathcal{J}_c$ near the origin. For $n = -2k$, instead, $\Lambda(\mathcal{J})$ is real but larger that 1, so that the correlation length is *negative* for all values of the temperature: the correlation function of the spin thus increases by increasing their separation.

The above analysis aimed to show the possibility of studing the behavior of the model by varying continuously the number n of the components of the vector $\vec{S}_i$. From this point of view, the one-dimensional $O(n)$ model is a paradigm of an important class of models that we will meet again in the following chapters and that will allow us to make progress in important fields of theoretical physics.

2.7 Models with Z_n Symmetry

Beside the generalizations of the Ising models given by the Potts and the $O(n)$ models, there is another possible extension provided by the Z_n models. In this case, the spins are planar vectors of unit length, which can be identified by their discrete angles θ_i with respect to the horizontal axes

$$\alpha^{(k)} = \frac{2\pi k}{n}, \quad k = 0, 1, 2, \ldots, n-1. \tag{2.7.1}$$

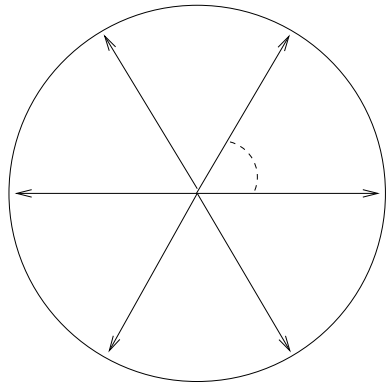


Fig. 2.15 *Possible values of the spins in the Z_6 model.*

They can be associated to the n (complex) roots of unity as in Fig. 2.15. The hamiltonian of the Z_n model is defined by

$$\mathcal{H} \,=\, -J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j \,=\, -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j), \tag{2.7.2}$$

and is invariant under the abelian group Z_n generated by the discrete rotations of the angles θ_i. In terms of the index k defined in (2.7.1), this symmetry is implemented by the transformations

$$k \,\rightarrow\, k + m \pmod n, \quad m = 0, 1, \ldots n. \tag{2.7.3}$$

For some particular values of n, the Z_n models coincide with previously defined models. For instance, when $n = 2$, one recovers the familiar Ising model or, equivalently, the two-state Potts model. When $n = 3$, the Z_3 model is equivalent to the three-state Potts model: it is sufficient to put $J = \frac{2}{3} J_{Potts}$ to have the coincidence of the two hamiltonians. Finally, when $n \rightarrow \infty$ the Z_n model becomes equivalent to the $O(2)$ model, i.e. that model invariant under an arbitrary rotation of the spins.

In the one-dimensional case, the solution of the Z_n model can be achieved by using the recursive method. Let us consider firstly the partition function of N spins

$$Z_N \,=\, \sum_{\theta_1=0}^{n-1} \cdots \sum_{\theta_N=0}^{n-1} \exp\left[\mathcal{J} \sum_{i=0}^{N-1} \cos\left(\frac{2\pi}{n}(\theta_i - \theta_{i+1}\right)\right]. \tag{2.7.4}$$

For $N = 1$, Z_1 is equal to the number of possible states of the system, i.e. $Z_1 = n$. Adding a new spin to the chain, one has

$$Z_{N+1} \,=\, Z_N \sum_{\theta_{N+1}=0}^{n-1} \exp\left[\mathcal{J} \cos\left(\frac{2\pi}{n}(\theta_N - \theta_{N+1}\right)\right],$$

where the last sum is independent of θ_N. Indeed, whatever the value taken by this variable, the sum over the angle θ_{N+1} in the argument $(\theta_N - \theta_{N+1})$ implies that this

quantity spans all possible values (2.7.1), i.e. θ_N can be eliminated by a simple change of variable. Hence, the partition function satisfies the recursive equation

$$Z_{N+1} \;=\; \mu_1(\mathcal{J}, n)\; Z_N, \tag{2.7.5}$$

where we have defined

$$\mu_1(\mathcal{J}, n) \;\equiv\; \sum_{k=0}^{n-1} \exp\left[\mathcal{J}\cos\frac{2\pi k}{n}\right]. \tag{2.7.6}$$

By iterating (2.7.5), with the initial condition $Z_1 = n$, we get

$$Z_N \;=\; n\;[\mu_1(\mathcal{J}, n)]^{N-1}. \tag{2.7.7}$$

It is easy to compute the correlation function of two spins

$$G(r) \;=\; \langle \vec{S}_i \cdot \vec{S}_{i+r}\rangle \;=\; \langle \cos(\theta_i - \theta_{i+r})\rangle.$$

For this, one needs the identity

$$\sum_{\{\vec{S}\}} \vec{S}\, e^{\mathcal{J}\vec{S}\cdot\vec{S}'} \;=\; \mu_2(\mathcal{J}, n)\,\vec{S}',$$

where the sum is over all discrete values of the vector $\vec{S}$ of the $Z(n)$ model and

$$\mu_2(\mathcal{J}, n) \;=\; \frac{\partial}{\partial \mathcal{J}}\mu_1(\mathcal{J}, n).$$

Following the same steps as the Ising and the $O(n)$ models, one has

$$G(r) \;=\; \left(\frac{\mu_2}{\mu_1}\right)^r. \tag{2.7.8}$$

When $n = 2$, both the partition function and the correlator coincide with those of the Ising model. When $n \to \infty$, a finite result is obtained by properly rescaling the sum over the states, i.e. multiplying the sum by $2\pi/n$ and then taking the limit. In this way, the previous formula becomes

$$\lim_{n\to\infty} \frac{2\pi}{n} \sum_{k=0}^{\infty}[\,\ldots\,] \;\longrightarrow\; \int_0^{2\pi} d\alpha\,[\,\ldots\,].$$

Hence

$$\lim_{n\to\infty} \mu_1(\mathcal{J}, n) \;=\; 2\pi I_0(\mathcal{J}), \qquad \lim_{n\to\infty} \mu_2(\mathcal{J}, n) \;=\; 2\pi I_1(\mathcal{J}), \tag{2.7.9}$$

where $I_0(x)$ and $I_1(x)$ are the Bessel functions. It is evident that one recovers the results of the $O(2)$ model.

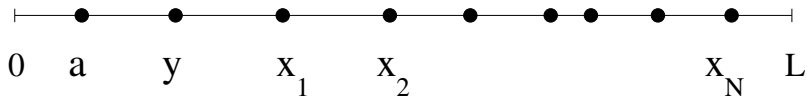


Fig. 2.16 *Feynman gas.*

2.8 Feynman Gas

In this section we discuss a particular one-dimensional gas, known as *Feynman's gas*. Even though it does not belong to the class of systems related to the Ising model, we will see in Chapter 20 that the thermodynamics of this system provides useful information on the spin–spin correlation fucntion of the bidimensional Ising model! For that reason, but also for the peculiarity of this gas, it is useful to present its exact solution.

Let us consider a set of N particles, forced to move along an interval of length L. Let $x_1, x_2, \ldots, x_n$ be their coordinates, while $V(\mid x_i - x_j \mid)$ is their interaction potential. We assume that $V(r)$ is a short-range potential, so that we will consider only the interactions among particles which are close to each other, neglecting all the rest. In this case, the partition function of the system can be written as[5]

$$Z_N(L) \;=\; \int_{0<x_1<x_2\cdot x_N<L} e^{-\beta[V(x_1-x_2)+V(x_2-x_3)+\cdots V(x_{N-1}-x_N)]}\, dx_1 dx_2 \cdots dx_N. \tag{2.8.1}$$

In order to solve the model and find its equation of state, it is natural to find a recursive equation that links Z_N to Z_{N+1}. It is convenient to modify slightly the original problem:[6] this consists of inserting an extra particle in the system but kept fixed at the position y (see Fig. 2.16). The partition function of the new version of the problem, denoted by $P_N(y, L)$, is expressed by

$$P_N(y, L) \;=\; \int_{y<x_1<x_2\cdot x_N<L} e^{-\beta[V(y-x_1)+V(x_1-x_2)+\cdots V(x_{N-1}-x_N)]}\, dx_1 \cdots dx_N. \tag{2.8.2}$$

Let's add a new particle in the position a. It only couples to the particle placed at y, a variable that can now vary on the total interval. Hence we obtain the recursive equation

$$\begin{aligned} P_{N+1}(a, L) &= \int_{a<y} e^{-\beta V(a-y)} dy \int_{y<x_1<x_2\cdot x_N<L} e^{-\beta[V(y-x_1)+V(x_1-x_2)+\cdots]}\, dx_1 \cdots dx_N \\ &= \int_{a<y<L} e^{-\beta V(a-y)}\, P_N(y, L)\, dy. \end{aligned} \tag{2.8.3}$$

[5]The integral over the moments p_i of the particles is gaussian and can be done straightforwardly. It leads to a normalization constant of the partition function, put equal to 1 for simplicity.

[6]It is easy to prove that, in the thermodynamic limit, this new version does not change the macroscopic properties of the system.

The grand canonical partition function of the gas of particles placed in the interval (y, L) is given by[7] (see Appendix A of Chapter 1)

$$\mathcal{Z}(y, L, z) \;=\; \sum_{N=0}^{\infty} z^N \, P_N(y, L), \tag{2.8.4}$$

where $z = e^{\beta\mu}$ (μ is the fugacity of the gas). In the grand canonical ensemble we have the equation of state

$$\mathcal{Z} \;=\; e^{\beta \, P(z) \, V},$$

where $P(z)$ is the pressure of the gas and V its volume. Since $V = L - y$, we have

$$e^{\beta \, P(z) \, (L-y)} \;=\; \sum_{N=0}^{\infty} z^N \, P_N(y, L). \tag{2.8.5}$$

Multiply both the left and right terms of (2.8.3) by z^{N+1} and sum over N. By using (2.8.4) and its expression given by (2.8.5), one obtains the integral equation

$$e^{\beta \, P(z) \, (L-a)} \;=\; z \int_x^L e^{\beta \, P(z) \, (L-y)} \, e^{-\beta V(a-y)} \, dy.$$

Simplifying the common terms present in the left and right sides, in the thermodynamic limit $L \to \infty$ this expression can be written as

$$z^{-1} \;=\; \int_0^{\infty} e^{-\beta \, P(z) \, x} \, e^{-\beta V(x)} \, dx. \tag{2.8.6}$$

This is the central equation of the Feynman gas: it permits us to find the pressure P as a function of z and then to derive all the thermodynamic quantities. For instance, the mean density of particles per unit length is expressed by

$$\rho(z) \;=\; \lim_{L\to\infty} \frac{\langle N \rangle}{L} \;=\; \frac{z}{\beta} \frac{dP(z)}{dz}. \tag{2.8.7}$$

For physical considerations, $\rho(z) > 0$. Hence, $P(z)$ is a monotonic increasing function of z and varying z^{-1} in the left-hand side of (2.8.6), there is only one solution for $P(z)$, as shown in Fig. 2.17.

Appendix 2A. Special Functions

In this appendix we present some properties of the special functions used in the text, namely the $\Gamma(z)$ function and the Bessel functions $I_\nu(x)$.

[7] In this expression the term $1/N!$ is absent for the ordering of the coordinates in the integrals.

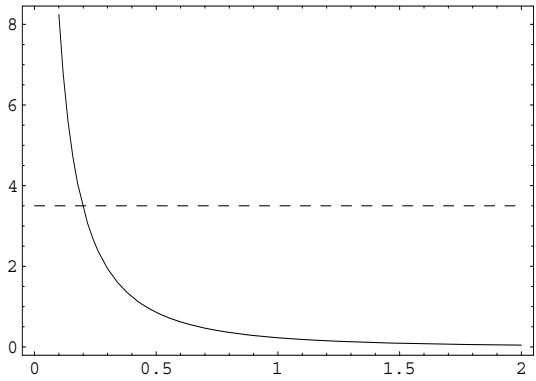


Fig. 2.17 *Graphical solution of (2.8.6). The solid line is the right-hand side of this equation, whereas the dashed line corresponds to a given value of the left-hand side of (2.8.6), i.e.* z^{-1}.

The $\Gamma(z)$ function

The $\Gamma(z)$ function is an analytic function of the complex variable z. For $\mathrm{Re}\, z > 0$ it is defined by the integral representation

$$\Gamma(z) \;=\; \int_0^\infty dt\, t^{z-1}\, e^{-t}. \tag{2.A.1}$$

To obtain $\Gamma(z)$ for other values of z, we can use its analytical continuation. This can be implemented by using the functional equation

$$\Gamma(z+1) \;=\; z\,\Gamma(z) \tag{2.A.2}$$

that is satisfied in the domain of convergence of the integral (2.A.1). In fact, by integrating by part the expression in (2.A.1), one has

$$\Gamma(z) \;=\; \left.\frac{t^z e^{-t}}{z}\right|_0^\infty + \frac{1}{z}\int_0^\infty dt\, t^z e^{-t},$$

and the first term on the right-hand side vanishes when $\mathrm{Re}\, z > 0$, so that we arrive at (2.A.2). By using eqn (2.A.2), we have

$$\Gamma(z) \;=\; \frac{\Gamma(z+1)}{z}.$$

Since $\Gamma(z+1)$ is defined when $\mathrm{Re}\, z > -1$, we have obtained the analytic continuation of the $\Gamma(z)$ function in the strip $-1 < \mathrm{Re}\, z < 0$. Repeating the same reasoning, we can further extend its definition in the next strip $-2 < \mathrm{Re}\, z < -1$, and in all other points of the half-plane $\mathrm{Re}\, z < 0$ as well. For instance, if we wish to compute $\Gamma\left(-\frac{1}{2}\right)$, by using (2.A.2), we have

$$\Gamma\left(-\frac{1}{2}\right) = -2\,\Gamma\left(\frac{1}{2}\right),$$

and the right-hand side can be computed by using (2.A.1). Since $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, we have $\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi}$.

Note that (2.A.2) implies $\Gamma(n+1) = n!$, when n is a positive integer: since $\Gamma(1) = \Gamma(2) = 1$, the iterative application of (2.A.2) gives

$$\Gamma(n+1) \;=\; n\,\Gamma(n) \;=\; n(n-1)\cdots\Gamma(1) \;=\; n!.$$

For that reason, the $\Gamma(z)$ function is considered the generalization, to all real and complex numbers, of the factorial.

The $\Gamma(z)$ function has a pole at $z = 0$. In fact, $\Gamma(1) = 1$, and from (2.A.2), we have

$$\Gamma(z) \;\simeq\; \frac{1}{z}, \qquad z \to 0.$$

We can again use (2.A.2) to prove that $\Gamma(z)$ has other simple poles at $z = -n$, with $n = 1, 2, 3, \ldots$ This result can also be obtained as follows: since the convergence of the integral (2.A.1) only depends on the behavior near the origin, it is convenient to split the interval of the integration as

$$\Gamma(z) \;=\; \int_0^1 dt\, t^{z-1} e^{-t} \;+\; \int_1^\infty dt\, t^{z-1} e^{-t}.$$

The second integral is always convergent, while the first one converges only when $\mathrm{Re}\, z > 0$. In this interval we can use the series expansion

$$e^{-t} \;=\; \sum_{k=0}^{\infty} (-1)^k \,\frac{t^k}{k!},$$

and, if we integrate term by term, we have

$$\Gamma(z) \;=\; \sum_{k=0}^{\infty} (-1)^k \frac{1}{k!(z+n)} \;+\; \int_1^\infty dt\, t^{z-1} e^{-t}. \tag{2.A.3}$$

This expression coincides with the original definition of $\Gamma(z)$ in the domain $\mathrm{Re}\, z > 0$, but it is also valid for other values of z, i.e. it provides the analytic continuation of the $\Gamma(z)$ function for all complex values of z. Equation (2.A.3) makes explicit the presence of the poles of this function for all negative integers, with residues given by

$$\lim_{z \to -n} (z+n)\,\Gamma(z) \;=\; \frac{(-1)^n}{n!}.$$

The plot of $\Gamma(z)$, for real values of z, is given in Fig. 2.18.

When z is very large, $\Gamma(z)$ admits the asymptotic expansion

$$\Gamma(z) \;\simeq\; z^{z-1/2}\, e^{-z} \sqrt{2\pi}, \qquad z \to \infty. \tag{2.A.4}$$

The simplest way to derive this formula is to start from the integral representation (2.A.1) and use the saddle point method, discussed in Appendix A of Chapter 3. With

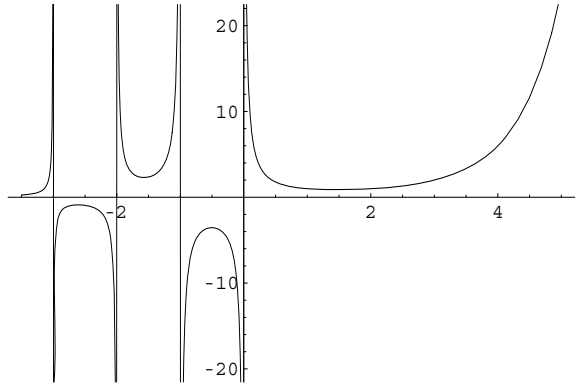


Fig. 2.18 *Plot of the* $\Gamma(x)$ *function for real values of* x.

the change of variable $t = zx$ in (2.A.1), we have

$$\Gamma(z+1) \;=\; z^{z+1} \int_0^\infty dx\, e^{z(\log x - x)}. \tag{2.A.5}$$

The expression in the exponential

$$\varphi(x) \;\equiv\; \log x - x,$$

goes to minus infinity both at $x \to 0$ and $x \to \infty$ and has a maximum at $x = 1$. When $z \to \infty$, the integral (2.A.5) is then dominated by the values around the maximum of $\varphi(x)$. Expanding around $x = 1$

$$\varphi(x) = -1 - \frac{(x-1)^2}{2} + \cdots$$

and substituting in (2.A.5), one has

$$\Gamma(z+1) \;\simeq\; z^{z+1} e^{-z} \int_0^\infty e^{-z\frac{(x-1)^2}{2}} \;\simeq \sqrt{2\pi} z^{z+1/2} e^{-z}.$$

Using $\Gamma(z+1) = z\Gamma(z)$, one arrives at eqn (2.A.4). When $z = n$, with n integer, one has

$$n! \;\simeq\; n^n\, e^{-n}\, \sqrt{2\pi n},$$

i.e. the Stirling approximation of the factorial.

The $\Gamma(z)$ function satisfies many mathematical identities that can often be proved by making use of the functional equation (2.A.2) and its analytic properties. One of them is the so-called *reflection formula*

$$\Gamma(z)\,\Gamma(1-z) \;=\; \frac{\pi}{\sin \pi z}. \tag{2.A.6}$$

To prove it, let's define $\Phi(z) \equiv \Gamma(z)\,\Gamma(1-z)$. Using (2.A.2), it is easy to see that $\Phi(z)$ is a periodic function, with a period equal to 2, i.e. $\Phi(z+2) = \Phi(z)$. Moreover, $\Phi(z)$

has simple poles at all integer values of z, i.e. $z = 0, \pm 1, \pm 2, \ldots$ Hence, the function obtained by multiplying $\Phi(z)$ with the infinite product

$$z \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2}\right).$$

which has precisely simple zeros at $z = 0, \pm 1, \pm 2, \ldots$, is a function without any singularities in the complex plane. Therefore, according to Liouville's theorem, it is a constant, in this case equal to 1, as can be seen by taking the limit $z \to 0$. Hence, using the identity

$$\frac{\sin \pi z}{\pi} \;=\; z \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2}\right), \tag{2.A.7}$$

we arrive at eqn (2.A.6). It is interesting to note that this formula permits us to prove

$$\sum_{k=1}^{\infty} \frac{1}{k^2} \;=\; \frac{\pi^2}{6}. \tag{2.A.8}$$

In fact, using the series expansion

$$\sin \pi z \simeq \pi z - \frac{(\pi z)^3}{3!} + \cdots$$

and comparing the z^3 terms of the right- and left-hand sides of (2.A.7), one gets (2.A.8). This result was originally obtained by Euler.

The Bessel functions $I_\nu(x)$

The Bessel functions $I_\nu(z)$ are defined by the series expansion

$$I_\nu(z) \;=\; \left(\frac{z}{2}\right)^{\nu} \sum_{k=0} \frac{1}{k!\Gamma(k+\nu+1)} \left(\frac{z}{2}\right)^{2k}. \tag{2.A.9}$$

For generic ν, they are regular functions of z in the complex plane, with a branch cut along the real negative semi-axis. When $\nu = \pm n$, they are instead entire functions of z. The functions $I_\nu(z)$ are solutions of the differential equation

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} - (z^2 + \nu^2) w \;=\; 0. \tag{2.A.10}$$

$I_\nu(z)$ and $I_{-\nu}(z)$ are linearly independent except if $\nu = n$, when

$$I_{-n}(z) \;=\; I_n(z). \tag{2.A.11}$$

In fact, substituting $\nu = -n$ in (2.A.9), the first n terms of the series expansion vanish, since these are the poles of $\Gamma(x)$ and therefore one has the identity (2.A.11).

For $\mathrm{Re}\,\nu > \frac{1}{2}$, $I_\nu(z)$ admits the integral representation

$$I_\nu(z) \;=\; \frac{\left(\frac{z}{2}\right)^\nu}{\sqrt{\pi}\,\Gamma\left(\nu+\frac{1}{2}\right)} \int_0^\pi e^{\pm z\cos\theta}\,\sin^{2\nu}\theta\,d\theta. \tag{2.A.12}$$

When $z \to 0$ (with fixed ν) $I_\nu(z)$ has the behavior

$$I_\nu(z) \;\simeq\; \frac{1}{\Gamma(\nu+1)}\left(\frac{z}{2}\right)^\nu, \qquad z\to 0 \quad (\nu \neq -n) \tag{2.A.13}$$

as can be seen from the series expansion (2.A.9), whereas for $z \to \infty$ (with fixed ν)

$$I_\nu(z) \;\simeq\; \frac{e^z}{\sqrt{2\pi z}}\left[1 - \frac{(4\nu^2-1)}{8z} + \frac{(4\nu^2-1)(4\nu^2-9)}{2(8z)^2} + \cdots\right], \qquad z\to\infty. \tag{2.A.14}$$

This behavior can be easily checked by the asymptotic expansion around $z = \infty$

$$I_\nu(z) \;\simeq\; e^z \sum_{n=0}^{\infty} \frac{a_n}{z^{n+\alpha}}, \qquad z\to\infty$$

where α and the coefficients a_n are fixed by substituting this expression into the differential equation (2.A.10). The plot of these functions is shown in Fig. 2.19.

For large values of the index ν, one has

$$I_\nu(\nu x) \;\simeq\; \frac{1}{\sqrt{2\pi\nu}}\,\frac{e^{\nu(\sqrt{1+x^2}-\xi^{-1}}}{(1+x^2)^{1/4}}, \qquad \nu\to\infty, \tag{2.A.15}$$

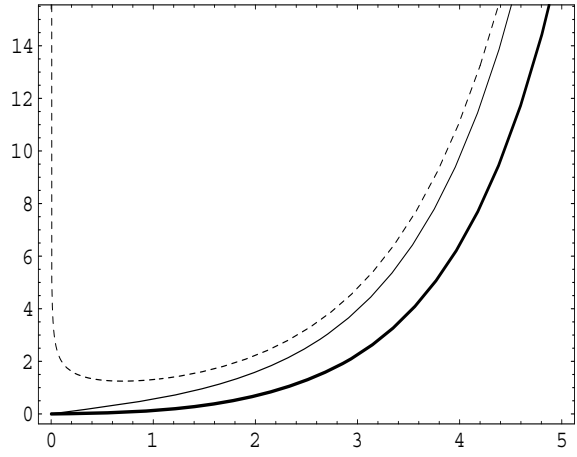


Fig. 2.19 *$I_\nu(x)$ for real values of x ($\nu = -1/3$ dashed line, $\nu = 1$ central line, $\nu = 2$ thick line).*

with

$$\xi^{-1} \;=\; \ln\left(\frac{1+\sqrt{1+x^2}}{x}\right).$$

Using (2.A.9), it is easy to prove that the Bessel functions satisfy the recursive equations

$$\begin{aligned} I_{\nu-1}(z) - I_{\nu+1}(z) \;&=\; \frac{2\nu}{z}\, I_\nu(z) \\ I_{\nu-1}(z) + I_{\nu+1}(z) \;&=\; 2\,\frac{dI_\nu}{dz}(z). \end{aligned} \tag{2.A.16}$$

For $\nu = n$, an integer number, their generating function is given by

$$e^{(z/2)\,(t+1/t)} \;=\; \sum_{n=-\infty}^{+\infty} I_n(z)\, t^n. \tag{2.A.17}$$

The Bessel functions $K_\nu(x)$

In other chapters of this book we will need the modified Bessel functions $K_\nu(x)$. These are also solutions of the differential equation (2.A.10) and can be expressed as linear combinations of the $I_\nu(x)$

$$K_\nu(x) \;=\; \frac{I_{-\nu}(x) - I_\nu(x)}{\sin \pi\nu}. \tag{2.A.18}$$

When ν is an integer, their definition involves the limit $\nu \to n$ of the above expression. The series expansion of these functions directly comes from that of the $I_\nu(x)$. The lowest orders for $x \to 0$ are given by

$$\begin{aligned} K_0(x) \;&=\; -\log x - \gamma + \log 2 + \cdots \\ K_\nu(x) \;&=\; 2^{\nu-1}\Gamma(\nu)\, x^{-\nu} + \cdots \end{aligned} \tag{2.A.19}$$

For large values of x, their aymptotic behavior is

$$K_\nu(x) \;\simeq\; \sqrt{\frac{\pi}{2x}}\, e^{-x} \left[1 + \frac{(4\nu^2-1)}{8x} + \frac{(4\nu^2-1)(4\nu^2-9)}{2(8x)^2} + \cdots\right]. \tag{2.A.20}$$

They satisfy the recursive equations

$$\begin{aligned} K_{\nu-1}(x) - K_{\nu+1}(x) \;&=\; -\frac{2\nu}{x}\, K_\nu(x) \\ K_{\nu-1} + K_{\nu+1}(x) \;&=\; -2K'_\nu(x). \end{aligned} \tag{2.A.21}$$

Their integral representation is given by

$$K_\nu(x) \;=\; \int_0^\infty e^{-z\,\cosh t}\, \cosh(\nu t)\, dt. \tag{2.A.22}$$

Appendix 2B. *n*-dimensional Solid Angle

The solid angle in n-dimensional space is expressed by the formula

$$\Omega(n) = \int d\Omega^{(n)} = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}. \tag{2.B.1}$$

This is an analytic function of n and, for this reason, can be computed for arbitrary values of n, not necessarily integers. However, its behavior is rather peculiar: for the presence of the $\Gamma(x)$ in its denominator (that diverges both at $x = 0$ and $x \to \infty$) the n-dimensional solid angle goes to zero at the two edges of the positive semi-axis, and presents a maximum at $n_m \simeq 7.2$ (see Fig. 2.20).

While it is easy to understand that $\Omega(n)$ should vanish when $n \to 0$ (if there is no space, the solid angle between its axes must vanishes as well), it is apparently paradoxical that it also vanishes when the number of the dimensions increases to infinity: our geometrical intuition would rather suggest that the solid angle should increase by increasing the dimensions, in sharp contrast however with the formula (2.B.1). There is a way to understand the asymptotic behavior of $\Omega(n)$ for large n. Using the asymptotic expansion

$$\Gamma(x) \simeq x^{x-1/2}\, e^{-x}\sqrt{2\pi}, \qquad x \to \infty$$

we have

$$\Omega(n) \simeq \frac{2\pi^{n/2}}{\left(\frac{n}{2}\right)^{\frac{n-1}{2}}} \frac{e^{n/2}}{\sqrt{2\pi}}, \qquad n \to \infty. \tag{2.B.2}$$

Let us consider an n-dimensional cube, with side of length 1: its volume is given by

$$V = 1 \times 1 \times 1 \cdots \times 1 = 1.$$

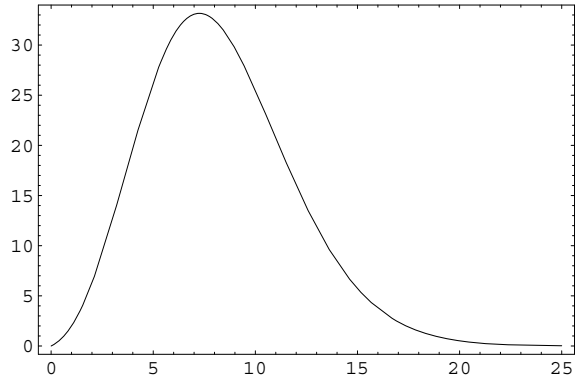


Fig. 2.20 *The solid angle* $\Omega(n)$ *as a function of* n.

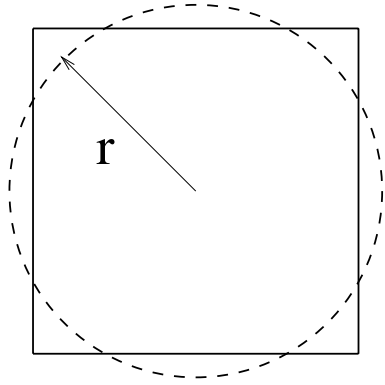


Fig. 2.21 *Cube and sphere of unit volume.*

On the other hand, the estimate of such a volume can be reached by using a sphere of radius r, of the order of the distance from the center of the cube to one of its vertices,[8] namely (see Fig. 2.21)

$$r \simeq \sqrt{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}} = \sqrt{\frac{n}{2}}. \tag{2.B.3}$$

Let us call a_n the proportional constant

$$r(n) = a_n \sqrt{\frac{n}{2}}, \tag{2.B.4}$$

Since the volume of an n-dimensional sphere of radius R is given by

$$V(n) = \frac{R^n}{n}\,\Omega(n),$$

posing $V(n) = 1$ and using the asymptotic behavior of $\Omega(n)$, it is easy to see that the decreasing behavior of the solid angle is compensated by the increasing of the term $[r(n)]^n$. This argument provides the explanation of the unusual behavior of the solid angle $\Omega(n)$ for large n. It is interesting to note that, for $n \to \infty$, we have

$$a_\infty = \frac{1}{\sqrt{\pi e}}.$$

Appendix 2C. The Four-color Problem

The four-color problem has been, for a long time, one of the most famous open problems of mathematics and it has played an important role in the development of modern

[8]This formula shows the counter-intuitive behavior of n-dimensional spaces: in fact, increasing the number of the dimension, the volume of the unit cube is always equal to 1, whereas its diagonal diverges as $\sqrt{n}$.

applied mathematics. Since 1852, several ideas pursued for its solution have contributed to the progress of many other fields of mathematics: graph theory, for instance, or topology have both received an enormous stimulus from this problem.

As with many other famous problems (Fermat's last theorem, for instance), its statement is rather simple: the *four-color conjecture* states that four colors are enough to color any geographical map (drawn on a plane or a sphere), with the constraint that regions with a common border have a different color.

Let's make a few comments. First of all, the conjecture refers to maps consisting of contiguous regions, i.e. regions that share a border not made of a single point. Otherwise, as a map consisting of sliced pies for instance, it would be necessary to use as many colors as the number of slices (see Fig. 2.22).

The second comment is that the four-color conjecture does not refer to maps that consist of disconnected regions. In this case, in fact, it is easy to show examples of maps with five regions which need five colors to distinguish them. One of these examples is in Fig. 2.23: the region indicated by the letter E is considered a unique region, even though it is made by two disconnected parts.

After these remarks, let us discuss in detail the four-color conjecture, starting from its history. The problem owes its origin to the mathematician Francis Guthrie in 1852: a few months after he graduated at University College, London, he wrote a letter to his brother Frederick, who was completing his studies at the same college under the supervision of the mathematician Augustus De Morgan. In this letter, he

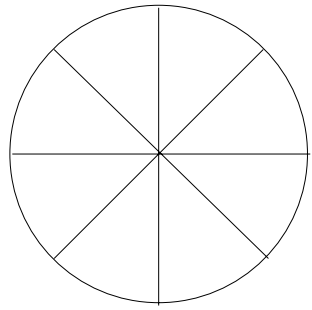


Fig. 2.22 *This map is excluded from the four-color conjecture.*

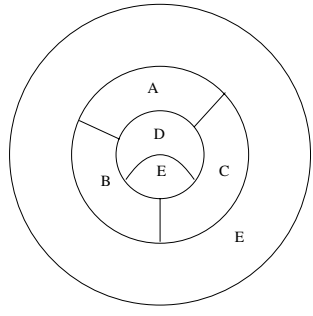


Fig. 2.23 *Map with disconnected regions that does not fall into the four-color theorem.*

expressed his feeling that four colors seem sufficient to distinguish the regions of a planar geographical map and asked him whether he was aware of any proof of that statement. Frederick's answer was negative and also De Morgan, asked about the problem, was not aware of any proof of the statement. However, De Morgan quickly showed that four colors are indeed necessary to color any planar map, pointing out an explicit example for which three colors are not sufficient. This map, shown in Fig. 2.24, is made of four regions, each of them adjacent to all the others, so that three colors are not sufficient.

As discussed below, using Euler's formula for the graphs, De Morgan was also able to prove that it is impossible to draw a planar map with five regions, each of them adjacent to the remaining four. Following this result, he stated that the case in which it is necessary to use five colors never occurs, namely that the four-color conjecture is true. However, he also realized that his conclusion was wrong: the fact that five mutual adjacent regions in the plane cannot exist is not equivalent to proving the four-color conjecture.

It is easy to see that it is not correct to assume, as done by De Morgan, that the number of colors necessary to draw a map is equal to the number of countries mutually adjacent. A counter-example is shown in Fig. 2.25. Although there are no more than three regions mutually adjacent in this map, nevertheless it is necessary to use four colors: three for the external ring and one for the central region.

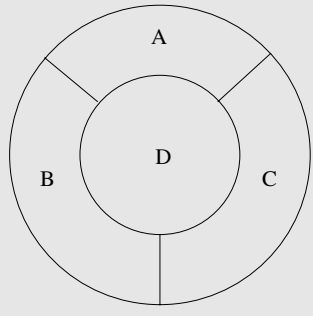


Fig. 2.24 *A map that needs four colors.*

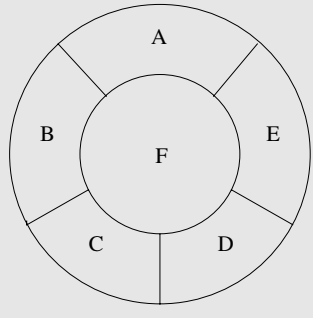


Fig. 2.25 *A counter-example to the presumed proof by De Morgan of the four-color theorem.*

To summarize: the map of Fig. 2.24 points out a *local obstruction* to draw a map using three colors, whereas the map shown in Fig. 2.25 provides an example of a *global obstruction* to color a map with only three colors. Therefore, the result by De Morgan on the impossibility to draw a map in which five regions are mutually adjacent, only addresses the *local* aspect of the problem and not the *global* one. Many solutions of the four-color theorem that have been proposed during the years show the same mistake made by De Morgan. The official birthday of the four-color conjecture is 13 June 1878, when the English mathematician Arthur Cayley discussed the problem at the London Mathematical Society. His contribution was published in the proceedings of the Society and this was the first written version of the problem and its official birthday.

As we have seen, the four-color conjecture is easy to state but this simplicity is deceptive. It has surely been deceptive for many mathematicians, even famous, who have attempted to prove it. Hermann Minkowski, for instance, told his students that, in his opinion, the conjecture was not proved yet simply because only mediocre mathematicians had been involved in it and he would have been surely able to achieve its solution. But, after a long period spent on the problem, he was forced to admit his failure. *God punishes me for my arrogance, my proof has unfortunately many flaws*, was his last comment.

The difficulty of the proof is because the statement refers to *all possible maps*. From this point of view, to establish that a given map can be colored by four colors is not particularly useful. Let us see how topology and graph theory can help in a clearer formulation of the problem. As far as the theorem is concerned, it is obvious that the shapes and the actual sizes of the regions are completely irrelevant: they can be deformed with continuity without altering the nature of the problem. The only thing that matters is the relation among the different regions, i.e. the topology of the map. This can be highlighted by means of the *dual graph* (see Fig. 2.26). As we will see in other parts of the book, this concept plays an important role in statistical physics and enters the solution of the Ising model and other similar models (see, for

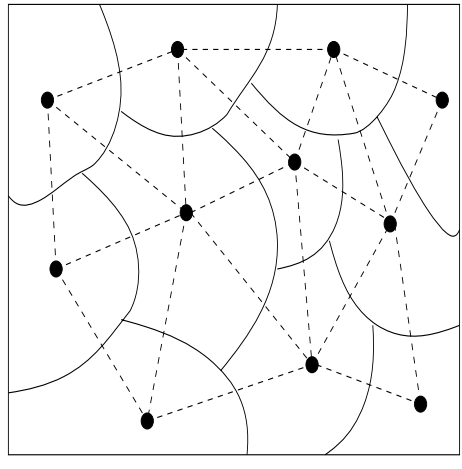


Fig. 2.26 *Dual graph of a map.*

instance, Chapter 4). The dual graph of a map is obtained as follows. Firstly, inside each region let's draw a point, denoted as a *vertex*; secondly, let's join the vertices by a line according to the rule that two vertices will be linked to each other if and only if their regions share a common border. At the end of this procedure, we have obtained the *dual graph* of the original map. The dual graph shows immediately the topological properties of the initial map and, at the same time, points out the close relationship between the four-color problem and the Potts model discussed in the text. In fact, the coloring problem of a map can be easily formulated in terms of the dual graph as follows: color the vertices of the dual graph in such a way that each pair of connected vertices has a different color. Note that an important property of the dual graphs is that all its links can never cross. For that reason, they are only a subclass of more general graphs.

Graph theory allows us to obtain some useful information on the dual graph. Let F be the number of *faces* (i.e. the different regions of the area of the plane spanned by the graph), V the number of vertices, and E the number of edges of a given graph. Euler found a formula that links these quantities together:

$$V - E + F = 1. \tag{2.C.1}$$

It is easy to check the validity of Euler's formula by analyzing a few examples, such as those shown in Fig. 2.27. The proof of this theorem is not difficult and it is interesting to follow its main steps. It consists of simplifying the graph by means of two procedures which never alter the combination $V - E + F$, till we arrive at the most elementary graph made of only one point, for which it trivially holds that $V - E + F = 1$.

Starting from an arbitrary graph, the first procedure consists of removing one of its external edges (see Fig. 2.28). In this way, E decreases by 1, the same happens

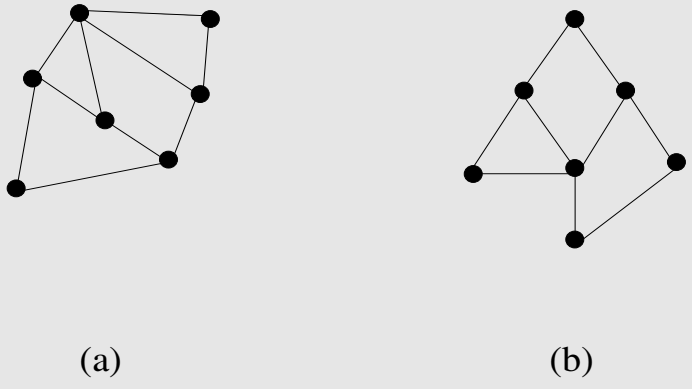


Fig. 2.27 *For every graph, it holds that* $V - E + F = 1$. *In the graph (a), we have* $V = 7$, $E = 10$, *and* $F = 4$, *while in the second* $V = 7$, $E = 9$, *and* $F = 3$.

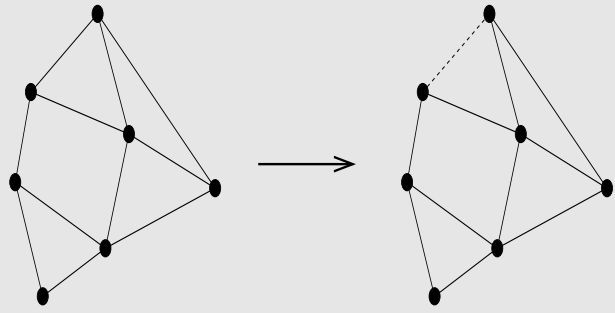


Fig. 2.28 *When an external edge is removed, the quantity $V - E + F$ does not change its value.*

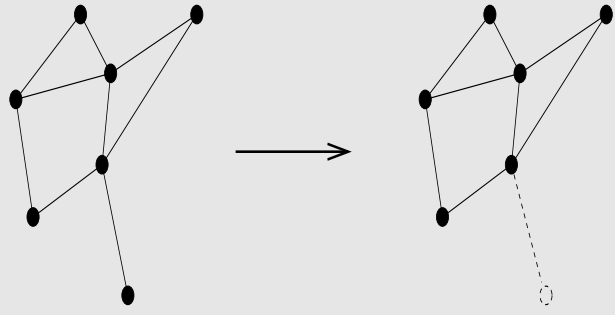


Fig. 2.29 *If an isolated vertex and the edge that links it to the rest of the graph are removed, the quantity $V - E + F$ does not change its value.*

for F, while F remains unchanged. Hence, in this procedure the expression $V-E+F$ does not change.

Let's image that, by iterating this procedure we have a vertex that is linked to the rest of the graph by only one edge. At this point, there is the second procedure to use: it consists of removing both the vertex and the edge (see Fig. 2.29). In this operation V decreases by 1, the same happens for E, while F instead remains unchanged. Hence, also in this case, the combination $V - E + F$ is the same before and after the procedure.

Starting from the initial graph and acting by the two procedures, we can systematically remove all the external edges and vertices, until we reach a graph made of only one vertex. The final graph has $V = 1$ with $E = F = 0$. Since in all these operations the quantity $V - E + F$ is always the same, we thus arrive at Euler's formula (2.C.1).

By using Euler's formula, it is easy to prove the theorem by De Morgan previously mentioned: there does not exist any planar map in which there are five countries, each of them that shares a border with the remaining four (see Fig. 2.30). This means that

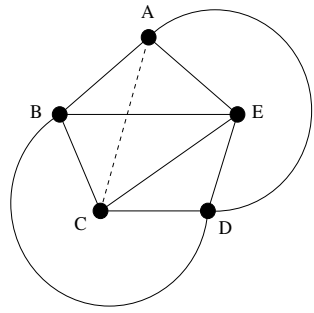


Fig. 2.30 *De Morgan's theorem: It is impossible to draw a planar graph of five vertices so that each vertex is linked to the others. However one proceeds, there will always be two vertices that cannot be linked by an edge without crossing the previous edges (in the figure these are the A and C vertices).*

it is impossible to draw a graph with five vertices, each of them linked to the other four, without crossing its edges.

The proof is by contradiction. Let us assume that there exists a graph made of five vertices, each of them linked to the other four. It is convenient to consider the region outside the graph as an additional face, so that each edge always separates two faces. If we do so, Euler's formula gets modified as follows:

$$V - E + F = 2. \tag{2.C.2}$$

In the graph under investigation $V = 5$. Since, by hypothesis, each of them is linked to the other four vertices, we have $E = 5 \cdot 4/2 = 10$. Hence, from Euler's formula one should find $F = 7$. Let us use this last result to do the computation differently: since there are seven faces, each of them with three edges (the external face is included as well in this computation), one should have $3 \cdot 7 = 21$ edges. However, since each of them separates two faces, the previous number should be twice the number of the edges. But 21 is an odd number, so that we arrive at a contradiction. This implies that the initial hypothesis is false, i.e. that it is impossible to draw a planar graph with five vertices, each of them linked to the others. The result found by De Morgan is not enough, though, to prove the four-color theorem. The history of all the different attempts to prove this theorem is an interesting chapter of mathematics, too long to be summarized.[9] Here we will only mention two important developments. The first refers to the study by Beraha who, during his PhD thesis, noticed that the chromatic polynomial of a planar graph often has zeros close to the sequence of numbers

$$B_n = 4\cos^2(\pi/n), \quad n = 2, 3, 4 \tag{2.C.3}$$

known in the literature as *Beraha numbers*. He was able to prove that any Beraha number that is not an integer (with the possible exception of B_{10} for which his proof

[9] The interested reader may read the article by Kenneth Appel and Wolfgang Haken, *The solution of the four color problem*, published in *Scientific American, n. 237*, October 1977. Additional references are given at the end of the chapter.

did not apply) could be a zero of the chromatic polynomial of a finite graph. Vice versa, he also proved that certain graphs with the shape of a strip of length L and finite width have zeros that converge to each Beraha number (not an integer), when $L \to \infty$. In other work made in collaboration with Kahane, he also observed that some families of graphs have the zeros of their chromatic polynomials that converge to the value $q = 4$, which seems to be their accumulation point when $L \to \infty$.

The original idea of Beraha and Kahane was to find a planar graph with a zero of the chromatic polynomial at $q = 4$. The discovery of such a graph would have been a counter-example of the four-color conjecture. By a twist of fate, between the submission of the article (1976) and its publication, the conjecture turned into a theorem! In fact, in 1976 two mathematicians at the University of Illinois, Kenneth Appel and Wolfgang Haken, announced that they were able to prove the validity of the four-color conjecture, although along a different path with respect to traditional mathematical demonstrations. Indeed, one of the main steps of the proof relied on the crucial use of computer analysis. Besides the remarkable result on the four-color problem, the important aspect of the proof by Appel and Haken is precisely the unusual method that was adopted for the proof. This has radically changed the concept of *mathematical proof*. According to their own words, "Obviously it would be possible that, some day, someone will find a shorter proof of the four color theorem. However, it may also be that such a demonstration does not exist at all: in this case, a new kind of theorem has emerged, a theorem that does not admit a traditional demonstration. Even though the four color theorem will not belong to such a category, nevertheless it provides a good example for a theorem of such a kind and there is no reason to think that this is an isolated case, there may be a large number of problems where such an analysis is indeed needed."

References and Further Reading

In his book, quoted at the beginning of this chapter, E.A. Abbott provides a vivid testimony of the Victorian Era in which he lived. This book is a classic of the scientific literature, filled with social satire, as well as *Gulliver's Travels*, and rich in deep theoretical hints on the nature of the space.

E. A. Abbott, *Flatland. A Romance of Many Dimensions*, Barnes & Noble, New York, 1963.

To learn about the percolation phenomenon, we refer to:

D. Stauffer, A. Aharony, *Introduction to Percolation*, Taylor & Francis, London, 1991.

The physics of polymers is discussed in the books:

P.G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press 1979.

J. de Cloizeaux, G. Jannick, *Polymers in Solution*, Claredon Press, Oxford, 1989.

Further references on the four-color problem are the following:

T.L. Saary, P.C. Kainen, *The Four-Color Problem*, Dover, New York, 1977.

R. Wilson, *Four Colours Suffice*, Allen Lane, London, 2002.

For the relation between the Potts model and the four-color problem see:

H. Saleur, *The antiferromagnetic Potts model in two dimensions: Berker–Kadanoff phase, anti-ferromagnetic transition, and the role of Beraha numbers*, Nucl. Phys. B 360 (1991), 219.

H. Saleur, *Zeros of chromatic polynomials: A new approach to Beraha conjecture using quantum groups*, Comm. Math. Phys. 132 (1990), 657.

Problems

1. Ising model with vacancies

Consider the one-dimensional Ising model with the familiar hamiltonian

$$\mathcal{H} = -J \sum_{i=1}^{N} S_i \, S_{i+1}$$

but with the values of the spins $S_i = \{0, \pm 1\}$. The null value may be interpreted as a vacancy of the corresponding site. Use the transfer matrix method to compute the partition function of the system.

2. Multiple couplings

Consider the one-dimensional Ising model with N spins $\sigma_i = \pm 1$, with the hamiltonian given by

$$\mathcal{H} = -J_1 \sum_{i=1}^{N-1} \sigma_i \, \sigma_{i+1} - J_2 \sum_{i=1}^{N-2} \sigma_i \sigma_{i+2}.$$

Compute the free energy of such a system and analyze its behavior by varying the parameters J_1 and J_2.

Hint. It is convenient to define $\tau_i = \sigma_i \sigma_{i+1}$ which assume values $\tau_i = \pm 1$.

3. Grand canonical partition function

Consider a one-dimensional lattice gas on M sites. The variable t_i has values 0 and 1, according to whether the relative site is occupied or not. The Hamiltonian is given by

$$\mathcal{H} = -J \sum_{i=1}^{M} t_i \, t_{i+1}$$

with $t_{i+M} = t_i$. Moreover, we assume that

$$\sum_{i=1}^{M} t_i \;=\; N.$$

a Compute the grand canonical partition function $\mathcal{Z}(z, M, T)$ of such a system.

b Show that for $J > 0$ the zeros of $\mathcal{Z}(z, M, T)$ are along the circle $\mid z \mid = e^{-\mathcal{J}}$ in the complex plane in the variable z, where $\mathcal{J} = \beta J$.

4. Ising model in a magnetic field

Let's consider the one-dimensional Ising model with nearest neighbor interactions in the presence of a magnetic field. Show that the partition function $Z_N(T, B)$ is an even function of B and compute its first non-vanishing term of the series expansion with respect to B.

5. Ising model in a purely imaginary magnetic field

Consider the one-dimensional Ising model with next neighbor interaction in the presence of a purely imaginary magnetic field $B = ih$. Study the correlation length of the model, expressed by eqn (2.2.28), and show that there exist an infinite number of pairs $(\mathcal{J},\, h)$ for which it diverges.

6. Correlation functions

For the one-dimensional Ising model with next neighbor interaction:

a compute the correlation function $\langle \sigma_i\, \sigma_j \sigma_k \sigma_l \rangle$, with $i > j > k > l$;

b use the expression of the quantity above to compute the derivative of the correlation function $\langle \sigma_i\, \sigma_j \sigma_k \rangle$ with respect to the magnetic field B, for $B = 0$.

7. Potts model

Consider the one-dimensional q-state Potts model, whose hamiltonian is given by

$$\mathcal{H} \;=\; -J \sum_{i=1}^{N} \delta(\sigma_i, \sigma_{i+1}).$$

a Compute the expectation value $\langle \delta(\sigma_k, \sigma_{k+1}) \rangle$.

b Study this quantity for negative values of q.

8. One-dimensional Tonks gas

Consider a gas made of molecules of size a that can move freely along a line of length L. Their interaction potential is given by

$$U(x_i - x_j) \;=\; \begin{cases} +\infty\,, & \text{if } \mid x_i - x_j \mid \,<\, a \\ 0 \quad\;\; , & \text{if } \mid x_i - x_j \mid \,>\, a. \end{cases}$$

Compute the exact expression of the partition function of the system and its equation of state.

9. One-dimensional gas

Considerar a one-dimensional gas of N particles, with potential interaction

$$V(x_1, x_2, \ldots, x_N) \;=\; -\sum_{i<j} \log\, \tanh^2 \left(\frac{x_i - x_j}{2} \right).$$

a Discuss the validity of the next neighbor approximation for this potential.

b Give a numerical estimate of the gas pressure at $z = \frac{1}{2\pi}$.

c Compute the mean values of the number of particle per unit length and check the validity of the next neighbor approximation.

10. Thermodynamics of one-dimensional oscillators

Consider a set of N particles of mass m placed along a line of length L. Let p_i and q_i be their momentum and coordinate respecively. Their hamiltonian is

$$H \;=\; \sum_{i=1}^{N} \frac{p_i^2}{2m} + \frac{1}{2} k \sum_{i=1}^{N-1} (q_i - q_{i+1})^2.$$

a Determine the frequencies of the normal modes of the system.

b Compute the partition function.

3
Approximate Solutions

The approximate computations are a fundamental part of physical science.

Steven Weinberg

The exact solutions of one-dimensional systems discussed in the previous chapter are particularly simple and mathematically elegant. Unfortunately, they are more an exception than a rule: for higher dimensional lattices, the computation of the partition functions of statistical models usually poses a formidable problem from a mathematical point of view. For that reason, it is of the utmost importance to develop some approximate methods that permit us to analyze the most relevant physical aspects and to extract an estimate of critical exponents or other thermodynamic quantities. In this chapter we will discuss several approximate solutions of the Ising model and its generalizations.

3.1 Mean Field Theory of the Ising Model

In the Ising model, each spin interacts both with the external magnetic field and the one created by neighbor spins. The magnetic field created by the spins is obviously a dynamical variable, which cannot be controlled by external knobs, for its value changes with the fluctuations of the configurations. The mean field approximation consists of replacing the magnetic field created by the spins by its thermal average. This substitution gives rise to an interaction among *all* spins, so that the mean field solution is essentially equivalent to solve the model in the limit $d \to \infty$: for an infinite-dimensional lattice, the mean field solution is then an exact one. Although this limit may appear artificial and distant from the actual physical features of magnets, we shall nevertheless see that the mean field solution is able to capture the main properties of phase transition in the Ising model, in a particularly simple way that is also sufficiently accurate.

Consider the Hamiltonian of a d-dimensional lattice with N spins

$$\mathcal{H} = -\frac{J}{2} \sum_{\langle i,j \rangle} \sigma_i \, \sigma_j - B \sum_i \sigma_i. \tag{3.1.1}$$

Let's introduce the magnetization, defined by

$$m = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sigma_i \right\rangle, \tag{3.1.2}$$

and let's express the product of the spins as

$$\begin{aligned}\sigma_i\,\sigma_j &= (\sigma_i - m + m)(\sigma_j - m + m)\\ &= m^2 + m\,(\sigma_i - m) + m\,(\sigma_j - m) + (\sigma_i - m)(\sigma_j - m).\end{aligned}$$

The last term is quadratic in the spins, of the form $(\sigma_i - \langle\sigma\rangle)(\sigma_j - \langle\sigma\rangle)$. The mean field approximation consists of completly neglecting this term, replacing the previous Hamiltonian with

$$\begin{aligned}\mathcal{H} &= -\frac{J}{2}\sum_{\langle i,j\rangle}\sigma_i\,\sigma_j - B\sum_i \sigma_i\\ &\simeq -\frac{J}{2}\sum_{\langle i,j\rangle}[-m^2 + m(\sigma_i + \sigma_j)] \;-\; B\sum_i \sigma_i. \qquad (3.1.3)\end{aligned}$$

Let z be the *coordination number* of the lattice, i.e. the number of neighbor spins with whom each spin interacts.[1] The first term in (3.1.3) can be expressed as

$$-\frac{J}{2}\sum_{\langle i,j\rangle}(-m^2) \;=\; \frac{1}{2}\,Jzm^2\sum_i \;=\; \frac{1}{2}\,JNzm^2,$$

while

$$-\frac{J}{2}\,m\;\sum_{\langle i,j\rangle}(\sigma_i + \sigma_j) \;=\; -Jzm\sum_i \sigma_i.$$

In the mean field approximation, the hamiltonian thus becomes

$$\mathcal{H}_{cm} \;=\; \frac{1}{2}\,NJzm^2 - (Jz\,m + B)\sum_i \sigma_i. \qquad (3.1.4)$$

Since all spins are decoupled, it is simple to compute the partition function

$$\begin{aligned}Z_N^{cm}(T,B) &= \sum_{\{\sigma\}} e^{-\beta\mathcal{H}_{cm}} = e^{-\frac{1}{2}\,N\mathcal{J}zm^2}\left(\sum_{\sigma=\pm1} e^{(\mathcal{J}zm+\mathcal{B})\sigma}\right)^N\\ &= e^{-\frac{1}{2}\,N\mathcal{J}zm^2}\,\left[2\cosh(2\mathcal{J}m + \mathcal{B})\right]^N. \qquad (3.1.5)\end{aligned}$$

The free energy per spin is expressed by

$$F^{cm}(T,B) \;=\; -\frac{1}{\beta N}\ln Z_N \;=\; \frac{1}{2}\,Jzm^2 - \frac{1}{\beta}\ln\left[2\cosh(\mathcal{J}m + \mathcal{B})\right]. \qquad (3.1.6)$$

Since the magnetization is given by the derivative of F with respect to B, it must satisfy the self-consistency equation

$$m \;=\; -\frac{\partial F}{\partial B} \;=\; \tanh(\mathcal{J}zm + \mathcal{B}). \qquad (3.1.7)$$

This equation was first obtained by Bragg and Williams and, for this reason, the mean field solution is also known as the Bragg–Williams approximation. To see whether there

[1] For example, a two-dimensional square lattice, $z = 4$, whereas in a three-dimensional cubic lattice $z = 6$.

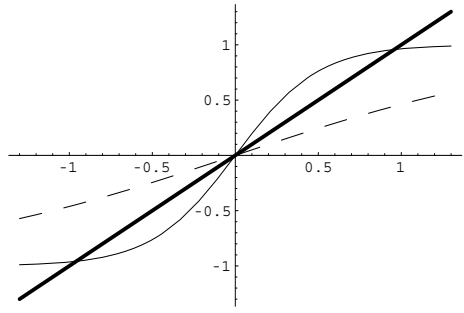


Fig. 3.1 *A graphical solution of eqn (3.1.8).*

is a spontaneous magnetization as $B \to 0$, one needs to look for non-zero solutions of the trascendental equation

$$m \,=\, \tanh(\mathcal{J}zm). \tag{3.1.8}$$

This problem can be solved by a graphical method, i.e. by checking whether the plot of the function on the right-hand side of (3.1.8) has intersections or not with the straight line $y = m$, as shown in Fig. 3.1. The value $m = 0$ is obviously always a solution of (3.1.8). However, when the derivative at the origin of the function $\tanh(\mathcal{J}zm)$ on the right-hand side of (3.1.8) is greater than 1, there are two other solutions (of different sign but equal in absolute value) $\pm m_0$. Hence, for $\mathcal{J}z = 1$, i.e. when the temperature becomes equal to

$$T_c \,=\, \frac{Jz}{k}, \tag{3.1.9}$$

the system undergoes a phase transition, from a disordered to an ordered phase. For $T < T_c$ the spins are oriented along the direction of the external magnetic field before it will be set to zero: hence, one has $m = m_0$ if $B \to 0^+$ while $m = -m_0$ in the other limit $B \to 0^-$. If the mean value of the magnetization is different from zero, the Z_2 symmetry of the model is spontaneously broken: in the ordered phase, the only effect left from the original Z_2 symmetry of the hamiltonian is the mapping between the two different solutions, i.e. $m_0 \to -m_0$ and vice versa.

Let us now compute, in the mean field approximation, the critical exponents. Using the variable

$$t = \frac{T - T_c}{T_c},$$

with T_c given in (3.1.9), eqn (3.1.7) can be written more conveniently as

$$m \,=\, -\frac{B}{kT_c} + (1 + t)\,\mathrm{arcth}\, m. \tag{3.1.10}$$

Consider firstly the case $B = 0$. Close to the critical point $t \approx 0$ and the spontaneous magnetization is also small. By expanding the right-hand side of the equation above,

we have

$$m_0 \;=\; (1+t)\,\mathrm{arcth}\, m_0 = (1+t)\,\left[m_0 + \frac{1}{3}m_0^3 + \frac{1}{5}m_0^5 + \cdots\right].$$

Solving with respect to m_0, we obtain

$$m_0 \;=\; (-3t)^{\frac{1}{2}}\{1+\mathcal{O}(t)\}. \tag{3.1.11}$$

Hence, the critical exponent β is given by

$$\beta \;=\; \frac{1}{2}. \tag{3.1.12}$$

In order to compute the magnetic susceptibility, we derive both terms of (3.1.10) with respect to B. Since $\chi = \frac{\partial m_0}{\partial B}$, we have

$$\chi \;=\; -\frac{1}{kT_c} + (1+t)\,\left(\frac{1}{1+m^2}\right)\,\chi.$$

For $B = 0$ and $T > T_c$, one has $m_0 = 0$ and therefore χ satisfies the equation

$$\chi \;=\; -\frac{1}{kT_c} + (1+t)\chi.$$

Hence

$$\chi \;=\; \frac{1}{kT_c}\,t^{-1}.$$

In the same way, at $B = 0$ but $T < T_c$, one obtains

$$\chi \;=\; \frac{1}{2kT_c}\,(-t)^{-1}.$$

Hence, for the critical exponent γ we get the value

$$\gamma \;=\; 1. \tag{3.1.13}$$

With the above computation we can also determine the universal ratio

$$\frac{\chi_+}{\chi_-} \;=\; 2.$$

To obtain the exponent δ, consider the equation of state (3.1.10) at $t = 0$. By using the series expansion of the hyperbolic function and simplifying the result, one has

$$\frac{B}{kT_c} \simeq \frac{1}{3}m^3 + \mathcal{O}m^5,$$

i.e. $m \simeq B^{1/3}$ and therefore

$$\delta \;=\; 3. \tag{3.1.14}$$

Finally, to obtain the exponent α it is convenient to consider the free energy (3.1.6) in the vicinity of the critical point. Using eqn (3.1.7) and the identity

$$\cosh x \;=\; \frac{1}{(1-\tanh^2 x)^{1/2}},$$

the free energy can be equivalently expressed as

$$F^{cm}(T,B) \;=\; \frac{1}{2}\,Jzm^2 - \frac{1}{2\beta}\ln\left[\frac{4}{(1-m_0^2)}\right]. \tag{3.1.15}$$

Let us take $B=0$. For $T>T_c$, one has $m_0=0$ and the free energy is simply equal to

$$F^{cm}(T,0) \;=\; -\frac{1}{\beta}\ln 2.$$

For $T<T_c$, $m_0\neq 0$ and by series expanding (3.1.15) we have

$$F^{cm}(T,0) \;=\; -\frac{1}{\beta}\ln 2 - \frac{1}{2\beta}m_0^2(1-\mathcal{J}z) + \cdots$$

Using (3.1.11), for t sufficiently small and negative, the free energy is given by

$$F^{cm}(T,0) \simeq -\frac{1}{\beta}\ln 2 - \frac{3}{4}t^2 + \cdots$$

Since $F \simeq t^{2-\alpha}$, for the critical exponent α we get

$$\alpha \;=\; 0. \tag{3.1.16}$$

Note that in the mean field approximation both the free energy and the mean value of the internal energy do not have a discontinuity at $T=T_c$, while the specific heat has a jump. Since each spin interacts with all the others, the spin–spin correlation function does not depend on their separation, so that $\eta=0$. The last critical exponent ν can be extracted by the scaling laws and its value is $\nu=1/2$.

In summary, the mean field approximation is efficient in showing the existence of a phase transition in the Ising model and in predicting its qualitative features. However, there are many aspects that are unsatisfactory from a quantitative point of view. For instance, it predicts the occurrence of a phase transition even for the case $d=1$, that is excluded by the exact analysis of Chapter 2. Moreover, even when there is a phase transition, as in $d=2$ or $d=3$, the mean field theory gives an estimate of the critical temperature that is higher than its actual value and the critical exponents differ from their known values in both cases, as shown in Table 3.1.

The universality of the results obtained in this approximation is due to the absence of spin fluctuations: once we substitute the dynamical magnetization of the spins with its thermal average, the long-range correlation among all spins suppresses in fact their fluctuations with respect to their mean value. This long-range order favors the energy contribution in the free energy but does not take into proper account the entropy contribution: for this reason, one obtains a value of the critical temperature T_c higher than the actual one.

Table 3.1: *Critical exponents of the Ising model for various lattice dimensions.*

Exponents	Mean field	Ising $d = 1$	Ising $d = 2$	Ising $d = 3$
α	0	1	0	0.119 ± 0.006
β	1/2	0	1/8	0.326 ± 0.004
γ	1	1	7/4	1.239 ± 0.003
δ	3	∞	15	4.80 ± 0.05
ν	1/2	1	1	0.627 ± 0.002
η	0	1	1/4	0.024 ± 0.007

3.2 Mean Field Theory of the Potts Model

The mean field approximation for the q-state Potts model shows a novel aspect with respect to the Ising model: a second-order phase transition for $q \leq 2$ but a first-order phase transition for $q > 2$.

In the mean field theory, each of the N spins of the lattice interacts with all the remaining $(N-1)$ ones. In this approximation the Hamiltonian can be written as

$$\mathcal{H}_{mf} \;=\; -\frac{1}{N}\, J\, z \sum_{i<j} \delta(\sigma_i, \sigma_j), \tag{3.2.1}$$

where z is the coordination number of the lattice and we have introduced, for convenience, a factor $1/N$ in the coupling constant. To solve the model, let's firstly write the free energy of the system

$$F[\mathcal{C}] \;=\; U[\mathcal{C}] - TS[\mathcal{C}] \tag{3.2.2}$$

as a function of the configurations $\mathcal{C}$ of the spins and then let's proceed to determine its minimum. Such a computation can be simplified by noticing that the energy is a highly degenerate function of the system configurations. Consequently, it is more convenient to express the Hamiltonian in terms of a proper set of variables that makes explicit such a degeneracy. Given a configuration $\mathcal{C}$ of the spins, let us denote by $x_i = N_i/N$ the fraction of the spins that are in the i-th state, with $i = 1, 2, \ldots, q$. Obviously

$$\sum_{i=1}^{q} x_i \;=\; 1. \tag{3.2.3}$$

Since there are $\frac{1}{2}N_i(N_i - 1)$ couplings of the type i in the hamiltonian (3.2.1), the energy $U[\mathcal{C}]$ of this configuration is given by

$$U[\mathcal{C}] \;=\; -\frac{1}{2N}\, Jz \sum_{i=1}^{q} N_i(N_i - 1).$$

Dividing by the number of spins and taking the limit $N \to \infty$, one has

$$\frac{U[\mathcal{C}]}{N} \simeq -\frac{1}{2} J z \sum_{i=0}^{q} x_i^2. \tag{3.2.4}$$

Since there are

$$\frac{N!}{N_1! N_2! \dots N_q!}$$

different ways of organizing the spins without changing the energy, there is an entropy equal to

$$S[\mathcal{C}] = k \log \left(\frac{N!}{N_1! N_2! \dots N_q!} \right).$$

By using the Stirling approximation for each of these terms $\log z! \simeq z \log z$ (with $z \gg 1$) together with the definition of the x_i's, we have

$$\frac{S[\mathcal{C}]}{N} \simeq -k \sum_{i=1}^{q} x_i \log x_i. \tag{3.2.5}$$

Hence, the free energy per spin reads

$$F(x_i) = -\sum_{i=1}^{q} \left[\frac{Jz}{2} x_i^2 - k\, x_i \log x_i \right].$$

This expression has to be minimized but taking into account the normalization condition (3.2.3). The latter constraint can be identically satisfied by using the parameterization

$$\begin{aligned} x_1 &= \frac{1}{q} \left[1 + (q-1)s \right], \\ x_i &= \frac{1}{q} (1-s), \quad i = 2, 3, \dots, q, \end{aligned}$$

with $0 \leq s \leq 1$. In the ferromagnetic case ($J > 0$) this position takes into account the possible symmetry breaking of the permutation group $\mathcal{S}_q$ in the low-temperature phase. Substituting the expressions for x_i in (3.2.4) and (3.2.5), we have

$$\begin{aligned} \frac{\beta}{N} \left[F(s) - F(0) \right] &= \frac{q-1}{2q} \mathcal{J} z\, s^2 - \frac{1 + (q-1)s}{q} \log \left[1 + (q-1)s \right] \\ &\quad - \frac{q-1}{q} (1-s) \log(1-s) \\ &\simeq -\frac{q-1}{2q} (q - \mathcal{J} z) s^2 + \frac{1}{6} (q-1)(q-2) s^3 + \cdots \end{aligned} \tag{3.2.6}$$

where $\mathcal{J} = \beta J$. Expanding this function in powers of s, one sees that for $q = 2$ the cubic term changes its sign: it is negative for $q < 2$ while positive for $q > 2$. This means that there could be a first-order phase transition. Let us consider the two cases separately:

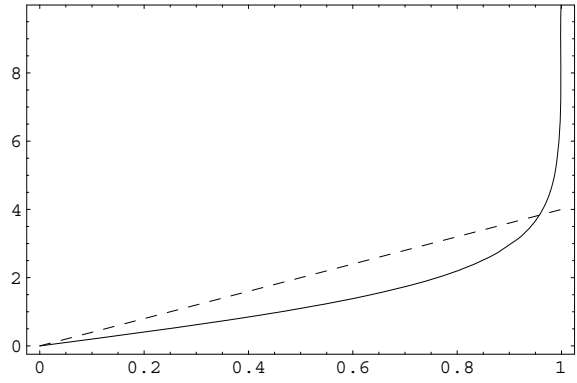


Fig. 3.2 *Graphical analysis of eqn (3.2.7).*

- $q < 2$. The minimum condition for the function in (3.2.6) is expressed by the equation

$$\mathcal{J} z\, s \;=\; \log\left[\frac{1+(q-1)s}{1-s}\right], \tag{3.2.7}$$

which always has $s = 0$ as a solution. For $\mathcal{J}z > q$ (where q is the derivative of the right-hand side at $s = 0$), there is however another solution $s \neq 0$, as can be easily seen graphically by plotting both terms of (3.2.7) as done in Fig. 3.2. The two solutions coincide when

$$\mathcal{J} \;=\; \mathcal{J}_c \;=\; \frac{q}{z}.$$

This condition identifies the critical value of the second-order phase transition that occurs for $q \leq 2$. Note that, for $q = 2$, we recover the critical temperature of the Ising model in the mean field approximation,[2] given by eqn (3.1.9). The plot of the free energy is shown in Fig. 3.3.

- $q > 2$. In this case we have a different situation: varying $\mathcal{J}$, there is a critical value at which the minimum of the free energy jumps from $s = 0$ to $s = s_c$, as shown in Fig. 3.4. This discontinuity is the fingerprint of a first order phase transition. In this case the critical values $\mathcal{J}_c$ and s_c are obtained by simultaneously solving the equations $F'(s) = 0$ and $F(s) = F(0)$, i.e.

$$z\,\mathcal{J}_c \;=\; \frac{2(q-1)}{q-2}\,\log(q-1),$$
$$s_c \;=\; \frac{q-2}{q-1}.$$

Computing the internal energy of the system, given by

$$U \;=\; -Jz\frac{q-1}{2q}\,s_{min}^2,$$

[2]To obtain the Ising model one has to make the substitution $\mathcal{J} \to 2\mathcal{J}$.

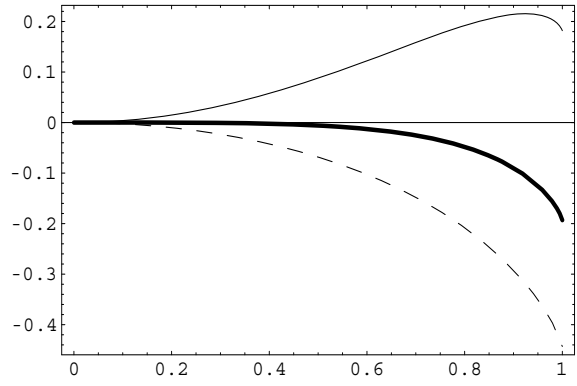


Fig. 3.3 *Plot of the free energy for $q < 2$: $\mathcal{J} > \mathcal{J}_c$ (upper curve), $\mathcal{J} = \mathcal{J}_c$ (black curve), and $\mathcal{J} < \mathcal{J}_c$ (lower curve).*

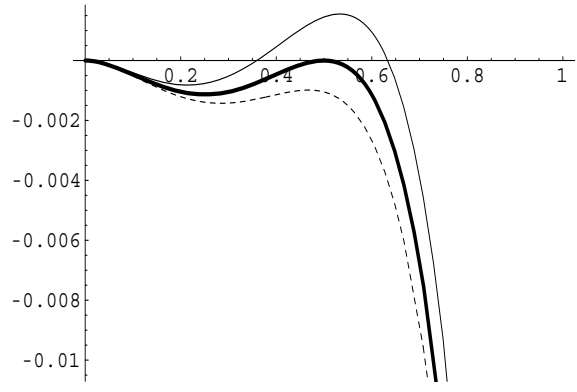


Fig. 3.4 *Plot of the free energy for $q > 2$: $\mathcal{J} > \mathcal{J}_c$ (upper curve), $\mathcal{J} = \mathcal{J}_c$ (black curve), and $\mathcal{J} > \mathcal{J}_c$ (lower curve).*

one sees that at $\mathcal{J} = \mathcal{J}_c$ this function has a jump that corresponds to a latent heat $\mathcal{L}$ per unit spin equal to

$$\mathcal{L} = Jz \frac{(q-2)^2}{2q(q-1)}.$$

3.3 Bethe–Peierls Approximation

The mean field approximation of the Ising model can be refined by adopting a formulation proposed by H.A. Bethe and R. Peierls. As for the Potts model, it is convenient to initially express the hamiltonian in terms of variables that take into account its degeneracy. For a given configuration of the spins, let us define

$$N_+ = \text{total number of spins with value } +1$$
$$N_- = \text{total number of spins with value } -1.$$

Each couple of nearest neighbor spins can only be one of the following types: (++), (−−), or (+−). Denote by N_{++}, N_{--}, and N_{+-} the total number of these pairs. These quantities are not independent: besides the obvious relationship

$$N_+ + N_- = N,$$

we also have

$$\begin{aligned} zN_+ &= 2N_{++} + N_{+-}; \\ zN_- &= 2N_{--} + N_{+-}, \end{aligned} \tag{3.3.1}$$

where z is the coordination number of the lattice. These identities can be proved as follows: once a site where the spin with value +1 is selected, draw a line that links this site to all the nearest neighbor ones, so that there are z lines. Repeating the same procedure for all those sites where the spins have value 1, we then have zN_+ lines. However, the pairs of next neighbor spins of the type (++) will have *two* lines while those of the type (+−) have *only one*, so that we reach the first formula in (3.3.1). Repeating the same argument for the spins with value −1, one obtains the second relationship. Eliminating N_{+-}, N_{--}, and N_- from the previous equations we have

$$\begin{aligned} N_{+-} &= zN_+ - 2N_{++}; \\ N_- &= N - N_+; \\ N_{--} &= \frac{z}{2}N + N_{++} - zN_+. \end{aligned}$$

Since

$$\sum_{\langle ij \rangle} \sigma_i \sigma_j = N_{++} + N_{--} - N_{+-} = 4N_{++} - 2zN_+ + \frac{z}{2}N,$$

$$\sum_i \sigma_i = N_+ - N_-,$$

the hamiltonian of the model can be expressed as

$$\begin{aligned} \mathcal{H} &= -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - B \sum_i \sigma_i \\ &= -4JN_{++} + 2(Jz - B)N_+ - \left(\frac{1}{2} Jz - B \right) N. \end{aligned} \tag{3.3.2}$$

The energy of the system depends only on the two quantities N_{++} and N_+ (the total number of the spins N is considered fixed) and therefore it is a degenerate function of the spin configurations. It is convenient to define an order parameter L (relative to the large-distance properties of the system) and an order parameter c (relative to its short distances):

$$\frac{N_+}{N} \equiv \frac{1}{2}(L+1) \quad (-1 \le L \le +1) \tag{3.3.3}$$

$$\frac{N_{++}}{\frac{1}{2}zN} \equiv \frac{1}{2}(c+1) \quad (-1 \le c \le 1). \tag{3.3.4}$$

In terms of these order parameters we have

$$\sum_{\langle ij \rangle} \sigma_i \sigma_j \;=\; \frac{1}{2} z N (2c - 2L + 1),$$
$$\sum_{i=1}^{N} \sigma_i \;=\; NL,$$

and the energy per unit spin can be written as

$$\frac{1}{N} E(L, c) \;=\; -\frac{1}{2} J z (2c - 2L + 1) - BL. \tag{3.3.5}$$

After these general considerations, let us discuss the Bethe–Peierls method, focusing on the case $B = 0$. Consider an elementary cell of the lattice, i.e. a site where the spin is in a state s together with its z neighbor sites. Denote by $P(s, n)$ the probability that n of these spins are in the state $+1$. If $s = +1$, then $P(s, n)$ is also equal to the probability to have n pairs $(++)$ and $(z-n)$ pairs $(+-)$. Vice versa, if $s = -1$, $P(s, n)$ is the probability to have n pairs $(+-)$ and $(n - z)$ of the type $(--)$. Given n, there are $\binom{z}{n} = z!/((n!(z-n)!)$ ways of selecting n among the z next neighbor spins. Let's assume that these probabilities can be written as

$$P(+1, n) \;=\; \frac{1}{q} \binom{z}{n} e^{\mathcal{J}(2n - z)} \rho^n; \tag{3.3.6}$$

$$P(-1, n) \;=\; \frac{1}{q} \binom{z}{n} e^{\mathcal{J}(z - 2n)} \rho^n, \tag{3.3.7}$$

where q is a normalization factor while ρ is a quantity that takes into account the overall effects of the lattice. While ρ will be determined later, q is obtained by imposing the normalization of the total probability

$$\sum_{n=0}^{z} [P(+1, n) + P(-1, n)] \;=\; 1,$$

namely

$$\begin{aligned} q &= \sum_{n=0}^{z} \left[\left(\rho e^{2\mathcal{J}} \right)^n e^{-\mathcal{J} z} + \left(\rho e^{-2\mathcal{J}} \right)^n e^{\mathcal{J} z} \right] \\ &= \left(e^{\mathcal{J}} + \rho e^{-\mathcal{J}} \right)^z + \left(\rho e^{\mathcal{J}} + e^{-\mathcal{J}} \right)^z . \end{aligned}$$

Using the order parameters L and c defined by eqns (3.3.3) and (3.3.4), and employing $P(+1, n)$, one has

$$\frac{N_+}{N} \;=\; \frac{1}{2}(L + 1) \;=\; \sum_{n=0}^{z} P(+1, n) \;=\; \frac{1}{q} \left(e^{\mathcal{J}} + \rho e^{-\mathcal{J}} \right)^z , \tag{3.3.8}$$

$$\frac{N_{++}}{\frac{1}{2}zN} = \frac{1}{2}(c+1) = \frac{1}{z}\sum_{n=0}^{z} nP(+1,n) = \frac{\rho}{q}\, e^{\mathcal{J}} \left(e^{-\mathcal{J}} + \rho e^{\mathcal{J}}\right)^{z-1}. \tag{3.3.9}$$

We can now proceed to directly compute the magnetization. Note that

$$\sum_{n=0}^{z} P(+1,n) = \begin{cases} \text{probability to find a spin} \\ \text{with value } +1 \text{ in the center} \end{cases}$$

$$\frac{1}{z}\sum_{n=0}^{z} n\,[P(+1,n) + P(-1,n)] = \begin{cases} \text{probability to find a spin with value} \\ +1 \text{ among the next neighbor sites.} \end{cases}$$

To have a consistent formulation, these two probabilities must be equal. Using (3.3.6) and (3.3.7), one arrives at the equation for the variable ρ

$$\rho = \left(\frac{1+\rho e^{2\mathcal{J}}}{\rho + e^{2\mathcal{J}}}\right)^{z-1}. \tag{3.3.10}$$

Assuming we have solved this equation and found the value of ρ, $\langle L\rangle$ and $\langle c\rangle$ can be obtained through eqns (3.3.8) and (3.3.9):

$$\langle L\rangle = \frac{\rho^x - 1}{\rho^x + 1}, \tag{3.3.11}$$

$$\langle c\rangle = \frac{2\rho^2}{(1+\rho e^{-2\mathcal{J}})(1+\rho^x)} - 1, \tag{3.3.12}$$

where $x \equiv z/z - 1$. The internal energy is given by

$$\frac{1}{N}U(T) = -\frac{1}{2}Jz\left(2\langle c\rangle - 2\langle L\rangle + 1\right), \tag{3.3.13}$$

whereas the spontaneous magnetization is expressed by

$$\frac{1}{N}\left\langle \sum_{i=1}^{N} \sigma_i \right\rangle = \langle L\rangle. \tag{3.3.14}$$

It is now necessary to solve eqn (3.3.10). Note that this equation has the following properties:

1. $\rho = 1$ is always a solution;
2. if ρ_0 is a solution, then also $1/\rho_0$ is a solution;
3. interchanging ρ with $1/\rho$ is equivalent to interchange $\langle L\rangle \to -\langle L\rangle$;
4. $\rho = 1$ corresponds to $\langle L\rangle = 0$, while $\rho = \infty$ corresponds to $\langle L\rangle = 1$.

To find the solution of eqn (3.3.10), it is useful to use a graphical method, similarly to the mean field solution: one plots the right- and the left-hand side functions of eqn (3.3.10) and determines the points of their intersection, as shown in Fig. 3.5.

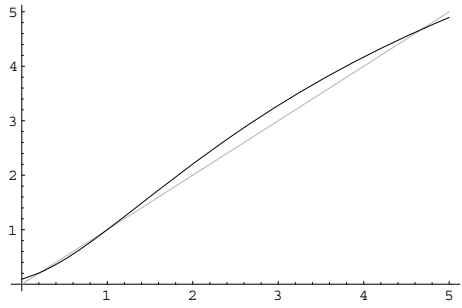


Fig. 3.5 *Graphical solution of eqn (3.3.10).*

An important quantity is the value of the derivative of the function on the right-hand side, computed at $\rho = 1$

$$g = \frac{(z-1)(e^{4\mathcal{J}}-1)}{(1+e^{2\mathcal{J}})^2}. \tag{3.3.15}$$

In fact, if $g < 1$, the only solution consists of $\rho = 1$. Vice versa, if $g > 1$, there are three solutions, $\rho = 1$, ρ_0, and $1/\rho_0$. Excluding the solution $\rho = 1$ (which corresponds to $\langle L \rangle = 0$) and $1/\rho_0$ (which is equivalent to exchanging the spins $+1$ with those of -1), the only physically relevant solution is given by ρ_0. In this approach, the critical temperature is given precisely by the condition $g = 1$, namely

$$kT_c = \frac{2J}{\ln\left[z/(z-2)\right]}. \tag{3.3.16}$$

For $T > T_c$ we have

$$\begin{aligned} \rho &= 1; \\ \langle L \rangle &= 0; \\ \langle c \rangle &= \frac{1}{2(1+e^{-2\mathcal{J}})}. \end{aligned} \tag{3.3.17}$$

For $T < T_c$, we have instead $\rho > 1$ and $\langle L \rangle > 0$, i.e. there is a spontaneous magnetization in the system.

The expression (3.3.16) of the critical temperature predicted by the Bethe–Peierls approximation correctly predicts that, in one dimension where $z = 2$, $T_c = 0$. In two dimensions, for a square lattice ($z = 4$), it provides the estimate $kT_c/J = 2/\ln 2 = 2.885$, which is smaller than the one obtained in the mean field approximation $kT_c/J = 4$ but still higher than the exact value $kT_c/J = 2/\ln(1+\sqrt{2}) = 2.269$ which we will determine in Chapter 4.

3.4 The Gaussian Model

In the Ising model, the computation of the partition function is based on the sums of the discrete variables $\sigma_i = \pm 1$. Notice that such a discrete sum can be written as an

integral on the entire real axis by using the Dirac delta function[3]

$$\sum_{\sigma_i=\pm 1} [\cdots] \;=\; \int_{-\infty}^{+\infty} d\sigma_i \, \delta(\sigma_i^2 - 1) \; [\cdots] \, .$$

Using the properties of $\delta(x)$, the Ising model can then be regarded as a statistical model where the spins assume all continuous values of the real axis but with a probability density given by

$$P_I(\sigma_i) \;=\; \frac{1}{2} \; [\delta(\sigma_i - 1) + \delta(\sigma_i + 1)] \, . \tag{3.4.1}$$

With the above notation, the sum on the configurations of a single spin assumes the form

$$\sum_{\sigma_i=\pm 1} [\cdots] \;=\; \int_{-\infty}^{+\infty} d\sigma_i \, P_I(\sigma_i) \; [\cdots] \, ,$$

and the usual mean values of the Ising model are given by

$$\begin{aligned} \langle \sigma_i \rangle \;&\equiv\; \int_{-\infty}^{+\infty} d\sigma_i \, P_I(\sigma_i) \, \sigma_i \;=\; 0, \\ \langle \sigma_i^2 \rangle \;&\equiv\; \int_{-\infty}^{+\infty} d\sigma_i \, P_I(\sigma_i) \, \sigma_i^2 \;=\; 1. \end{aligned} \tag{3.4.2}$$

We can now conceive to approximate the Ising model by substituting the probability density $P_I(\sigma_i)$ – given by eqn (3.4.1) – with another probability density $P(\sigma_i)$ that shares the mean values $\langle \sigma_i \rangle$ and $\langle \sigma_i^2 \rangle$ of (3.4.2). A function with such a property is, for instance, the gaussian curve[4] (see Fig. 3.6)

$$P_G(\sigma) \;=\; \sqrt{\frac{1}{2\pi}} \, \exp\left[-\frac{\sigma^2}{2} \right] . \tag{3.4.3}$$

The spin model defined by this new probability density is known as *the gaussian model*.

Since thermal averages are computed according to the formula

$$\langle A \rangle \;=\; \frac{1}{Z} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{i=1}^{N} P(\sigma_i) \; A \, e^{-\beta \mathcal{H}} \, d\sigma_1 \, \cdots \, d\sigma_N ,$$

[3]The Dirac delta function $\delta(x)$, with x real, satisfies the properties:

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ +\infty & x = 0 \end{cases}$$

with $\int_{-\infty}^{+\infty} \delta(x) = 1$. Moreover, $\delta[f(x)] = \sum_i \frac{1}{|f'(x_i)|} \, \delta(x - x_i)$, where x_i are the roots of the equation $f(x) = 0$.

[4]Although $P_I(\sigma)$ and $P_G(\sigma)$ give rise to the same mean values of eqn (3.4.2), they nevertheless differ for what concerns the mean values of the higher powers of the spins. For $P_I(\sigma)$ we have $\langle \sigma^{2n} \rangle = 1$, while for $P_G(\sigma)$, $\langle \sigma^{2n} \rangle = [1 \cdot 3 \cdot 5 \cdots (2n-1)]$.

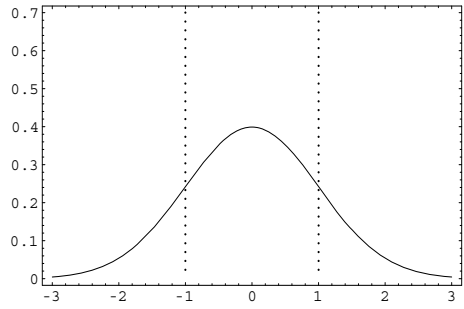


Fig. 3.6 *Probability density $P(\sigma)$: from the Ising model to the gaussian model.*

where

$$Z \,=\, \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_{i=1}^{N} P(\sigma_i)\;\, e^{-\beta \mathcal{H}}\, d\sigma_1 \,\cdots\, d\sigma_N,$$

it is obvious that the presence of $P(\sigma)$ in the sum over the states can be equivalently interpreted as a new term in the hamiltonian, so that

$$\mathcal{H} \,\longrightarrow\, \mathcal{H}' \,=\, \mathcal{H} - \frac{1}{\beta} \sum_{i=1}^{N} \log[P(\sigma_i)].$$

The thermal averages computed with the new Boltzmann factor

$$\langle A \rangle \,=\, \frac{1}{Z} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} A\, e^{-\beta \mathcal{H}'}\, d\sigma_1 \,\cdots\, d\sigma_N,$$

clearly coincide with the previous ones. Hence, we can reformulate the gaussian model as a system where the spins assume a continuous set of values, with an interaction given, up to a constant, by the hamiltonian

$$\mathcal{H} \,=\, \frac{1}{2\beta} \sum_{i=1}^{N} \sigma_i^2 - J \sum_{\langle ij \rangle} \sigma_i\, \sigma_j - B \sum_{i=1}^{N} \sigma_i. \tag{3.4.4}$$

Let us now proceed to the computation of its partition function

$$Z_N \,=\, \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2} \sum_{i=1}^{N} \sigma_i^2 + \mathcal{J} \sum_{\langle kl \rangle} \sigma_k \sigma_l + \mathcal{B} \sum_{l} \sigma_l \right] d\sigma_1 \cdots d\sigma_N,$$

where $\mathcal{J} = J/kT$ and $\mathcal{B} = B/kT$. To simplify the formulas, it is convenient to introduce a matrix notation: let σ be an N-component vector $\sigma = (\sigma_1, \sigma_2 \ldots \sigma_N)$, and let $\mathbf{V}$ be a $N \times N$ matrix defined by

$$\sigma^T\, \mathbf{V}\, \sigma \,=\, \frac{1}{2} \sum_{l} \sigma_l^2 - \mathcal{J} \sum_{\langle kl \rangle} \sigma_k \sigma_l.$$

Moreover, let $\mathcal{B}$ be an N-dimensional vector, with all its components equal to $\mathcal{B}$. In terms of these new notations, the partition function is written as

$$Z_N \;=\; \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp\left[-\sigma^T \mathbf{V}\sigma + \mathcal{B}^T\sigma\right]\, d\sigma_1 \cdots d\sigma_N.$$

The integral over the variables σ_j is gaussian and can be performed using the formula[5]

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp\left[-x^T \mathbf{V} x + h^T x\right]\, dx_1 \cdots dx_N = (\pi)^{N/2}\, [\det \mathbf{V}]^{-\frac{1}{2}} \exp\left[\frac{1}{4} h^T \mathbf{V}^{-1} h\right] \quad (3.4.5)$$

so that we arrive at

$$Z_n \;=\; \pi^{\frac{N}{2}}\, [\,\det \mathbf{V}]^{-\frac{1}{2}}\, \exp\left[\frac{1}{4}\mathcal{B}^{\mathbf{T}}\mathbf{V}^{-\mathbf{1}}\mathcal{B}\right].$$

Cyclic matrix. It is necessary, however, to verify if there are eigenvalues of the matrix V with a real part that is either zero or negative. Their explicit expression clearly depends on the nature of the matrix V, namely from the underlying lattice structure of the gaussian model. Let us consider, for simplicity, a d-dimensional cubic lattice of length L in all its directions, with periodic boundary conditions. In this case, $N = L^d$. For the (discrete) translation invariance of the lattice, the matrix elements of V depends only on the difference of its indices

$$V_{\vec{i},\vec{j}} \;=\; V(\vec{i} - \vec{j}).$$

For a cubic lattice, the only components that are different from zero are those for which $\vec{i} - \vec{j} = 0$ and

$$\vec{i} - \vec{j} \;=\; \begin{cases} (\pm 1, 0, 0, 0, \ldots) \\ (0, \pm 1, 0, 0, \ldots) \\ (0, 0, \pm 1, 0, \ldots) \\ (0, 0, 0, \pm 1, \ldots) \\ \cdots \end{cases}$$

The periodic boundary conditions put the additional constraints

$$V(i + L, j) \;=\; V(i, j + L) \;=\; V(i, j).$$

A matrix that satisfies all the above properties is called a *cyclic matrix* and its eigenvalues can be easily determined by using the Fourier series. The result is

$$\lambda(\omega_1, \ldots, \omega_d) \;=\; \frac{1}{2} - \mathcal{J}(\cos\omega_1 + \ldots + \cos\omega_d), \quad (3.4.6)$$

[5]The validity of this formula relies on the condition that all the eigenvalues of V are positive. As we will see, when this condition is not satisfied, the system undergoes a phase transition.

where each frequency ω_j can take one of the L possible values $0, 2\pi/L, 4\pi/L, \ldots, 2\pi(L-1)/L$. From eqn (3.4.6) it follows that the eigenvalues have a positive real part only if

$$|\mathcal{J}| < \frac{1}{2d}. \tag{3.4.7}$$

This condition determines the range of validity of the gaussian model and identifies the critical temperature of the model, given by

$$\frac{J}{kT_c} = \frac{1}{2d}. \tag{3.4.8}$$

It is easy to understand the origin of this critical point by directly analyzing the hamiltonian (3.4.4). Note that the coupling constant of the first term explicitly depends on the temperature through the parameter $1/\beta$. At high temperatures, i.e. $\beta \to 0$, the minimum of the hamiltonian is reached by the configuration in which all spins have a zero value, $\sigma_i = 0$. In the opposite limit, $\beta \to \infty$, i.e. in the low-temperature phase, the second term may prevail over the first one. In such a case, the hamiltonian density per unit spin $\mathcal{H}/N$ can be made arbitrarily negative by allowing the spins to align with each other and grow without any bound on their modulus. Hence, in the low-temperature phase, the energy of the model is not bounded from below. The two physical pictures, so qualitatively different, obtained in the high- and low-temperature phases, clearly signal the existence of a phase transition, with the critical temperature given in (3.4.8). The low-temperature phase of the model is however pathological: as a matter of fact, the model is only defined for $T > T_c$ and therefore it only has a high-temperature phase. In the next section we will see how to get around this difficulty by posing a bound on the higher values of the spins.

Transfer matrix in 1-D. The analysis done is completely general and applies to arbitrary lattices. However, it is an interesting exercise to solve the one-dimensional gaussian model by using the transfer matrix. To this purpose, consider the hamiltonian of the one-dimensional gaussian model

$$\mathcal{H} = \frac{1}{2\beta}\sum_{i=1}^{N}\sigma_i^2 - J\sum_{i=1}^{N}\sigma_i\sigma_{i+1}.$$

The transfer matrix T of the model has a set of continuous indices: denoting by x and y the values of the spin of two neighbor sites, we have

$$\langle x\,|T|\,y\rangle = T(x,y) = \exp\left[-\frac{1}{4}(x^2+y^2) + \mathcal{J}xy\right].$$

To compute the partition function, we have to diagonalize this matrix by solving the integral equation

$$\int_{-\infty}^{+\infty} T(x,y)\,\psi(y)\,dy \;=\; \lambda\,\psi(x). \tag{3.4.9}$$

Note that the norm of the integral operator is finite only if

$$|\mathcal{J}| \;<\; \frac{1}{2}. \tag{3.4.10}$$

In fact, it is only in this interval that the kernel of the integral operator is square integrable

$$\begin{aligned} ||T||^2 &\equiv \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \langle x|T|y\rangle\,\langle y|T|x\rangle\,dx\,dy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}(x^2+y^2)+2\mathcal{J}xy\right]\,dx\,dy \\ &= \frac{2\pi}{\sqrt{1-4\mathcal{J}^2}}. \end{aligned} \tag{3.4.11}$$

The integral operator is symmetric in its indices $T(x,y) = T(y,x)$ and, for its reality, it is then a hermitian operator. The general theory of integral equations[6] tells us that the eigenvalues λ_n of such an operator are real and discrete, while the corresponding eigenfuctions ψ_n form a complete set of orthonormal functions. The operator T admits the spectral decomposition

$$T(x,y) \;=\; \sum_{n=0}^{\infty} \lambda_n\,\psi_n(x)\psi_n(y), \tag{3.4.12}$$

and its eigenvalues satisfy the identity

$$||T||^2 \;=\; \sum_{n=0}^{\infty} \lambda_n^2. \tag{3.4.13}$$

To determine the spectrum of the integral equation (3.4.9), we can use a very elegant algebraic method that is similar to the one used to find the spectrum of the quantum one-dimensional harmonic oscillator.[7] Let us introduce the differential operator $\mathbf{A}$

$$\mathbf{A} \;=\; \frac{1}{\sqrt{2}}\left(ux + u^{-1}\,\frac{d}{dx}\right), \tag{3.4.14}$$

[6]See, for instance, D. Porter and D. Stirling, *Integral Equations*, Cambridge University Press, Cambridge, 1990.

[7]See, for instance, C. Cohen–Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, John Wiley & Sons, New York, 1977.

and its hermitian conjugate[8] $\mathbf{A}^\dagger$

$$\mathbf{A}^\dagger \;=\; \frac{1}{\sqrt{2}}\left(ux - u^{-1}\,\frac{d}{dx}\right). \tag{3.4.15}$$

These operators satisfy the commutation relation

$$\left[\mathbf{A}, \mathbf{A}^\dagger\right] \;=\; 1. \tag{3.4.16}$$

Choosing

$$u^2 \equiv u_*^2 \;=\; \frac{1}{2}\sqrt{1-4\mathcal{J}^2}, \tag{3.4.17}$$

the differential operator $\mathbf{A}$ satisfies the operator equation

$$\mathbf{A}\,T \;=\; \xi\,T\,\mathbf{A}, \tag{3.4.18}$$

with the constant ξ given by

$$\xi \;=\; \frac{1-2u_*^2}{2\mathcal{J}} \;=\; \frac{1-\sqrt{1-4\mathcal{J}^2}}{2\mathcal{J}}. \tag{3.4.19}$$

To prove (3.4.18), let us initially apply this relation to an arbitrary function $\psi(x)$

$$\int_{-\infty}^{\infty}\left(ux + u^{-1}\,\frac{d}{dx}\right) T(x,y)\,\psi(y)\,dy \;=\; \xi\int_{-\infty}^{\infty} T(x,y)\,\left(uy + u^{-1}\,\frac{d}{dy}\right)\psi(y)\,dy.$$

Integrating by parts the last term on the right-hand side, and taking into account the arbitrariness of the function $\psi(x)$, we get

$$\left(ux + u^{-1}\,\frac{d}{dx}\right) T(x,y) \;=\; \xi\left(uy - u^{-1}\,\frac{d}{dy}\right) T(x,y).$$

Equating the terms that are proportional to x and y in both terms, we obtain

$$\begin{aligned}\left(u - \frac{1}{2u}\right) &= -\xi\,\frac{\mathcal{J}}{u},\\ \xi\left(u + \frac{1}{2u}\right) &= \frac{\mathcal{J}}{u}.\end{aligned} \tag{3.4.20}$$

Hence, we arrive at eqn (3.4.19) with u_*, given by (3.4.17), a value that comes from the consistency condition of the systems of equations (3.4.20).

Taking the hermitian conjugate of (3.4.18), the operator T also satisfies

$$T\,\mathbf{A}^\dagger \;=\; \xi\,\mathbf{A}^\dagger\,T. \tag{3.4.21}$$

[8]It is important to recall that the differential operator $\frac{d}{dx}$ is antihermitian.

The two functional equations (3.4.18) and (3.4.21) satisfied by T enable us to reach some important consequences about its spectrum. Suppose we have identified a real eigenfunction $\psi(x)$ of this operator, with eigenvalue λ. Applying to $\psi(x)$ the operator $\mathbf{A}^\dagger$, we see that $\psi'(x) = \mathbf{A}^\dagger \psi(x)$ is also an eigenfunction of T but with the eigenvalue $\lambda' = \xi\,\lambda$. Hence, the iterated application of $\mathbf{A}^\dagger$ to the eigenfunction $\psi(x)$ gives rise to the sequence of eigenfunctions $\left(\mathbf{A}^\dagger\right)^n \psi(x)$. Since for $|\mathcal{J}| < 1/2$ we have $\xi < 1$, we obtain the sequence of decreasing eigenvalues

$$\lambda > \lambda\xi > \lambda\xi^2 > , \ldots > \lambda\xi^n. \tag{3.4.22}$$

Vice versa, making use of (3.4.18), the iterated application of the operator $\mathbf{A}$ to the eigenfunction $\psi(x)$ also generates a sequence of eigenfunctions $\tilde{\psi}_n = \mathbf{A}^n\psi$, but this time with a sequence of *increasing* eigenvalues

$$\lambda < \lambda\xi^{-1} < \lambda\xi^{-2} < \ldots < \lambda\xi^{-n}. \tag{3.4.23}$$

Since the T is bounded in the interval (3.4.10), a maximum eigenvalue $\lambda_{\max} \equiv \lambda_0$ must necessarily exist. This implies that the sequence (3.4.23) must stop and the eigenfunction that corresponds to the maximum eigenvalue λ_0 satisfies the equation $\mathbf{A}\,\psi_0(x) = 0$, i.e.

$$\left(u_* x + \frac{1}{u_*}\frac{d}{dx}\right)\psi_0(x) = 0. \tag{3.4.24}$$

Therefore

$$\psi_0(x) = A_0 \exp\left[-\frac{u_*^2\,x^2}{2}\right],$$

where the constant $A_0 = \sqrt{\frac{u_*^2}{\pi}}$ is fixed by the normalization condition

$$\int_{-\infty}^{+\infty} \psi_0^2(x)\,dx = 1.$$

We could directly compute the maximum eigevalue λ_0 by substituting $\psi_0(x)$ in the integral equation. However, in order to control all the results obtained above, it is convenient to proceed in a more general way. Note that the application of the operator T to a generic gaussian function $g(x) = A\exp[-\lambda^2 x^2/2]$ produces another gaussian function

$$\begin{aligned} T\,g(x) &= A\int_{-\infty}^{+\infty} dy\,\exp\left[-\frac{1}{4}(x^2+y^2) + \mathcal{J}xy\right]\exp\left[-\frac{\lambda^2}{2}y^2\right] \\ &= A_0\int_{-\infty}^{+\infty} dy\,\exp\left[-\frac{1}{4}x^2 - \frac{1}{4}(1+2\lambda^2)y^2) + \mathcal{J}xy\right] \\ &= \tilde{A}\exp\left[-\frac{\tilde{\lambda}^2}{2}x^2\right], \end{aligned}$$

with a new exponent $\tilde{\lambda}^2$, given by

$$\tilde{\lambda}^2 \;=\; \frac{1}{2} - \frac{\mathcal{J}^2}{\lambda^2 + 1/2},$$

and a new normalization constant

$$\tilde{A}_0 \;=\; A_0 \sqrt{\frac{2\pi}{\lambda^2 + 1/2}}. \tag{3.4.25}$$

If the gaussian $g(x)$ is an eigenfunction of T, $\tilde{\lambda}^2$ should be obviously equal to the previous one λ^2 and we get the equation

$$\lambda^2 \;=\; \frac{1}{2} - \frac{\mathcal{J}^2}{\lambda^2 + 1/2}. \tag{3.4.26}$$

The solution is given by

$$\lambda^2 \;=\; u_*^2 \;=\; \frac{1}{2}\sqrt{1 - 4\mathcal{J}^2},$$

and coincides with the condition (3.4.17), previously obtained for the eigenfunction $\psi_0(x)$.

Substituting in (3.4.25) the value of u_*^2, we obtain the maximum eigenvalue

$$\lambda_0 \;=\; \sqrt{\frac{2\pi}{u_*^2 + 1/2}} \;=\; \sqrt{\frac{4\pi}{1 + \sqrt{1 - 4\mathcal{J}^2}}}. \tag{3.4.27}$$

The sequence of eigenvalues is now given by eqn (3.4.22), with $\lambda = \lambda_0$. In particular, it is easy to check the validity of the identity (3.4.13): in fact, for the right-hand side of this equation we have

$$\sum_{k=0}^{\infty} \lambda_k^2 \;=\; \lambda_0^2 \sum_{k=0}^{\infty} \xi^{2k} \;=\; \frac{\lambda_0^2}{1 - \xi^2},$$

and, substituting the two expressions (3.4.27) and (3.4.19), we precisely obtain the norm of the operator T, expressed by eqn (3.4.11).

Once the maximum eigenvalue is known, the free energy per unit spin of the model is given by

$$\beta F \;=\; -\lim_{N\to\infty} \frac{1}{N} \log Z_N \;=\; -\log \lambda_0(\mathcal{J}). \tag{3.4.28}$$

Note, that when the temperature tends to its critical value

$$\mathcal{J}_c \;\to\; \frac{1}{2},$$

correspondingly $\xi \to 1$. This implies a collapse of all eigenvalues of the transfer matrix. Since the transfer matrix of a classical statistical system can be associated to a hamiltonian $\mathbf{H}$ of a quantum system by means of the formula

$$T \equiv e^{-a\mathbf{H}},$$

the collapse of all eigenvalues of T corresponds to a very singular point of degeneracy of the quantum hamiltonian $\mathbf{H}$. At $\mathcal{J} = \mathcal{J}_c$ we have a significant mixing of all eigenstates of $\mathbf{H}$, with a drastic and discontinous change of the fundamental state of the system: the systems then undergoes a phase transition.

Using the spectral decomposition[9] of the operator T, eqn (3.4.12), is easy to see that the quantum hamiltonian $\mathbf{H}$ assumes the form

$$\mathbf{H} = -\frac{1}{a}\left[\left(\log\frac{1-\sqrt{1-4\mathcal{J}^2}}{2\mathcal{J}}\right)\mathbf{A}^\dagger\,\mathbf{A} + \frac{1}{2}\log\left(\frac{4\pi}{1+\sqrt{1-\mathcal{J}^2}}\right)\right]. \qquad (3.4.29)$$

In the limit $\mathcal{J} \to \mathcal{J}_c$, the coefficient in front of $\mathbf{A}^\dagger\,\mathbf{A}$ vanishes and, as expected, there is an infinite degeneration of the eigenvalues of $\mathbf{H}$.

To cure the pathological features of the low-temperature phase of the gaussian model, T.H. Berlin and M. Kac proposed a more sophisticated version of the model, the so-called *spherical model*. This model has the additional advantage of being more similar to the Ising model than the gaussian model itself.

3.5 The Spherical Model

The spherical model, introduced and solved by Berlin and Kac in 1952, consists of an interesting variant of the Ising model, or better, of the gaussian model. Like the last one, the N spins of the spherical model interact with their first neighbors and an eventual external field, and assume all real values. However, they are subject to the condition

$$\sum_{j=1}^{N}\sigma_j^2 = N. \qquad (3.5.1)$$

When there is homogeneity in the spins, this condition is equivalent to $\langle\sigma_i^2\rangle \simeq 1$, just like in the original Ising model. However it is obvious that there is a difference between these two models: in fact, while in the Ising model the sum over the spin configurations corresponds to a sum over all the vertices of an N-dimensional hypercube, in the spherical model this sum is replaced by an integral over the N-dimensional spherical surface that passes through them.

[9]The normalized eigenfunctions $\psi_n(x)$ are given by $\psi_n(x) = \frac{1}{\sqrt{n!}}\left(\mathbf{A}^\dagger\right)^n\psi_0(x)$, and we have $\langle\psi_m \mid \mathbf{A}^\dagger\,\mathbf{A} \mid \psi_n\rangle = n\,\delta_{nm}$.

Besides its intrinsic interest,[10] one could however doubt its physical content, inasmuch as the condition (3.5.1) depends on the dimension N of the system. This is in fact equivalent to having an interaction between all the spins. This objection has found, however, a valid answer in the equivalence (shown by H.E. Stanley in 1968) between the spherical model and a spin model with $O(n)$ symmetry and nearest neighbor interactions, in the limit in which $n \to \infty$. Namely, Stanley has proved that the model with Hamiltonian

$$\mathcal{H} \;=\; -J \sum_{\langle ij \rangle} \vec{\sigma}_i \cdot \vec{\sigma}_j,$$

where each spin is an n-dimensional vector satisfying

$$| \, \vec{\sigma}_i \, |^2 = \; n$$

in the limit $n \to \infty$, is equivalent to the spherical model.[11]

Let's now compute the partition function of the model and its equation of state. Although not particularly demanding, the following calculations require, however, a certain mathematical skill. The partition function is given by the multidimensional integral

$$Z_N \;=\; \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\sigma_1 \cdots d\sigma_N \delta \left(N - \sum \sigma_j^2 \right) \exp \left[\mathcal{J} \sum_{\langle kl \rangle} \sigma_k \sigma_l + \mathcal{B} \sum_l \sigma_l \right], \tag{3.5.2}$$

with $\mathcal{J} = J/kT$ and $\mathcal{B} = B/kT$. The constraint (3.5.1) is enforced by the Dirac delta function. Using

$$\delta(x) \;=\; \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{isx} \, ds,$$

and noting that we can insert in the integral the term

$$e^{\mu \left(N - \sum_l \sigma_l^2 \right)}$$

(which is equal to 1, thanks to eqn (3.5.1)), the partition function can be rewritten as

$$Z_N = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\sigma_1 \cdots d\sigma_N \int_{-\infty}^{+\infty} ds \tag{3.5.3}$$
$$\exp \left[\mathcal{J} \sum_{\langle kl \rangle} \sigma_k \sigma_l + \mathcal{B} \sum_l \sigma_l + (\mu + is)(N - \sum_l \sigma_l^2) \right].$$

[10] As we will see below it is exactly solvable, with a different behavior with respect to the mean field solution for $d \geq 3$, while for $d = 1$ and $d = 2$ it does not have a phase transition.

[11] Note that the model considered by Stanley differs from the one discussed in Section 2.6 since the modulus of the spin is $n^{1/2}$ instead of 1.

It is convenient to adopt the compact notation of the previous section. Let us define an $N \times N$ matrix V by means of

$$\sigma^T \mathbf{V} \sigma \;=\; (\mu + is) \sum_l \sigma_l^2 - \mathcal{J} \sum_{\langle kl \rangle} \sigma_k \sigma_l .$$

Hence

$$Z_N \;=\; \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} d\sigma_1 \cdots d\sigma_N \int_{-\infty}^{+\infty} ds \, \exp\left[-\sigma^T \mathbf{V} \sigma + \mathcal{B}^T \sigma + (\mu + is) N \right] . \tag{3.5.4}$$

We can choose a sufficiently large value of the arbitrary constant μ in such way that all the eigenvalues of the matrix V have a positive real part (we will specify this condition in more detail ahead, see eqn (3.5.7)). Under these conditions, we can exchange the integration order over the variables σ_j and s: the integration over the variable σ_j is gaussian and can be carried out thanks to the formula (3.4.5), so that

$$Z_n \;=\; \frac{1}{2} \pi^{\frac{N}{2}-1} \int_{-\infty}^{+\infty} ds \, \left[\det \mathbf{V} \right]^{-\frac{1}{2}} \, \exp\left[(\mu + is) N + \frac{1}{4} \mathcal{B}^T \mathbf{V}^{-1} \mathcal{B} \right] . \tag{3.5.5}$$

To proceed further, it is necessary to specify the nature of the matrix V. For simplicity, also in this case we choose a cubic lattice with $N = L^d$ and with periodic conditions along all directions. V is therefore a cyclic matrix and we can repeat the main steps of the analysis of the previous section. The eigenvalues of V are obtained in terms of the Fourier series, with the final result given by

$$\lambda(\omega_1, \ldots, \omega_d) \;=\; \mu + is - \mathcal{J}(\cos\omega_1 + \ldots + \cos\omega_d), \tag{3.5.6}$$

where each frequency ω_j assumes the L values $0, 2\pi/L, 4\pi/L, \ldots, 2\pi(L-1)/L$. From (3.5.6) it is easy to see that the real part is positive if the constant μ satisfies

$$\mu > \mathcal{J} d. \tag{3.5.7}$$

Since the determinant of a matrix is given by the product of its eigenvalues, we have

$$[\, \det \mathbf{V}] \;=\; \exp\left[\ln \det \mathbf{V} \right] = \exp\left[\sum_{\omega_1} \ldots \sum_{\omega_d} \ln \lambda(\omega_1, \ldots, \omega_d) \right] .$$

In the thermodynamic limit $L \to \infty$, the eigenvalues become dense and the sum over them can be converted into an integral

$$\ln \det \mathbf{V} \;=\; N \left[\ln \mathcal{J} + g(z) \right] ,$$

where we have defined

$$z \;=\; (\mu + is)/\mathcal{J},$$

and

$$g(z) \;=\; \frac{1}{(2\pi)^d} \int_0^{2\pi} \ldots \int_0^{2\pi} d\omega_1 \ldots d\omega_d \, \ln \left[z - \sum_{j=1}^{d} \cos\omega_j \right] . \tag{3.5.8}$$

The function $g(z)$ is analytic when $\mathrm{Re}\, z > d$ and has a singular point at $z = d$.

We can take further advantage of the cyclic nature of the matrix V to show that the constant vector $\mathcal{B}$ is the eigenvector of V corresponding to its minimum eigenvalue $\mu + is - \mathcal{J}d = \mathcal{J}(z-d)$. Hence

$$\mathcal{B}^T \mathbf{V}^{-1} \mathcal{B} = \mathcal{B}^T \frac{1}{\mathcal{J}(z-d)} \mathcal{B} = \frac{N\mathcal{B}^2}{\mathcal{J}(z-d)}.$$

Putting together the last formulas and making a change of variable from s to z, the partition function can be expressed as

$$Z_N = \left(\frac{\mathcal{J}}{2\pi i}\right)\left(\frac{\pi}{\mathcal{J}}\right)^{\frac{N}{2}} \int_{c-i^\infty}^{c+i^\infty} dz \exp[N\phi(z)], \tag{3.5.9}$$

where the function $\phi(z)$ is defined by

$$\phi(z) = \mathcal{J}z - \frac{1}{2}g(z) + \frac{\mathcal{B}^2}{4\mathcal{J}(z-d)}. \tag{3.5.10}$$

For the condition on the eigenvalues of V, the integration contour γ is chosen as in Fig. 3.7, with $c = (\mu - \mathcal{J}d)/\mathcal{J} > 0$. Since $\phi(z)$ is an analytic function in the semi-plane $\mathrm{Re}\, z > d$, the value of the integral in eqn (3.5.9) does not depend on the value of the constant c, as long as this constant is positive. In the thermodynamic limit $N \to \infty$, Z_N can be estimated by using the saddle point method, discussed in Appendix 3A. Consider the behavior of $\phi(z)$ when z is real and positive, in the case in which $\mathcal{J} > 0$ and $\mathcal{B} \neq 0$. It is easy to see that the function diverges both for $z \to d$ and $z \to \infty$, assuming positive values in between. Therefore the function $\phi(z)$ must have a *minimum* at some positive point z_0 and since $\phi''(z) > 0$, this is the only minimum. Let us choose the constant c to be *exactly* equal to z_0. Since $\phi(z)$ is an analytic function, along the direction of the new path of integration γ it will present a *maximum* at $z = z_0$. Such a maximum rules the behavior of the integral in the limit $N \to \infty$ and therefore the free energy is given by

$$-F/kT = \lim_{N\to\infty} \frac{1}{N} \ln Z_N = \frac{1}{2} \ln\left(\frac{\pi}{\mathcal{J}}\right) + \phi(z_0). \tag{3.5.11}$$

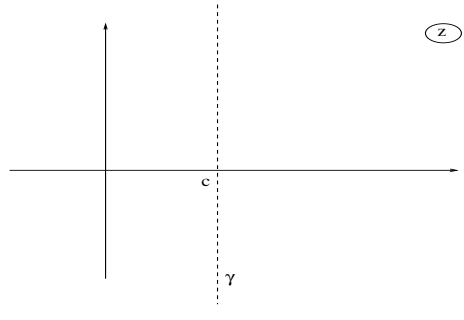


Fig. 3.7 *Contour of integration in the complex plane.*

The value z_0 is determined by the zero of the first derivative of $\phi(z)$ and is a solution of the saddle point equation

$$\mathcal{J} - \frac{\mathcal{B}^2}{4\mathcal{J}(z_0 - d)^2} = \frac{1}{2} g'(z_0). \tag{3.5.12}$$

Since there is a unique positive solution of this equation, it permits us to define F as a function of $\mathcal{J}$ and $\mathcal{B}$, for $\mathcal{J} > 0$ and $\mathcal{B} \neq 0$. In Appendix 3B we will show that this equation permits us to establish an interesting relation between the spherical model and brownian motion on a lattice.

Equation of state. The equation of state of the spherical model can be derived as follows. Let us first take a derivative of (3.5.11) with respect to $\mathcal{B}$, keeping $\mathcal{J}$ fixed. Based on (3.5.12) and taking into account that z_0 also depends on $\mathcal{B}$, one has

$$-\frac{d}{d\mathcal{B}}\left(\frac{F}{kT}\right) = \frac{\mathcal{B}}{2\mathcal{J}(z_0 - d)} + \phi'(z_0)\frac{dz_0}{dh}.$$

However, z_0 is exactly the value where the first derivative of $\phi(z)$ vanishes. Using the thermodynamic relation

$$M(B,T) = -\frac{\partial}{\partial H} F(B,T),$$

we have

$$M = \frac{\mathcal{B}}{2\mathcal{J}(z_0 - d)} = \frac{B}{2J(z_0 - d)}.$$

We can now eliminate the variable $(z_0 - d)$ by using the saddle point equation (3.5.12), with the result

$$2J(1 - M^2) = kTg'\left(\frac{B}{2JM}\right).$$

This is the exact equation of state of the spherical model that links the quantities M, B, and T.

Let us now discuss in more detail the saddle point equation (3.5.12) to see if there is a phase transition in the spherical model. The function $g'(z)$ is expressed by the multidimensional integral

$$g'(z) = \frac{1}{(2\pi)^d} \int_0^{2\pi} \cdots \int_0^{2\pi} \frac{1}{z - \sum_{j=1}^{d} \cos\omega_j} \, d\omega_1 \ldots d\omega_d. \tag{3.5.13}$$

Using the identity

$$\frac{1}{a} = \int_0^\infty e^{-at}\, dt,$$

and the integral representation (2.A.12) of the Bessel function $I_0(t)$, given in Appendix A of Chapter 2,

$$I_0(t) \,=\, \frac{1}{2\pi} \int_0^{2\pi} e^{t\cos\omega}\, d\omega,$$

$g'(z)$ can be expressed in a more convenient form as

$$g'(z) \,=\, \int_0^{\infty} e^{-tz}\, [I_0(t)]^d\, dt. \tag{3.5.14}$$

This formula has the advantage of showing the explicit dependence of the dimension d of the lattice, which can be regarded as a continuous variable and not necessarily restricted to integer values.

Let us study the main properties of $g'(z)$. From the asymptotic behavior of $I_0(t)$,

$$I_0(t) \,\simeq\, \frac{e^t}{\sqrt{2\pi t}}, \quad t \to \infty \tag{3.5.15}$$

it follows that the integral (3.5.14) converges when $\mathrm{Re}\, z > d$. Consequently, $g'(z)$ is an analytic function in this semiplane. For real z, $g'(z)$ is a positive function, that monotonically decreases toward its null value when $z \to \infty$. For $z \to d$, using once again (3.5.15), the integral diverges when $d \leq 2$, while it converges when $d > 2$

$$\lim_{z\to d} g'(z) \,=\, \begin{cases} \infty, & 0 < d < 2 \\ g'(d) < \infty, & d > 2. \end{cases}$$

This implies that there is a phase transition for $\mathcal{B} = 0$ only when $d > 2$. Consider, in fact, eqn (3.5.12) when $\mathcal{B} = 0$

$$2\mathcal{J} \,=\, g'(z). \tag{3.5.16}$$

If $g'(z)$ diverges for $z \to d$, however we vary the value of $\mathcal{J}$ (i.e. the value of the temperature), there is always a root z_0 of the equation that varies with continuity, as shown by its graphical solution of Fig. 3.8.

Vice versa, if $g'(z)$ converges towards the finite value $g'(d)$ when $z \to d$, there is a solution z_0 that varies with continuity as long as $\mathcal{J} < g'(d)$. However, when the function reaches the value $\mathcal{J} = g'(d)$, there is a discontinous change in the nature of the equation. Since the function $g'(z)$ cannot grow more than its limit value $g'(d)$, further increasing $\mathcal{J}$ the root z_0 of the equation remains fixed at its value $z_0 = d$, as shown in Fig. 3.9. The appearance of a spontaneous magnetization below the critical temperature may be regarded qualitatively as a condensation phenomenon akin to Bose–Einstein condensation of integer spins atoms (see Appendix B of Chapter 1). The phase transition point is identified, for $d > 2$, by the condition

$$\mathcal{J}_c \,=\, \frac{J}{kT_c} \,=\, \frac{1}{2}\, g'(d). \tag{3.5.17}$$

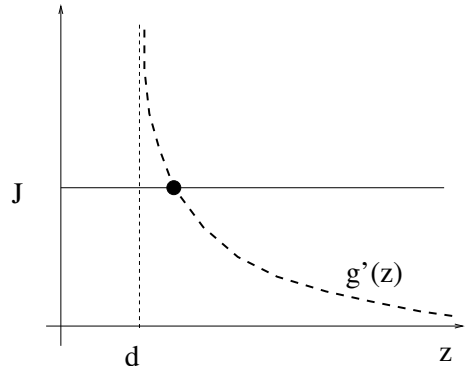


Fig. 3.8 *Graphical solution of the saddle point equation for* $d < 2$.

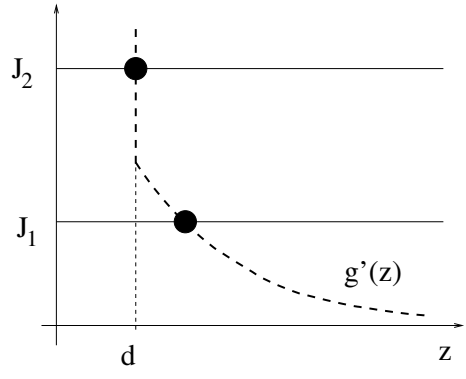


Fig. 3.9 *Graphical solution of the saddle point equation for* $d > 2$. *There is a phase transition when* $\mathcal{J} = g'(d)$.

From the detailed analysis of the model, as proposed in one of the problems at the end of the chapter, one arrives to the following conclusions: first of all, there is no phase transition for $d \leq 2$, while for $d > 2$ there is a phase transition with the values of the critical exponents as follows: α assumes the value

$$\alpha = \begin{cases} -(4-d)/(d-2), & 2 < d < 4, \\ 0, & d > 4, \end{cases}$$

while β is given by

$$\beta = \frac{1}{2}.$$

For the critical exponent γ we have

$$\gamma = \begin{cases} 2/(d-2), & 2 < d < 4, \\ 1, & d > 4. \end{cases}$$

Finally, the value of the critical exponent δ is

$$\delta = \begin{cases} (d+2)/(d-2), & 2 < d < 4, \\ 3, & d > 4. \end{cases}$$

Using these results, it is easy to establish the validity of the first two scaling laws (1.1.26). The other two scaling laws permit us to determine the critical exponent ν

$$\nu = \begin{cases} 1/(d-2), & 2 < d < 4, \\ 1/2, & d > 4. \end{cases}$$

and the critical exponent η

$$\eta = 0.$$

In conclusion, the spherical model has the interesting property that its critical exponents vary with the dimensionality d of the lattice in the range $2 < d < 4$, while they assume the values predicted by the mean field theory for $d > 4$. One expects to find the same behavior in the critical exponents of the Ising model, obviously with a different set of values for the two models.

Appendix 3A. The Saddle Point Method

In many mathematical situations, one faces the problem of estimating the asymptotic behavior of a function $J(s)$ when $s \to \infty$. Some examples were shown in the previous chapter (the asymptotic behavior of the $\Gamma(s)$ function or the Bessel functions $I_\nu(s)$) and in this chapter (the partition function of the spherical model). In this appendix we study how to solve this problem when the function $J(s)$ is expressed as an integral, of general form

$$J(s) = \int_{\mathcal{C}} g(z)\, e^{s f(z)}\, dz. \tag{3.A.1}$$

In the following we will consider the case in which s is a real variable. The contour $\mathcal{C}$ is chosen in such a way that the real part of $f(z)$ goes to $-\infty$ at both points of integration (so that the integrand vanishes in these regions) or as a closed contour in the complex plane.[12]

If the variable s assumes quite large positive values, the integrand is large when the real part of $f(z)$ is also large and, vice versa, is small when the real part of $f(z)$ is either small or negative. In particular, for $s \to +\infty$, the significant contribution of the integral comes from those regions in which the real part of $f(z)$ assumes its maximum positive value. To see this, expressing $f(z)$ as

$$f(z) = u(x,y) + i\, v(x,y),$$

[12] In the following we assume that the function $g(z)$ is significantly smaller than the term $e^{s f(z)}$ in the regions of interest.

one has

$$J(s) \;=\; \int_{\mathcal{C}} g(z)\, e^{su(x,y)}\, e^{isv(x,y)}\, dz.$$

If we make the hypothesis that the imaginary part of the exponent, $iv(x,y)$, is approximately constant in the region where the real part has its maximum, i.e. $v(x,y) \simeq v(x_0,y_0) = v_0$, one can approximate the integral as follows

$$J(s) \;\simeq\; e^{isv_0} \int_{\mathcal{C}} g(z)\, e^{su(x,y)}\, dz.$$

Far from the point of the maximum of the real part, the imaginary part can oscillate in an arbitrary way, for the integrand is anyway small and the phase factor quite irrelevant.

Let us now discuss the properties of the maximum point of $sf(z)$. The real part of $sf(z)$ has a maximum, at a given s, corresponding to the maximum of the real part of $f(z)$, i.e. $u(x,y)$. This point is determined by the equations

$$\frac{\partial u}{\partial x} \;=\; \frac{\partial u}{\partial y} \;=\; 0.$$

From the Cauchy–Riemann equations satisfied by the analytic functions, these equations can be expressed as

$$\frac{df(z)}{dz} \;=\; 0. \tag{3.A.2}$$

It is important to stress that the maximum of $u(x,y)$ is such *only along a particular contour*. In fact, for all points of the complex plane at a finite distance from the origin, neither the real nor the imaginary parts of an analytic function have an absolute maximum or an absolute minimum. This is a direct consequence of the Laplace equation satisfied by both functions u and v

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \;&=\; 0; \\ \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \;&=\; 0. \end{aligned}$$

If the second derivative with respect to x of one of the functions u or v is positive, its second derivative with respect to y is necessarily negative. Hence, none of the two functions can have an absolute maximum or minimum. The vanishing of the first derivative of $f(z)$, eqn (3.A.2), implies that we are in the presence of a saddle point: this is a stationary point that is a maximum of $u(x,y)$ along one contour, but a minimum along another (see Fig. 3.10).

The problem is then how to choose a path of integration $\mathcal{C}$ that satisfies the following conditions: (a) there exists a maximum of $u(x,y)$ along $\mathcal{C}$; (b) the contour passes through the saddle point, so that the imaginary part $v(x,y)$ has the smallest variation. From complex analysis, it is known that the curves associated to the equations u = constant and v = constant form a system of orthogonal curves, and the curve $v = c$ (where c is a costant) is always tangent to the gradient ∇u of u. Hence, this

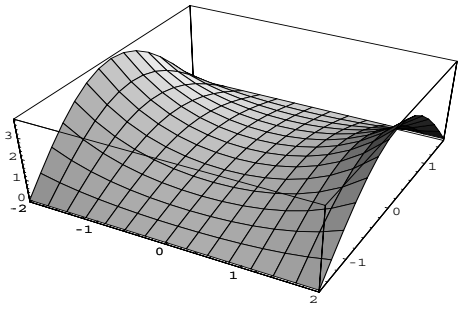


Fig. 3.10 *Saddle point of an analytic function.*

is the curve along which we have the maximum decreasing of the function each time we move away from the saddle. Therefore this is the curve to select as the contour of integration $\mathcal{C}$.

At the saddle point, the function $f(z)$ can be expanded in its Taylor series

$$f(z) \simeq f(z_0) + \frac{1}{2}(z - z_0)^2 f''(z_0) + \cdots$$

Along $\mathcal{C}$, the quadratic correction of the function is both real (the imaginary part is constant along the chosen path) and negative (since we are moving along the path of fastest decrease from the saddle point). Assuming $f(z_0) \neq 0$, we have

$$f(z) - f(z_0) \simeq \frac{1}{2}(z - z_0)^2 f''(z_0) \equiv -\frac{1}{2s} t^2,$$

where we have defined the new variable t. Expressing $(z - z_0)$ in polar coordinates

$$(z - z_0) = \delta e^{i\alpha},$$

(with the phase being fixed), we get

$$t^2 = -sf''(z_0)\, \delta^2 e^{2i\alpha}.$$

Since t is real, one has

$$t = \pm\delta \mid sf''(z_0) \mid^{1/2},$$

and substituting in (3.A.1), we obtain[13]

$$J(s) \simeq g(z_0)\, e^{sf(z_0)} \int_{-\infty}^{+\infty} e^{-t^2/2} \frac{dz}{dt}\, dt. \tag{3.A.3}$$

Since

$$\frac{dz}{dt} = \left(\frac{dt}{dz}\right)^{-1} = \left(\frac{dt}{d\delta}\,\frac{d\delta}{dz}\right)^{-1} = \mid sf''(z_0) \mid^{-1/2}\ e^{i\alpha},$$

[13]The integral has been extended to $\pm\infty$ since the integrand is small when t is large.

eqn (3.A.3) becomes

$$J(s) \simeq \frac{g(z_0)\, e^{sf(z_0)}\, e^{i\alpha}}{\mid sf''(z_0) \mid^{1/2}} \int_{-\infty}^{+\infty} e^{-t^2/2}\, dt. \tag{3.A.4}$$

The integral is now gaussian (equal to $\sqrt{2\pi}$) so the asymptotic behavior of $J(s)$ is given by

$$J(s) \simeq \frac{\sqrt{2\pi}\, g(z_0)\, e^{sf(z_0)}\, e^{i\alpha}}{\mid sf''(z_0) \mid^{1/2}}, \qquad s \to +\infty. \tag{3.A.5}$$

Two comments are in order. Sometimes the integration contour passes through two or more saddle points. In such cases, the asymptotic behavior of $J(s)$ is obtained by summing all the contributions (3.A.5) relative to the different saddle points. The second comment is about the validity of the method: in our discussion we have assumed that the only significant contribution to the integral comes from the region near the saddle point $z = z_0$. This means that one should always check that the condition

$$u(x, y) \ll u(x_0, y_0)$$

holds along the entire contour $\mathcal{C}$ away from $z_0 = x_0 + iy_0$.

Appendix 3B. Brownian Motion on a Lattice

In this appendix we will recall the basic notions of brownian motion on a d-dimensional lattice. We will also show the interesting relation between this problem and the spherical model discussed in the text.

Binomial coefficients. Let us initially consider the one-dimensional case, with lattice sites identified by the variable s, with $s = 0, \pm 1, \pm 2, \ldots$: the problem consists of studying the motion of a particle that, at each discrete time step t_n, has a probability p and $q = 1 - p$ to move respectively to the neighbor site on its right or on its left (see Fig. 3.11). Suppose that at $t_0 = 0$ the particle is at the origin $s = 0$: what is the probability $P_n(s)$ that at time t_n (after n steps) the particle is at site s? There are several way to determine such a quantity. One of the most elegant methods consists

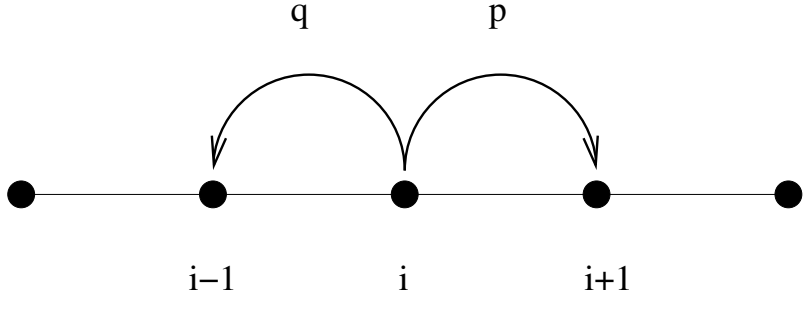


Fig. 3.11 *Brownian motion on a one-dimensional lattice.*

of assigning a weight $e^{i\phi}$ to the jump toward the right site and a weight $e^{-i\phi}$ to the one toward the left site and to consider the binomial

$$\left(pe^{i\phi} + qe^{-i\phi}\right)^n.$$

For $n = 1$, one has

$$\left(pe^{i\phi} + qe^{-i\phi}\right),$$

from which we can see that the coefficient p in front of $e^{i\phi}$ represents the probability that, after the first step, the particle is at site $s = 1$, placed to the right of the origin, whereas the coefficient q in front of the other exponential $e^{-i\phi}$ gives the probability that the particle is at site $s = -1$ on the left of the origin. Similarly, considering the expression

$$\left(pe^{i\phi} + qe^{-i\phi}\right)^2,$$

and expanding the binomial, the coefficient p^2 in front of the term $e^{2i\phi}$ gives the probability that the particle is at site $s = 2$ after two steps, the coefficient $2pq$ in front of $e^{0i\phi}$ gives the probability to find the particle at origin, while the coefficient q^2 in front of $e^{-2i\phi}$ expresses the probability to find the particle at site $s = -2$. More generally, we have that

$$P_n(s) \;=\; \text{coefficient in front of}\, e^{is\phi} \;\text{in}\; \left(pe^{i\phi} + qe^{-i\phi}\right)^n.$$

Thanks to the identity

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\phi a}\, d\phi \;=\; \delta_{a,0},$$

such a coefficient can be filtered by means of the Fourier transform, so that

$$P_n(s) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(pe^{i\phi} + qe^{-i\phi}\right)^n e^{-is\phi}\, d\phi. \tag{3.B.1}$$

In the symmetric case, $p = q = \frac{1}{2}$, we have

$$P_n(s) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} (\cos\phi)^n\, e^{-is\phi}\, d\phi \;=\; \frac{1}{2^n}\,\frac{n!}{\left[\frac{1}{2}(n+s)\right]!\,\left[\frac{1}{2}(n-s)\right]!}. \tag{3.B.2}$$

In this case, if n is even, the only possible values of s are also even, with $\mid s \mid \leq n$, while if n is odd then s is also odd, with $\mid s \mid < n$.

The origin of the binomial coefficient becomes evident by looking at Fig. 3.12. In fact, the computation of $P_n(s)$ is equivalent to counting the number of different paths that start from the origin and reach the point s after n steps. In these paths each turn to the right or to the left is weighted by p and q, respectively.

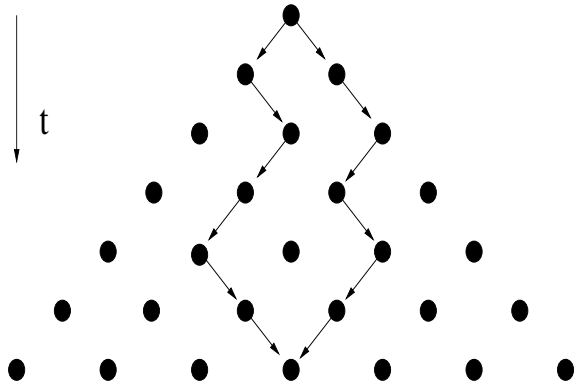


Fig. 3.12 *Two paths that lead to the same point after n steps.*

Continuous probability. When $n \to \infty$ the integral (3.B.6) can be estimated by the saddle point method. In this limit, the dominant term of the integral comes from the values of ϕ_i near the origin, so that expanding in series the term $\left[\frac{1}{d}(\cos\phi_1 + \cdots + \cos\phi_d)\right]^n$ and keeping only the quadratic terms, one has

$$\begin{aligned}\left[\frac{1}{d}\,(\cos\phi_1 + \cdots + \cos\phi_d)\right]^n &= \exp\left[n\,\log\frac{1}{d}\,(\cos\phi_1 + \cdots + \cos\phi_d)\right] \\ &\simeq \exp\left[-\frac{n}{2d}\,\left(\phi_1^2 + \phi_2^2 + \cdots + \phi_d^2\right)\right].\end{aligned}$$

Changing variables $x_i = \phi_i\, n^{1/2}$ and performing the integral (3.B.6), one obtains the gaussian distribution

$$P_n(\vec{s}) \simeq \left(\frac{d}{2\pi n}\right)^{d/2} \exp\left[-\frac{d}{2n}\,\vec{s}\cdot\vec{s}\right]. \tag{3.B.3}$$

If we now denote by a the lattice spacing and by τ the time interval between each transition, the variable $\vec{x} = a\vec{s}$ is the distance of the particle from the origin after the time $t = n\tau$. The function $P_n(\vec{s})$ in (3.B.3) is related to the *continuous probability density* $P(\vec{x}, t)$ to find the particle in the volume $d\vec{x}$ nearby the point $\vec{x}$

$$P(\vec{x}, t) \;=\; \frac{1}{(4\pi\mathcal{D}t)^{d/2}}\,\exp\left[-\frac{\vec{x}\cdot\vec{x}}{4\mathcal{D}t}\right], \tag{3.B.4}$$

where $\mathcal{D} = \frac{a^2}{2d\tau}$ is the diffusion constant. In fact, the function $P(\vec{x}, t)$ satisfies the differential equation of the diffusion process

$$\left(\frac{\partial}{\partial t} - \mathcal{D}\,\nabla^2\right)\,P(\vec{x}, t) \;=\; 0,$$

where ∇^2 is the laplacian operator in d dimensions. The dispersion of the probability density $P(\vec{x},t)$ is expressed by the mean value $\langle \mid \vec{x} \mid^2 \rangle$, computed with respect to the probability distribution (3.B.4): this quantity grows linearly with time:

$$\langle \mid \vec{x} \mid^2 \rangle \; = \; 2\mathcal{D}\, t. \tag{3.B.5}$$

Generalization. The analysis of the one-dimensional case can be easily generalized in higher dimensional lattices. Consider, for instance, a d-dimensional cubic lattice in which, at each discrete temporal step, there are $2d$ possible transitions to the neighbor sites. For simplicity, let us assume that all these probabilities are the same and equal to $\frac{1}{2d}$. Assigning the weight $e^{i\phi_i}$ for the jump ahead and $e^{-i\phi_i}$ for the jump back along the i-th direction, the probability of finding the walker at site s with coordinates $\vec{s} = (s_1, s_2, \ldots, s_d)$ after n steps is expressed by the d-dimensional Fourier transform

$$P_n(\vec{s}) \; = \; \frac{1}{(2\pi)^d} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \left[\frac{1}{d} (\cos\phi_1 + \cos\phi_2 + \cdots \cos\phi_d) \right]^n \; e^{-i\vec{s}\cdot\vec{\phi}}\, d^d\phi. \tag{3.B.6}$$

The problem can be easily generalized to the cases where there are transitions between arbitrary sites, not necessarily next neighbor. Let $\vec{s}_i$ and $\vec{s}_j$ be two sites of a d-dimensional lattice, with total number of sites equal to L^d. Assuming periodic boundary conditions along all directions, we have the equivalence relationships

$$(s_1, s_2, \ldots, s_d) \; \equiv \; (s_1 + L, s_2, \ldots) \; \equiv \; (s_1, s_2 + L, s_3, \ldots) \; \equiv \; \ldots$$

Let $p(\vec{s}_i - \vec{s}_j)$ be the probability of the transition $\vec{s}_j \longrightarrow \vec{s}_i$. For simplicity we assume that this probability is time independent and a function only of the distance between the two sites. Let us denote, as before, by $P_n(\vec{s})$ the probability that the particle is at site $\vec{s}$ after n steps. This function satisfies the recursive equation

$$P_{n+1}(\vec{s}) \; = \; \sum_{\vec{s}_j} p(\vec{s} - \vec{s}_j)\, P_n(\vec{s}_j), \tag{3.B.7}$$

with the initial condition

$$P_0(\vec{s}) \; = \; \delta_{\vec{s},\mathbf{0}}. \tag{3.B.8}$$

Due to their probabilistic nature, $P_n(\vec{s})$ and $p(\vec{s})$ satisfy the normalization conditions

$$\sum_{\vec{s}} P_n(\vec{s}) \; = \; 1, \qquad \sum_{\vec{s}} p(\vec{s}) \; = \; 1. \tag{3.B.9}$$

To solve the recursive equation (3.B.7) let us introduce the generating function[14]

$$G(\vec{s}, w) \; = \; \sum_{n=0}^{\infty} P_n(\vec{s})\, w^n. \tag{3.B.10}$$

[14] A brownian motion is *transient* if $G(0,1)$ is a finite quantity, while it is *recurrent* if $G(0,1)$ is instead divergent. The origin of this terminology will become clear below.

Let's now multiply eqn (3.B.7) by w^{n+1} and sum over n. Taking into account the definition of $G(\vec{s},w)$ and the initial condition (3.B.8), the generating function $G(\vec{s},w)$ satisfies the equation

$$G(\vec{s},w) - w\sum_{\vec{s}'} p(\vec{s}-\vec{s}')\,G(\vec{s}',w) \;=\; \delta_{\vec{s},\mathbf{0}}, \tag{3.B.11}$$

where the convolution term comes from the translation invariance of the lattice. This suggests finding its solution by expanding $G(\vec{s},w)$ in a Fourier series. Let $g(\vec{k},w)$ and $\lambda(\vec{k})$ be the Fourier transforms of $G(\vec{s},w)$ and $p(\vec{s})$:

$$\begin{aligned} g(\vec{k},w) &= \sum_{\vec{s}} G(\vec{s},w)\,\exp\left[i\vec{k}\cdot\vec{s}\right]; \\ \lambda(\vec{k}) &= \sum_{\vec{s}} p(\vec{s})\,\exp\left[i\vec{k}\cdot\vec{s}\right], \end{aligned} \tag{3.B.12}$$

with $\vec{k}=\frac{2\pi}{L}\vec{r}$ and $r_j = 0,1,2,\ldots,(L-1)$. In terms of these quantities, eqn (3.B.11) can be written as

$$g(\vec{k},w) \;-\; w\,\lambda(\vec{k})\;g(\vec{s},w) \;=\; 1,$$

from which

$$g(\vec{k},w) \;=\; \frac{1}{1-w\lambda(\vec{k})}. \tag{3.B.13}$$

Taking now the inverse Fourier transform, the solution of (3.B.11) is

$$G(\vec{s},w) \;=\; \frac{1}{L^d}\sum_{\{r_j=0\}}^{L-1}\frac{\exp\left(-2\pi i\vec{r}\cdot\vec{s}/L\right)}{1-w\,\lambda\left(2\pi\vec{r}/L\right)}. \tag{3.B.14}$$

Since $P_n(\vec{s})$ is the coefficient of w^n in $G(\vec{s},w)$, expanding in series the expression above we obtain

$$P_n(\vec{s}) \;=\; \frac{1}{L^d}\sum_{\{r_j=0\}}^{L-1}\left[\lambda\left(\frac{2\pi\vec{r}}{L}\right)\right]^n \exp\left(-2\pi i\vec{r}\cdot\vec{s}/L\right). \tag{3.B.15}$$

When $L\to\infty$, the generating function $G(\vec{s},w)$ is expressed by the integral

$$G(\vec{s},w) \;=\; \frac{1}{(2\pi)^d}\int_0^{2\pi}\cdots\int_0^{2\pi}\frac{\exp(-i\vec{s}\cdot\vec{k})}{1-w\,\lambda(\vec{k})}\,d\vec{k}. \tag{3.B.16}$$

If the transitions are only those between next neighbor sites of a cubic lattice, the function $\lambda(\vec{k})$ is given by

$$\lambda(\vec{k}) \;=\; \frac{1}{d}\sum_{j=1}^{d}\cos k_j,$$

and $G(\vec{s},w)$ can be written as

$$G(\vec{s},w) \;=\; \frac{1}{(2\pi)^d}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\frac{\exp(-i\vec{s}\cdot\vec{k})}{1-w\,d^{-1}\sum_{j=1}^{d}\cos k_j}\,d\vec{k}. \tag{3.B.17}$$

Note that $\tilde{G}(\vec{s},w) \equiv wG(\vec{s},w)$ satisfies the equation

$$\left[-\nabla_s^2 + (w^{-1}-1)\right]\tilde{G}(\vec{s},w) \;=\; \delta_{\vec{s},0},$$

where ∇_s^2 is the discrete version of the laplacian operator on the d-dimensional lattice

$$\nabla_s^2 f(\vec{s}) \;\equiv\; \frac{1}{2d}\sum_{\mu=1}^{d}\left[f(\vec{s}+\vec{e}_\mu)+f[\vec{s}-\vec{e}_\mu-2f(\vec{s})\right].$$

This function is analogous of the euclidean propagator of a free bosonic field of mass m: in fact, rescaling the quantities by the lattice space a according to $\vec{s}\to\vec{s}/a$, $\vec{k}\to\vec{k}a$ and imposing

$$w^{-1} \;=\; 1+\; m^2\frac{a^2}{2d},$$

we have

$$\mathcal{D}(\vec{s},m^2) \;=\; \lim_{a\to 0}\frac{1}{2da^{d-2}}\,\tilde{G}\left(\frac{\vec{s}}{a},w\right) \;=\; \int_{-\infty}^{+\infty}\frac{d^d\vec{k}}{(2\pi)^d}\,\frac{e^{i\vec{k}\cdot\vec{s}}}{k^2+m^2}. \tag{3.B.18}$$

The relationship between the spherical model and brownian motion should now be clear. In fact, the function $g'(z)$ defined by (3.5.13) and entering the saddle point equation of model (3.5.16) is nothing else but the generating function of the brownian motion on a cubic lattice! More generally, the spherical model with coupling constants J_{ij} is related to the brownian motion with a probability transitions $p(\vec{s}_i-\vec{s}_j)$ proportional to J_{ij}. There is, in fact, the following identity

$$g'(z) \;=\; \frac{1}{z}\,G(0,dz^{-1}). \tag{3.B.19}$$

Transient and recurrent brownian motion. As discussed in the text, there is a phase transition in the spherical model only if $g'(d)$ is finite. For brownian motion, this condition implies that the corresponding brownian motion is *transient* and not *recurrent*. For $d=1$ and $d=2$ the brownian motion is always recurrent:[15] this means that a brownian motion that starts from the origin will always come back to the origin with probability equal to 1. For $d\geq 3$, $G(0,1)$ is a finite quantity and this implies that the brownian motion is *transient*: this means that there is a finite probability that the walker never comes back to the origin. These results are part of the famous problem posed by Polya about the probability of the random walk to return to a given site and its dependence on the dimensionality of the lattice. To derive these results, in general it is useful to introduce the following functions:

- $P_n(\vec{s},\vec{s}_0)$ = probability to be at site $\vec{s}$ after n steps, where $\vec{s}_0$ is the starting point;
- $F_n(\vec{s},\vec{s}_0)$ = probability to be at site $\vec{s}$ *for the first time* after n steps, where $\vec{s}_0$ is the starting point;

[15] In the two-dimensional case, this result gives support to the popular saying *All roads lead to Rome*.

together with their corresponding generating functions

$$\begin{aligned} G(\vec{s},\vec{s}_0;w) &= \delta_{\vec{s},\vec{s}_0} + \sum_{n=1}^{\infty} P_n(\vec{s},\vec{s}_0)\, w^n, \\ \mathcal{F}(\vec{s},\vec{s}_0;w) &= \sum_{n=1} F_n(\vec{s},\vec{s}_0)\, w^n. \end{aligned} \tag{3.B.20}$$

The functions P_n and F_n satisfy

$$\begin{aligned} P_0(\vec{s},\vec{s}_0) &= \delta_{\vec{s},\vec{s}_0}, \\ P_n(\vec{s},\vec{s}_0) &= \sum_{k=1}^{n} P_{n-k}(\vec{s},\vec{s})\, F_k(\vec{s},\vec{s}_0). \end{aligned} \tag{3.B.21}$$

In fact, the particle can reach the site $\vec{s}$ for the first time after k steps and can come back later to the same site in the remaining $(n-k)$ steps. So, the sum over k corresponds to all independent ways to implement the transition $\vec{s}_0 \to \vec{s}$ in n steps. Multiplying these equations by w^n, summing on n, and using the generating functions, one has

$$\begin{aligned} G(\vec{s},\vec{s}_0;w) &= \sum_{n=0} P_n(\vec{s},\vec{s}_0)\, w^n \\ &= \delta_{\vec{s},\vec{s}_0} + \sum_{n=1}^{\infty}\sum_{k=1}^{n} w^k\, F_k(\vec{s},\vec{s}_0)\; w^{n-k}\, P_{n-k}(\vec{s},\vec{s}) \\ &= \delta_{\vec{s},\vec{s}_0} + G(\vec{s},\vec{s},w)\, \mathcal{F}(\vec{s},\vec{s}_0,w). \end{aligned} \tag{3.B.22}$$

Hence

$$\begin{aligned} &\mathcal{F}(\vec{s}_0,\vec{s}_0,w) = 1 - [G(\vec{s}_0,\vec{s}_0,w)]^{-1}, \\ &\mathcal{F}(\vec{s},\vec{s}_0,w) = G(\vec{s},\vec{s}_0,w)/G(\vec{s},\vec{s},w) \qquad \text{if } \vec{s} \neq \vec{s}_0. \end{aligned} \tag{3.B.23}$$

These formulas can now be used to study the nature of the brownian motion on different lattices. If we have translation invariance, $G(\vec{s},\vec{s}_0;w) = G(\vec{s}-\vec{s}_0;w)$ and analogously for $\mathcal{F}$. Note that $\mathcal{F}(0,1)$ is exactly the probability that a particle comes back soon or later to its starting point. In fact

$$\mathcal{F}(0,1) = F_1(0) + F_2(0) + \cdots \tag{3.B.24}$$

and therefore this quantity corresponds to the sum of the probabilities of all independent ways to come back to the origin, i.e. for the first time after one step, two steps, etc. On the other hand, from (3.B.23) one has

$$\mathcal{F}(0,1) = 1 - [G(0,1)]^{-1}, \tag{3.B.25}$$

so that the particle has probability equal to 1 to come back to the origin if $G(0,1)$ is a divergent quantity, as we saw it happen for $d = 1$ and $d = 2$. On the contrary, in

three dimension and for a cubic lattice we have

$$\begin{aligned} G(0,1) &= \frac{1}{(2\pi)^3}\int_0^{2\pi}\int_0^{2\pi}\int_0^{2\pi}\frac{d^3\vec{k}}{1-\frac{1}{3}(\cos k_1+\cos k_2+\cos k_3)} \\ &= \left(\frac{\sqrt{6}}{32\pi^3}\right)\Gamma\left(\frac{1}{24}\right)\Gamma\left(\frac{5}{24}\right)\Gamma\left(\frac{7}{24}\right)\Gamma\left(\frac{11}{24}\right) \\ &= 1.516386059.... \end{aligned} \tag{3.B.26}$$

so that the probability to return to the origin is equal to

$$\mathcal{F}(0,1) \;=\; 0.34053733... \tag{3.B.27}$$

Number of distinct points visited in the brownian motion. Denoting by S_n the mean value of the distinct points visited by the walker after n steps, let's now derive the following asymptotic value when $n \to \infty$ for various lattices

$$S_n \;\simeq\; \begin{cases} \left(\frac{8n}{\pi}\right)^{\frac{1}{2}} & d \;=\; 1, \\ \frac{\pi n}{\log n} & d \;=\; 2, \\ C_d\, n & d \geq 3, \end{cases} \tag{3.B.28}$$

where the constant C_d depends on the structure of the lattice. For their derivation, observe that

$$S_n \;=\; 1 + {\sum_{\vec{s}}}' \left[F_1(\vec{s}) + F_2(\vec{s}) + \cdots + F_n(\vec{s})\right],$$

where the sum is over all sites of the lattice but the origin. The first term of this expression is related to the origin, i.e. to the initial condition of the particle. With the definition previously given for $F_n(\vec{s})$, each term in the sum represents the probability that a site of the lattice has been visited at least once in the first n steps. Consider now

$$\Delta_k \;=\; S_k - S_{k-1}, \quad k = 1, 2, \ldots$$

Since $S_0 = 1$ and $S_1 = 2$, one has $\Delta_1 = 1$. Moreover

$$\Delta_n \;=\; {\sum_{\vec{s}}}' F_n(\vec{s}) \;=\; -F_n(0) + \sum_{\vec{s}} F_n(\vec{s}),$$

where the sum is now extended to all lattice sites. The generating function of Δ_n is given by

$$\Delta(w) \;=\; \sum_{n=1}^{\infty} w^n\, \Delta_n \;=\; -\mathcal{F}(0,w) + \sum_{\vec{s}} \mathcal{F}(\vec{s},w).$$

From (3.B.22) one has

$$\mathcal{F}(\vec{s},w) \;=\; \frac{G(\vec{s},w)-\delta_{\vec{s},0}}{G(0,w)}.$$

Since for any n

$$\sum_{\vec{s}} P_n(\vec{s}) \;=\; 1,$$

one gets

$$\sum_{\vec{s}} G(\vec{s},w) \;=\; 1+w+w^2+\cdots \;=\; \frac{1}{1-w}.$$

Therefore

$$\Delta(w) \;=\; -1+\frac{1}{(1-w)\,G(0,w)}. \tag{3.B.29}$$

Taking into account that

$$\begin{aligned} S_0 \;&=\; 1, \qquad S_1 \;=\; 2 \\ S_n \;&=\; 1+\Delta_1+\Delta_2+\cdots+\Delta_n, \qquad n\geq 1 \end{aligned}$$

the generating function of S_n is expressed by

$$\begin{aligned} S(w) &= \sum_{n=0}^{\infty} w^n\, S_n \\ &= (1-w)^{-1}\,\left[1+w\Delta_1+w^2\Delta_2+\cdots\right] \\ &= (1-w)^{-1}\,\left[1-\Delta(w)\right]^{-1} \;=\; \left[(1-w)^2\,G(0,w)\right]^{-1}. \end{aligned} \tag{3.B.30}$$

Consider $G(0,w)$ for various lattices. For $d=1$, we have

$$G(0,w) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{dk}{1-w\cos k} \;=\; \frac{1}{\sqrt{1-w^2}}. \tag{3.B.31}$$

For $d=2$ we have

$$G(0,w) \;=\; \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\frac{dk_1\,dk_2}{1-\frac{1}{2}w(\cos k_1+\cos k_2)} \;=\; \frac{2}{\pi}\,K(w), \tag{3.B.32}$$

where

$$K(w) \;=\; \int_0^{\pi/2}\frac{d\alpha}{\sqrt{1-w^2\sin^2\alpha}} \tag{3.B.33}$$

is the elliptic integral of first kind. For $w\to 1$, $K(w)$ has a logarithmic singularity, so that

$$G(0,w) \;\simeq\; -\frac{1}{\pi}\,\log(1-w)+\mathcal{O}(1), \qquad z\to 1. \tag{3.B.34}$$

For $d\geq 3$, $G(0,w)$ has a finite limit for $w\to 1$, in particular for $d=3$ it is given by (3.B.26).

To derive the asymptotic behavior of S_n we need the following theorem.

Theorem 3.1 *Let* $U(y) = \sum_n a_n e^{-ny}$ *be a convergent series for all values* $y > 0$, *with* $a_n > 0$. *If, for* $y \to 0$, $U(y)$ *behaves as*

$$U(y) \sim \Phi(y^{-1}),$$

where $\Phi(x) = x^\sigma L(x)$ *is an increasing positive function of* x *that goes to infinity when* $x \to \infty$, *with* $\sigma \geq 0$ *and* $L(cx) \sim L(x)$ *for* $x \to \infty$, *then*

$$a_1 + a_2 + \cdots + a_n \sim \frac{\Phi(n)}{\Gamma(\sigma + 1)}. \tag{3.B.35}$$

If we now substitute the different expressions of $G(0,1)$ in (3.B.29), put $z = e^{-y}$, and study the limit $y \to 0$, we have

$$\Delta(y) \simeq \begin{cases} (2/y)^{1/2} & d = 1, \\ \pi/(y \log 1/y) & d = 2, \\ 1/yG(0,1) & d \geq 3. \end{cases} \tag{3.B.36}$$

Hence

$$\begin{aligned} &d = 1,\ \sigma = \tfrac{1}{2},\ L(x) = 2^{1/2}, \\ &d = 2,\ \sigma = 1,\ L(x) = \pi/\log x, \\ &d \geq 3,\ \sigma = 1,\ L(x) = 1/G(0,1). \end{aligned} \tag{3.B.37}$$

Putting now $a_i = \Delta_i$ in (3.B.35), we obtain the asymptotic behavior (3.B.28) of the mean value of the distinct sites visited after n steps, in the limit $n \to \infty$.

Relation with prime numbers. It is interesting to note that for $d = 2$ the number of distinct sites visited in n steps is proportional to the number of prime numbers less than the integer n. This quantity has been estimated originally by Gauss: denoting by $\Pi(n)$ the number of primes less than n, Gauss found the asymptotic form of such a function:

$$\Pi(n) \simeq \frac{n}{\log n}. \tag{3.B.38}$$

The coincidence between this aspect of number theory and brownian motion has an elementary explanation that clarifies some important aspects of the prime numbers.

Gauss's law can be derived in a simple way by employing the sieve of Eratosthenes. Let us denote by $P(n)$ the probability that an integer n is a prime number. Since a generic integer n has probability $1/p_i$ of being divisible by p_i (this comes directly from the sieve of Eratosthenes), the probability that the number n is not divisible by p_i is equal to $(1 - 1/p_i)$. Assuming that there is no correlation between the prime numbers, the probability that the number n is not divisible for all prime p_i less than $n/2$ (i.e. the probability that n is itself a prime) is given by

$$P(n) \simeq \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\cdots = \prod_{p_i < n}\left(1 - \frac{1}{p_i}\right). \tag{3.B.39}$$

By taking the logarithm of both sides, we have

$$\log P(n) \simeq \sum_{p_i<n} \log\left(1-\frac{1}{p_i}\right) \simeq -\sum_{p_i<n}\frac{1}{p_i}.$$

The last sum over the primes p_i can be written as a sum over the integers by using the probability $P(k)$ and therefore

$$\log P(n) \simeq -\sum_{k}^{n}\frac{P(k)}{k} \simeq \int_1^n \frac{P(k)}{k}\,dk.$$

Taking the derivative with respect to n of both terms we see that $P(n)$ satisfies the differential equation

$$\frac{dP(n)}{dn} = -\frac{P^2(n)}{n},$$

whose solution is the celebrated probability of the prime numbers $P(n) = 1/\log n$.

The key point of this derivation is the iterative nature of the sieve of Eratosthenes. From an algorithmic point of view, this procedure consists of the following steps: (a) once a prime number p_i has been identified, one cancels all its multiples; (b) the next prime p_{i+1} is then the first integer number after p_i that has survived step (a). One then repeats iteratively steps (a) and (b). In light of this observation, any other sieve on the natural numbers implemented with the same rules will produce the same probability law $P(n) \simeq 1/\log n$. We can adopt, for instance, a sieve purely based on probabilistic laws. For instance, if we select the number 2 as a prime number, we can proceed to cancel all next numbers with probability 1/2. After this procedure, we take the first number n_i that has survived this sieve and proceed to cancel all numbers $n > n_i$ with probability $1/n_i$, and so on. After all steps of this procedure have been completed, the set $\mathcal{E}$ of the survived numbers are prime numbers only *on average*, although they follow the same Gauss's law (3.B.38). In other words, the probability that a generic integer number n belongs to the set $\mathcal{E}$ is given by Gauss's law (3.B.38).

The arithmetic situation then looks similar to statistical mechanics, where there are general laws that are easy to establish if one considers a large number of similar ensembles but whose origin may appear mysterious if analyzed on a particular sample.

Note that, for large n, eqn (3.B.39) can be cast as the probability that the next two numbers belong to the same set, i.e.

$$P(n+1) \simeq P(n)\left[1-\frac{1}{n}P(n)\right]. \tag{3.B.40}$$

This equation has the following interpretation. For large n one expects that $P(n+1)$ and $P(n)$ are almost equal, but the difference is that the number n can cancel the next number with probability $1/n$ but it is not true vice versa! The probability that n belongs to the set and simultaneously that cancels $(n+1)$ is given by the product of the probabilities $\frac{1}{n} \times P(n)$. Therefore, the probability that this does not happen is given by $\left[1-\frac{1}{n}P(n)\right]$, i.e. the factor present on the right-hand side of (3.B.40).

It should be noted that (3.B.40) is the equation satisfied by the probability to visit a new site in two-dimensional brownian motion. In fact, in a brownian motion the mean square distance from the origin after n steps is $\overline{r^2} \simeq n$. Therefore, in two-dimensional brownian motion, after n steps the particle is localized in a circle of radius $\sqrt{n}$. Let $P(n)$ be the probability that the particle visits a new site at the n-th step. What is the probability that the particle will visit a new site at the $(n+1)$-th step? Going from n to $(n+1)$, the area A of the circle varies as $\Delta A/A = 1/n$ and the probability that these sites have already been visited is given by the product $P(n) \times \Delta A/A = P(n) \times 1/n$. The probability that these sites have not already been visited is given by $1 - \frac{1}{n}P(n)$. Since for large n, $P(n+1)$ and $P(n)$ must be proportional, we arrive at eqn (3.B.40) and therefore $P(n) \simeq 1/\log n$. The average number of distinct sites that have been visited is given by the total area $A \simeq \pi n$ mutiplied by the probability $P(n) = 1/\log n$, i.e. $S_n \simeq \pi n/\log n$.

We leave as an exercise the elementary derivation of the asymptotic behavior of S_n in $d = 1$ and $d \geq 3$.

References and Further Reading

The mean field approximation is discussed in many books on statistical mechanics. It may be useful to consult:

L.E. Reichl, *Modern Course in Statistical Mechanics*, Arnold Publishers, London, 1980.

The Bethe–Peierls approximation was proposed in the articles:

H.A. Bethe, *Statistical theory of superlattices*, Proc. Roy. Soc. A 150 (1935), 552.

R. Peierls, *Statistical theory of superlattices with unequal concentration*, Proc. Roy. Soc. 154 (1936), 207.

The transfer matrix solution of the one-dimensional gaussian model is discussed in:

L. P. Kadanoff, *Statistical Physics. Statics, Dynamics and Renormalization*, World Scientific, Singapore, 2000.

The spherical model was proposed in the article:

T.H. Berlin, M. Kac, *The spherical model of a ferromagnet*, Phys. Rev. 86 (1952), 821.

The equivalence between the spherical model and a model with $O(n)$ symmetry has been shown by E. Stanley in the paper:

H.E. Stanley, *Spherical model as the limit of infinite spin dimensionality*, Phys. Rev. 176 (1968), 718.

Brownian motion is one of the most important topics of statistical mechanics. The interested reader may find a detailed discussion of this topic in:

C. Itzykson, J.M. Drouffe, *Statistical Field Theory*, Cambridge University Press, Cambridge, 1989.

E.W. Montroll, B.J. West, *Studies in Statistical Mechanics VII, Fluctuation Phenomena*, North Holland, Amsterdam, 1979.

S. Chandrasekhar, *Stochastic problems in physics and astronomy*, Rev. Mod. Phys. 15 (1943), 1.

The magnum opus on brownian motion is the two-volume book

B. D. Hughes, *Random Walks and Random Enviroments*, Oxford Science Publications, Oxford, 1995.

Problems

1. Mean field theory for antiferromagnetic systems

a Prove that the mean field theory can be applied to antiferromagnetic systems defined on a square lattice.

b Generalize this result to all lattices that admit two sublattices such that the next neighbor sites of one are sites of the other and vice versa. This occurs, for instance, in the two-dimensional hexagonal or square lattices.

c What happens if one considers a triangular lattice?

2. Mean field theory for coupled lattices

Consider a system of Ising spins made of two-dimensional square lattices A and B, coupled together. Let J_a and J_b be the coupling constants between next neighbor spins in the lattices A and B respectively, and J the coupling constant betwen the next neighbor spins of the two lattices, as in Fig. 3.13.

a Generalize the mean field theory to this system and show that the spontaneous magnetizations of the two lattices satisfy the coupled system of equations

$$\begin{aligned} M_a &= f(J_a M_a + J M_b) \\ M_b &= f(J_b M_b + J M_a). \end{aligned}$$

Determine the explicit form of the function f.

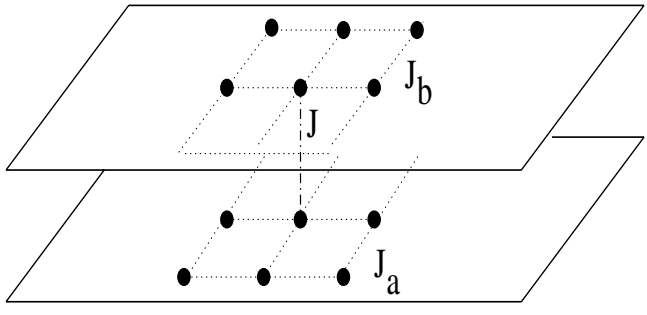


Fig. 3.13 *Layers of Ising spins.*

b Estimate the critical temperature T_c in the mean field approximation.

c Show that, for $T > T_c$, the magnetic susceptibility is expressed by the ratio of a linear and quadratic polynomial in T.

d Discuss the case $J_a = J_b > 0$ while $J < 0$ and study the magnetic susceptibility for these values of the couplings.

e Discuss the limits $J \to \pm\infty$.

3. Spontaneous magnetization at low temperature

Show that, when T is much less than T_c, the mean field theory of a ferromagnet predicts a spontaneous magnetization that differs from its saturation value for terms that are exponentials in $-1/T$.

4. Quantum magnets

Consider the hamiltonian

$$\mathcal{H} = -\frac{1}{2}\sum_{R,R'} J(R-R')\,\vec{S}(R)\cdot\vec{S}(R')$$

where $J(R-R') > 0$ and $\vec{S}(R)$ is the quantum operator of spin S.

a Prove initially the following result: the largest (smallest) diagonal element that a hermitian operator can have is equal to its largest (smallest) eigenvalue.

b Use this result to prove that, for $R \neq R'$, $\langle \vec{S}(R)\cdot\vec{S}(R')\rangle \leq S^2$.

c Let $\mid S\rangle_R$ be the eigenvectors $S_z(R)$ with the maximum eigenvalue

$$S_z(R) \mid S\rangle_R = S \mid S\rangle_R.$$

Prove that the state $\mid 0\rangle = \prod_i \mid S\rangle_R$ is an eigenvector of the hamiltonian with eigenvalue

$$E_0 = -\frac{1}{2}S^2 \sum_{R,R'} J(R-R').$$

Hint. Express the hamiltonian in terms of the ladder operators $S_\pm(R) = (S_x \pm iS_y)(R)$ *and use the condition* $S_+(R) \mid S\rangle_R = 0$.

d Use the result of **a** to show that E_0 is the smallest eigenvalue of the hamiltonian.

5. Critical exponents of the spherical model

Use the equation of state and the other relations discussed in the text to derive the critical exponents of the spherical model.

6. Brownian motion with boundary conditions

Let

$$P(s,t) = \frac{1}{\sqrt{4\Pi\mathcal{D}t}}\,\exp[-s^2/(4\mathcal{D}t)]$$

be the probability distribution in the continuum limit of the one-dimensional brownian motion.

a Find the probability distribution $P_a(s,t)$ when the $s = s_0 > 0$ is an absorbent point, i.e. when $P_a(s_0,t) = 0$ for all $t \geq 0$.

b Find the probability distribution $P_r(s,t)$ when the point $s = s_0 > 0$ is a pure reflecting point, i.e. when it holds the condition $\frac{\partial P_r}{\partial x}(s_0,t) = 0$ for all $t \geq 0$.

Hint. Use $P(s,t)$ and the linearity of the problem to set up the method of solution.

7. Distinct points visited in brownian motion

Give a physical argument for the mean value of the number of distinct sites visited in brownian motion and show that this mean value depends on the dimensionality of the lattice as predicted by eqn (3.B.28).

8. Markov processes

Brownian motion is a particular example of a general class of stochastic processes known as *Markov processes*, characterized by a transition probability $w(i \to j) = w_{ij}$ between the discrete states $\{A\} = \{a_1, a_2, a_3, \ldots, a_n\}$ of a stochastic variable A ($w_{ij} \geq 0$ and $\sum_{j=1}^{n} w_{ij} = 1$). These transitions take place at discrete time steps $t_n = n$. Denoting by $P_i(n)$ the probability to be in the i-th state at time n, it satisfies the recursive equation

$$P_i(n+1) = \sum_{j=1}^{n} w_{ij}\, P_j(n).$$

Using a matrix formalism, it can be expressed as $P(n+1) = WP(n)$.

a Prove that the eigenvalues of the matrix W satisfy the condition $\mid \lambda_i \mid \leq 1$.

b Show that the system reaches an equilibrium distribution $P_i(\infty)$ for $t \to \infty$ that is independent of the initial condition if and only if the matrix W has only one eigenvalue of modulus 1.

c Assuming that the conditions of the point **b** are satisfied, prove that

$$\lim_{n\to\infty} (W)^n = M = \begin{pmatrix} m_1, m_2, \ldots, m_n \\ m_1, m_2, \ldots, m_n \\ . \\ . \\ . \\ m_1, m_2, \ldots, m_n \end{pmatrix}$$

with $\lim_{n\to\infty} P_i(n) = m_i$

9. Brownian motion on a ring

Consider brownian motion on a ring of N sites, with a transition rate to next neighbor sites equal to $1/2$. Let $P_n(s)$ be the probability to find the walker at site s at time n ($s = 1, 2, \ldots, N$). Show that if N is an odd number, there is a unique stationary probability distribution $n \to \infty$. Vice versa, if N is an even number, for $n \to \infty$, the distribuition probability can oscillate between two different probability distributions.

10. Langevin equation and brownian motion

Consider a particle of mass m in motion in a fluid (for simplicity we consider one-dimensional motion), subjected to a frictional force proportional to the velocity and a random force $\eta(t)$ due to the random fluctuations of the fluid density. Denoting by $x(t)$ and $v(t)$ the position and the velocity of the particle at time t, the equations of motion of the particle are

$$\frac{dv(t)}{dt} = -\frac{\gamma}{m}v(t) + \frac{1}{m}\eta(t),$$
$$\frac{dx(t)}{dt} = v(t),$$

where $\gamma > 0$ is the friction coefficient. Assume that $\eta(t)$ is a random variable, with zero mean and delta-correlated

$$\langle \eta(t) \rangle_\eta \, = \, 0, \qquad \langle \eta(t_1)\eta(t_2) \rangle_\eta \, = \, 2\gamma k_B T \, \delta(t_1 - t_2)$$

where k_B is the Boltzmann constant, T is the temperature, and the average $\langle \, \rangle_\eta$ is with respect to the probability distribution of the stochastic variable $\eta(t)$.

a Let x_0 and v_0 be the position and velocity of the particle at $t = 0$. Integrating the equations of motion and taking the average with respect to η, show that the correlation function of the velocity is

$$\langle v(t_2)v(t_1) \rangle_\eta \, = \, \left(v_0^2 - \frac{k_B T}{m} \right) e^{-(\gamma/m)(t_1+t_2)} + \frac{k_B T}{m} \, e^{-(\gamma/m)(t_2-t_1)}.$$

with $t_2 > t_1$.

b Compute the variance of the displacement and show that

$$\langle (x(t) - x_0)^2 \rangle_\eta = \frac{m^2}{\gamma} \left(v_0^2 - \frac{k_B T}{m} \right) \left(1 - e^{-(\gamma/m)t} \right)^2 + \frac{2k_B T}{\gamma} \left[t - \frac{m}{\gamma} \left(1 - e^{-(\gamma/m)t} \right) \right].$$

c Assuming that the particle is in thermal equilibrium, we can now average over all possible initial velocities v_0. Let's denote this thermal average by $\langle \, \rangle_T$. By the equipartion theorem we have $\langle v_0^2 \rangle_T \, = \, k_B T/m$. Show that, for $t \gg m/\gamma$, the thermal average of the variance of the displacement becomes

$$\langle \langle (x(t) - x_0)^2 \rangle_\eta \rangle_T \, \simeq \, (2k_B T/\gamma)t.$$

This page intentionally left blank

Part II

Bidimensional Lattice Models

This page intentionally left blank

4

Duality of the Two-dimensional Ising Model

Being dual is in the nature of things.

Elias Canetti

In this chapter we will begin our study of the Ising model on the two-dimensional lattice. In two dimensions the model has a phase transition, with critical exponents that have different values from those obtained in the mean field approximation. For this reason, it provides an important example of critical phenomena. As we will see in great detail in this chapter and in the next, among all exactly solved models of statistical mechanics, the two-dimensional Ising model is not only the one that has been most studied but it is also the model that has given a series of deep mathematical and physical results. Many solutions of the model stand out for the ingenious methods used, such as the theory of determinants, combinatorial approaches, Grassmann variables, or elliptic functions. Many results have deeply influenced the understanding of critical phenomena and have strongly stimulated new fields of research. Ideas that have matured within the study of the two-dimensional Ising model, such as the duality between its high- and low-temperature phases, have been readily generalized to other systems of statistical mechanics and have also found important and fundamental applications in other important areas such as, for instance, quantum field theory. Equally fundamental is the discovery that in the vicinity of the critical point, the dynamics of the model can be described from the relativistic Dirac equation for Majorana fermions.

This chapter is devoted to the study of some properties of the model that can be established by means of elementary considerations. We will discuss, in particular, the argument by Peierls that permits us to show the existence of a phase transition in the model. We will also present the duality relation that links the expressions of the partition functions in the low- and high-temperature phase of a square lattice, and the partition functions of the triangle and hexagonal-lattices. In the last case, it is necessary to make use of an identity, known as the *star–triangle equation*, that will be useful later on to study the commutativity properties of the transfer matrix. At the end of the chapter, we will also discuss the general formulation of the duality transformations for lattice statistical models.

4.1 Peierls's Argument

In 1936 R. Peierls published an article with the title *On the Model of Ising for the Ferromagnetism* in which he proved that the Ising model in two or higher dimensions has a low-temperature region in which the spontaneous magnetization is different from zero. Since at high temperature the system is disordered, it follows that there must exist a critical value of the temperature at which a phase transition takes place.

Peierls's argument starts with the initial observation that to each configuration of spins there corresponds a set of closed lines that separate the regions in which the spins assume values $+1$ from those in which they assume values -1, as shown in Fig. 4.1. If it is possible to prove that at sufficiently low temperatures the mean value of the regions enclosed by the closed lines is only a small fraction of the total volume of the system, one has proved that the majority of the spins is prevalently in the state in which there is a spontaneous magnetization.

There are several versions of the original argument given by Peierls. The simplest generalizes the argument already used in the one-dimensional case (see Chapter 2, Section 2.1) and concerns the stability of the state with a spontaneous magnetization. Let's consider the two-dimensional Ising model at low temperatures and suppose that it is in the state of minimal energy in which all the spins have values $+1$. The thermal fluctuations create domains in which there are spin flips, such as the domain in Fig. 4.1. The creation of such domains clearly destabilizes the original ordered state. There is an energetic cost to the creation of the domain shown in Fig. 4.1, given by

$$\Delta E = 2J\,L, \tag{4.1.1}$$

where L is the total length of the curve. There are, however, many ways of creating a closed curve of a given total perimeter L. In fact, the domain in which the spins are flipped can be placed everywhere in the lattice and moreover can assume different shapes. To estimate the number of such configurations, imagine that the closed line is created by a random motion on the lattice of total number of steps equal to L. If we

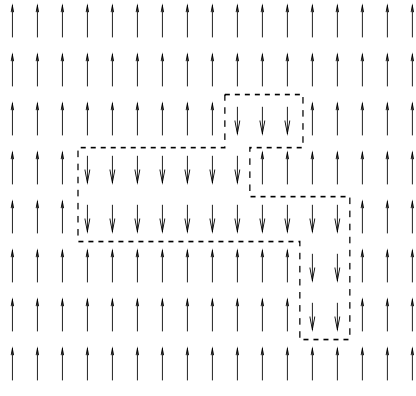


Fig. 4.1 *Closed lines that enclose a region with a flipped value of the spin.*

assume that at each step of this motion there are only two possibilities,[1] we have 2^L ways of drawing a closed curve of length L. The corresponding variation of the entropy is given by

$$\Delta S \,=\, k\,\ln(2^L). \tag{4.1.2}$$

Hence the total variation of the free energy associated to the creation of such a domain is

$$\begin{aligned}\Delta F = \Delta E - T\Delta S \,&=\, 2JL - kT\ln(2^L) \\ &= L(2J - kT\ln 2).\end{aligned} \tag{4.1.3}$$

Therefore the system is stable with respect to the creation of such domains of arbitrary length L (i.e. $\Delta F \geq 0$) if

$$T \leq T_c \,=\, \frac{2J}{k\ln 2} \,=\, 2.885\,\frac{J}{k}. \tag{4.1.4}$$

Note that such an estimate is surprisingly close to the exact value of the critical temperature $T_c = 2.269...J/k$ that we will determine in the next section.

4.2 Duality Relation in Square Lattices

Peierls's argument shows that the two-dimensional Ising model has two different phases: the high-temperature phase in which the system is disordered and the low-temperature phase in which the system is ordered, with a non-zero spontaneous magnetization. The exact value of the critical temperature at which the phase transition happens was first determined by H.A. Kramers and G.H. Wannier by using a duality relation between the high- and the low-temperature partition functions.[2] The self-duality of the two-dimensional Ising model on a square lattice is one of its most important properties, with far-reaching consequences on its dynamics. To prove it, we need to study the series expansions of the high/low-temperature phase of the model. We will see that these expansions have an elegant geometrical interpretation in terms of a counting problem of the polygons that can be drawn on a lattice. In the next section we will consider the square lattice and, in later sections, the triangle and hexagonal lattices.

4.2.1 High-temperature Series Expansion

Consider a square lattice $\mathcal{L}$ with M horizontal links and M vertical links. In the thermodynamical limit $M \to \infty$, M coincides with the total number N of the lattice sites. In the following we will consider a Hamiltonian with different coupling constants,

[1] On a square lattice, starting from a given site, one can move in four different directions. However, taking four instead of two as possible directions of the motion gives an upper estimate of the entropy, since it does not take into account that the final curve is a closed contour.

[2] The self-duality of the model that we are going to discuss only holds in the absence of an external magnetic field.

along the horizontal and vertical directions. Let J and J' be these coupling constants, respectively. For the partition function of the model at zero magnetic field we have

$$Z_N = \sum_{\{\sigma\}} \exp\left[K \sum_{(i,j)} \sigma_i \sigma_j + L \sum_{(i,k)} \sigma_i \sigma_k\right], \tag{4.2.1}$$

where the first sum is on the spins along the horizontal links and the second sum along the vertical links, with

$$K = \beta J; \quad L = \beta J'.$$

By using the identity

$$\exp\left[x \sigma_i \sigma_l\right] = \cosh x \,(1 + \sigma_i \sigma_l \tanh x), \tag{4.2.2}$$

the partion function can be written as

$$Z_N = (\cosh K \, \cosh L)^M \sum_{\{\sigma\}} \prod_{(i,j)} (1 + v \sigma_i \sigma_j) \prod_{(i,k)} (1 + w \sigma_i \sigma_k), \tag{4.2.3}$$

with

$$v = \tanh K; \quad w = \tanh L.$$

Both parameters v and w are always less than 1 for all values of the temperature, except for $T = 0$ when their value is $v = w = 1$. In particular, they are small parameters in the high-temperature phase and it is natural to look for a series expansion of the partition function near $T = \infty$.

If we expand the two products in (4.2.3), we have 2^{2M} terms, since there are $2M$ factors (one for each segment), and each of them has two terms. We can set up a graphical representation for this expansion associating a line drawn on the horizontal link (i, j) to the factor $v \sigma_i \sigma_j$ and a line on the vertical link (i, k) to the factor $w \sigma_i \sigma_k$. No line is drawn if there is instead the factor 1. Repeating this operation for the 2^{2M} terms, we can establish a correspondence between these terms and a graphical configuration on the lattice $\mathcal{L}$. The generic expression of these terms is

$$v^r w^s \sigma_1^{n_1} \sigma_2^{n_2} \sigma_3^{n_3} \ldots$$

where r is the total number of horizontal lines, s the total number of vertical lines, while n_i is the number of lines where i is the final site. It is now necessary to sum over all spins of the lattice in order to obtain the partition function. Since each spin σ_i assumes values ± 1, we have a null sum unless all $n_1, n_2, \ldots, n_N$ are even numbers and, in this case, the result is $2^N v^r w^s$. Based on these considerations, the partition function can be expressed as

$$Z_N = 2^N (\cosh K \cosh L)^M \sum_P v^r w^s, \tag{4.2.4}$$

where the sum is over all the line configurations on $\mathcal{L}$ with an even number of lines at each site, i.e. *all closed polygonal lines* P of the lattice $\mathcal{L}$. Therefore, apart from a

prefactor, the partition function is given by the geometrical quantity

$$\Phi(v,w) \;=\; \sum_{P} v^r w^s. \tag{4.2.5}$$

It is easy to compute the first terms of this function. The first term is equal to 1 and corresponds to the case in which there are no polygons on the lattice. The second term corresponds to the smallest closed polygon on the lattice $\mathcal{L}$, i.e. a square with unit length, as shown in Fig. 4.2. The number of such squares is equal to N, since they can be placed on any of the N sites of the lattice. Each of them has a weight $(vw)^2$, hence the second term of the sum (4.2.5) is equal to $N(vw)^2$. The next closed polygonal curve is a rectangle of six sides: there are two kinds of them, as shown in Fig. 4.3, each with a degeneracy equal to N, and width v^4w^2 for the first and v^2w^4 for the second.

Using the first terms, the function $\Phi(v,w)$ is given by

$$\Phi(v,w) \;=\; 1 + N(vw)^2 + N(v^4w^2 + v^2w^4) + \cdots \tag{4.2.6}$$

The computation of the next terms becomes rapidly more involved although it can be clearly performed in a systematic way: presently, the first 40 terms of such a series are known. For our purposes it is not necessary to introduce all these terms, since the duality properties can be established just by exploiting the geometrical nature of the sum (4.2.5).

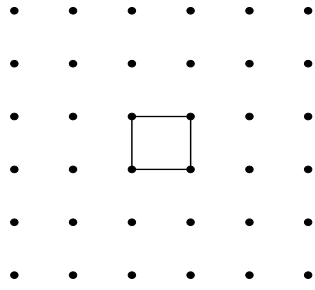


Fig. 4.2 *Second term of the high-temperature expansion.*

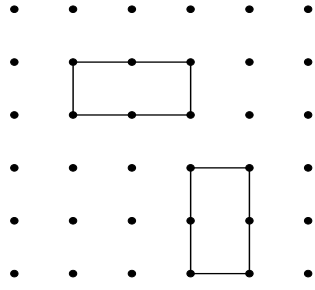


Fig. 4.3 *Third term of the high-temperature expansion.*

4.2.2 Low-temperature Series Expansion

In the low-temperature phase, according to Peierls's argument, the spins tend to align one with another. The series expansion of the partition function in this phase can be obtained as follows. For a given configuration of the spins, let r and s be the numbers of vertical and horizontal links in which the two adjacent spins are antiparallel. Since M is the total number of vertical links as well as of the horizontal ones, we have $(M - r)$ vertical links and $(M - s)$ horizontal links in which the adjacent spins are parallel. The contribution to the partition function of such a configuration is

$$\exp\left[K(M - 2s) + L(M - 2r)\right].$$

Besides a constant, this expression depends only on the number of links in which the spins are antiparallel. These segments will be called *antiparallel links*.

It is now convenient to introduce the concept of a *dual lattice*. This notion, which is familiar in crystallography, has already been met in the discussion of the four-color problem (see Appendix C of Chapter 2). For any planar lattice $\mathcal{L}$, we can define another lattice $\mathcal{L}_D$ that is obtained by placing its sites at the center of the original lattice $\mathcal{L}$ and joining pairwise those relative to adjacent faces, i.e. those sharing a common segment. It is easy to see that the dual lattice of a square lattice is also a square lattice, simply displaced by a half-lattice space with respect to the original one (see Fig. 4.4), while the dual lattice of a triangular lattice is a hexagonal one and vice versa.

Given the geometrical relation between the dual and the original lattices, it is easy to see that the spins can be equivalently regarded as defined on the sites of the original lattice $\mathcal{L}$ or at the center of the faces of the dual lattice $\mathcal{L}_D$. This allows us to introduce a useful graphical formalism. Given a configuration, we can associate to its antiparallel links a set of lines of the dual lattice by the following rule: if two next neighbor spins are antiparallel, then draw a line along the segment of $\mathcal{L}_D$ that crosses them, draw no line if they are parallel. By applying this rule, on the dual lattice $\mathcal{L}_D$ there will be r horizontal lines and s vertical lines. However, it is easy to see that there should always

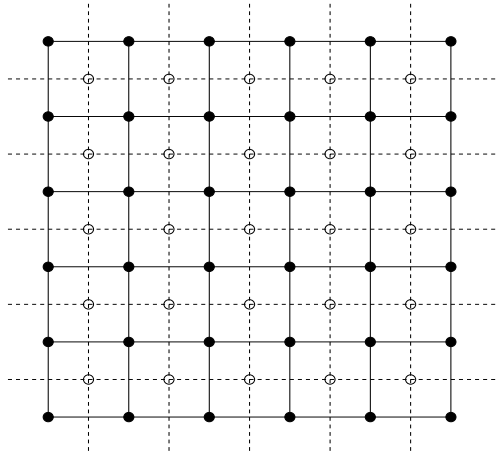


Fig. 4.4 *Dual square lattices.*

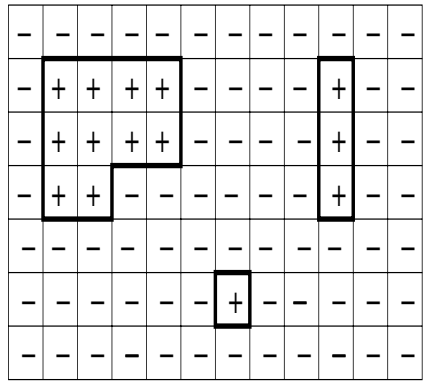


Fig. 4.5 *Polygons that separate the domains with spins* $+1$ *and* -1.

be an even number of lines passing through each site, since there is an even number of next successive changes among the adjacent faces. The drawn lines must therefore form closed polygons on the dual lattice $\mathcal{L}_D$, as illustrated in Fig. 4.5.

It is evident that the closed polygons that have been obtained in this way are nothing else that the perimeters of the different magnetic domains where, inside them, all spins are aligned in the same direction. Since for any given set of polygons there are two corresponding configurations (one obtained from the other by flipping all the spins), the partition function can be written as

$$Z_N \;=\; 2\exp[M(K+L)] \sum_{\tilde{P}} \exp[-(2Lr+2Ks)], \tag{4.2.7}$$

where the sum is over all closed polygons $\tilde{P}$ on the dual lattice $\mathcal{L}_D$. This is the low-temperature expansion, because when $T \to 0$, both K and L are quite large and the dominant terms are given by small values of r and s. Therefore, also in this case the partition function is expressed by a geometrical quantity

$$\tilde{\Phi}\left(e^{-2L}, e^{-2K}\right) \;=\; \sum_{\tilde{P}} \exp[-(2Lr+2Ks)]. \tag{4.2.8}$$

Consider the first terms of this series. The first term is equal to 1 and corresponds to the situation in which all spins assume the same value. The second term corresponds to the configuration in which there is only one spin flip: in this case there are two horizontal antiparallel links and two vertical antiparallel links that altogether form a square. The degeneracy of this term is equal to N, since the spin that has been flipped can be placed on any of the N sites of the lattice. The next term is given by the rectangle with six segments that can be elongated either horizontally or vertically: these rectangles correspond to next neighbor spins that are antiparallel to all other spins of the lattice. Taking into account the degeneracy N and the orientation of the rectangle, the contribution of this term to the partition function is $N(e^{-4L-8K} + e^{-8L-4K})$. With these first terms, the function $\tilde{\Phi}(e^{-2L}, e^{-2K})$ is expressed by

$$\tilde{\Phi}\left(e^{-2L}, e^{-2K}\right) \;=\; 1 + Ne^{-4L-4K} + N(e^{-4L-8K} + e^{-8L-4K}) + \cdots \tag{4.2.9}$$

From what was said above, it should now be clear that all terms of the function $\tilde{\Phi}$ have the same origin as those of the function Φ.

4.2.3 Self-duality

In the last two sections we have shown that the partition function of the two-dimensional Ising model on a square lattice can be expressed in two different series expansions, one that holds in the high-temperature phase, the other in the low-temperature phase, given in eqns (4.2.4) and (4.2.7), respectively. The final expressions involve a function that has a common geometric nature, i.e. a sum over all the polygonal configurations that can be drawn on the original lattice and its dual. For finite lattices, $\mathcal{L}$ and $\mathcal{L}_D$ differ only at the boundary. In the thermodynamical limit this difference disappears and the two expressions can be obtained one from the other simply by a change of variables. For $N \to \infty$ one has $M/N = 1$: substituting K and L in eqn (4.2.5) with $\tilde{K}$ and $\tilde{L}$ given by

$$\tanh \tilde{K} = e^{-2L}; \quad \tanh \tilde{L} = e^{-2K}, \tag{4.2.10}$$

and comparing with eqn (4.2.8), we have in fact

$$\tilde{\Phi}\left(e^{-2\tilde{K}}, e^{-2\tilde{L}}\right) = \Phi(v, w). \tag{4.2.11}$$

This implies the following identity for the partition function

$$\frac{Z_N[K, L]}{2^N (\cosh K \cosh L)^N} = \frac{Z_N[\tilde{K}, \tilde{L}]}{2 \exp[N(\tilde{K} + \tilde{L})]}. \tag{4.2.12}$$

Equation (4.2.10) can be expressed in a more symmetrical form:

$$\sinh 2\tilde{K} \sinh 2L = 1; \quad \sinh 2\tilde{L} \sinh 2K = 1. \tag{4.2.13}$$

Analogously, eqn (4.2.12) can be written as

$$\frac{Z_N[K, L]}{(\sinh 2K \sinh 2L)^{N/4}} = \frac{Z_N[\tilde{K}, \tilde{L}]}{(\sinh 2\tilde{K} \sinh 2\tilde{L})^{N/4}}. \tag{4.2.14}$$

These equations show the existence of a symmetry of the two-dimensional Ising model and establish the mapping between high- and low-temperature phases of the model. Large values of K and L are equivalent to small values of $\tilde{K}$ and $\tilde{L}$, and vice versa large values of $\tilde{K}$ and $\tilde{L}$ correspond to small values of K and L. It must be stressed that this correspondence between the two phases can also be useful from a computational point of view.

We can now identify the critical point. Let's consider first the isotropic case, i.e. $K = L$ and, correspondingly, $\tilde{K} = \tilde{L}$. At the critical point the partition function presents a divergence: assuming that this happens at the value K_c, the same should happen also at $\tilde{K} = K_c$ thanks to eqn (4.2.14). These two values can be different but, making the further hypothesis that there is only one critical point – a hypothesis that

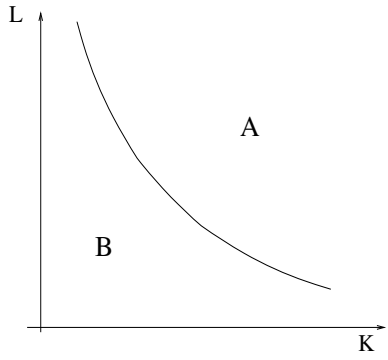


Fig. 4.6 *Critical curve.*

is fully justified from the physical point of view – these two values must coincide and the critical point is thus identified by the condition

$$\sinh 2K_c \,=\, 1; \qquad T_c^{square} \,=\, 2.26922...J. \tag{4.2.15}$$

The arguments presented above were given originally by Kramers and Wannier.

Let us consider now the general case in which there are two coupling constants. Note that combining eqn (4.2.13), we have

$$\sinh 2K \sinh 2L \,=\, \frac{1}{\sinh 2\tilde{K} \sinh 2\tilde{L}}. \tag{4.2.16}$$

This equation implies that, under the mapping $(K, L) \,\rightarrow\, (\tilde{K}, \tilde{L})$, the region A in Fig. 4.6 is transformed into the region B and vice versa, leaving invariant the points along the curve

$$\sinh 2K \sinh 2L \,=\, 1. \tag{4.2.17}$$

If there is a line of fixed points in A, there should be another line of fixed points also in B. Assuming that there is only one line of fixed points, this is expressed by eqn (4.2.17). Therefore this is the condition that ensures the criticality of the Ising model with different coupling constants along the horizontal and vertical directions. This equation plays an important role both in the solution proposed by Baxter for the Ising model and in the discussion of its hamiltonian limit.

4.3 Duality Relation between Hexagonal and Triangular Lattices

The duality transformation of the square lattice can be generalized to other lattices. In this section we discuss the mapping between the low- and high-temperature phases of the Ising model defined on the triangular and hexagonal lattices shown in Fig. 4.7.

Let us introduce the coupling constants K_i and L_i $(i \,=\, 1, 2, 3)$ relative to the triangle and hexagonal lattices, respectively, as shown in Fig. 4.8. In the absence of a

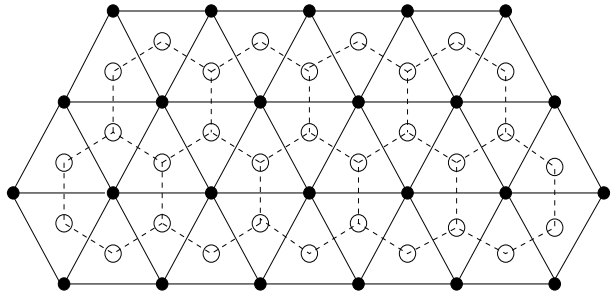


Fig. 4.7 *Dual lattices: hexagonal and triangular lattices.*

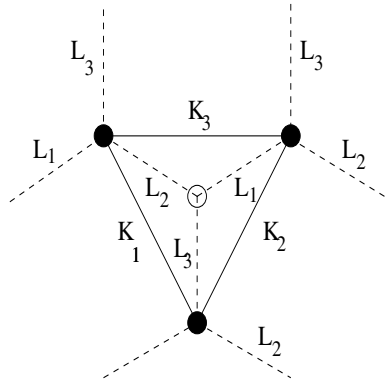


Fig. 4.8 *Coupling constants on the triangular and hexagonal lattices.*

magnetic field, the partition function of the hexagonal lattice is given by

$$Z_N^{\mathrm{H}}(\mathcal{L}) \;=\; \sum_{\{\sigma\}} \exp\left[\mathcal{L}_1 \sum \sigma_l\sigma_i + \mathcal{L}_2 \sum \sigma_l\sigma_j + \mathcal{L}_3 \sum \sigma_l\sigma_k\right], \tag{4.3.1}$$

with $\mathcal{L}_i = L_i/kT$. In the exponential term, the sums refer to all next neighbor pairs of spins along the three different directions of the hexagonal lattice. Similarly, in the absence of the magnetic field, we can write the partition function on the triangular lattice as

$$Z_N^{\mathrm{T}}(\mathcal{K}) \;=\; \sum_{\{\sigma\}} \exp\left[\mathcal{K}_1 \sum \sigma_l\sigma_i + \mathcal{K}_2 \sum \sigma_l\sigma_j + \mathcal{K}_3 \sum \sigma_l\sigma_k\right], \tag{4.3.2}$$

with $\mathcal{K}_i = K_i/kT$ and the sums in the exponentials on all next neighbor pairs of spins in the three different directions of the triangular lattice.

Let's consider the high-temperature expansion of the partition function on the triangular lattice. Put $v_i = \tanh \mathcal{K}_i$, we have

$$Z_N^{\mathrm{T}}(\mathcal{K}) \;=\; (2\cosh\mathcal{K}_1\,\cosh\mathcal{K}_2\,\cosh\mathcal{K}_3) \sum_P v_1^{r_1}\, v_2^{r_2}\, v_3^{r_3}, \tag{4.3.3}$$

where the sum is over all closed polygons on the triangular lattice, with the number of sides equal to r_i $(i = 1, 2, 3)$ along the three different directions.

Consider now the low temperature expansion of the partition function on the hexagonal lattice. This is obtained by drawing the lines corresponding to the antiparallel links on the dual lattice. Since the triangular lattice of N sites is the dual of the hexagonal lattice with $2N$ sites, in this case we have[3]

$$Z_{2N}^{\mathrm{H}}(\mathcal{L}) \;=\; e^{[N(\mathcal{L}_1+\mathcal{L}_2+\mathcal{L}_3)]} \sum_P \exp[-2\mathcal{L}_1 r_1 + \mathcal{L}_2 r_2 + \mathcal{L}_3 r_3], \tag{4.3.4}$$

where the sum is over the closed polygons of the triangular lattice with the number of sides r_i $(i = 1, 2, 3)$ along the three directions.

Since in both expressions there is the same geometrical function given by the sum over polygons drawn on the triangular lattice, imposing

$$\tanh \mathcal{K}_i^* \;=\; \exp[-2\mathcal{L}_i], \quad i = 1, 2, 3 \tag{4.3.5}$$

the two partition functions are related as

$$Z_{2N}^{\mathrm{H}}(\mathcal{L}) \;=\; (2a_1 a_2 a_3)^{N/2}\, Z_N^{\mathrm{T}}(\mathcal{K}^*), \tag{4.3.6}$$

where

$$a_i \;=\; \sinh 2\mathcal{L}_i = 1/\sinh 2\mathcal{K}_i^*, \quad i = 1, 2, 3.$$

The relation (4.3.5) can be written in a more symmetrical way as

$$\sinh 2\mathcal{L}_i \, \sinh 2\mathcal{K}_i^* \;=\; 1. \tag{4.3.7}$$

As in the square lattice, the duality relation (4.3.7) implies that when one of the coupling constant is small, the other is large and vice versa. However, the duality relation alone cannot determine in this case the critical temperature of the two lattices, since they are not self-dual. Fortunately, there exists a further important identity between the coupling constants of the two lattices that permits us to identify the singular points of the free energies of both models. This identity is the *star–triangle identity* and, because of its importance, it is worth a detailed discussion.

4.4 Star–Triangle Identity

The star–triangle identity plays an important role in the two-dimensional Ising model. In addition to the exact determination of the critical temperature for triangular and hexagonal lattices, this identity also enables us to establish the commutativity of the transfer matrix of the model for special values of the coupling constants. This aspect will be crucial for the exact solution of the model discussed in Chapter 6.

To prove such an identity, first observe that the sites of the hexagonal lattice split into two classes, i.e. the hexagonal lattice is bipartite. The sites of type A interact only with those of type B and vice versa, while there is no direct interaction between sites of the same type (see Fig. 4.9). The generic term that enters the sum in the partition

[3]For large N, the number of links along each of the three directions is equal to N.

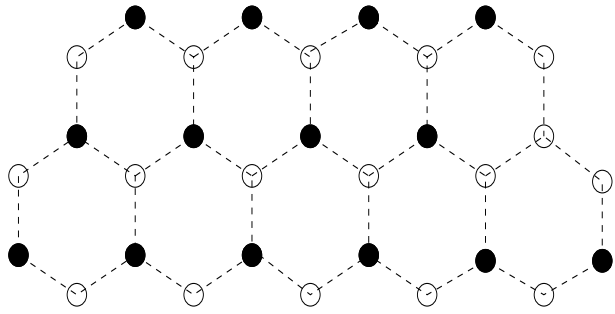


Fig. 4.9 *Bipartition of the hexagonal lattice: site of type A (black sites) and type B (white sites).*

function (4.3.1) can be written as

$$\prod_b W(\sigma_b;\sigma_i,\sigma_j,\sigma_k), \tag{4.4.1}$$

where the product is over all sites of type B and the above quantity is expressed by the Boltzmann weight

$$W(\sigma_b;\sigma_i,\sigma_j,\sigma_k) \,=\, \exp\left[\sigma_b(\mathcal{L}_1\sigma_i + \mathcal{L}_2\sigma_j + \mathcal{L}_3\sigma_k\right]. \tag{4.4.2}$$

Since each spin of type B appears only once in (4.4.1), it is simple to sum on them in the expression of the partition function, with the result

$$Z_N^{\mathrm{E}}(\mathcal{L}) \,=\, \sum_{\sigma_a} \prod_{i,j,k} w(\sigma_i,\sigma_j,\sigma_k), \tag{4.4.3}$$

where

$$w(\sigma_i,\sigma_j,\sigma_k) \,=\, \sum_{\sigma_b=\pm 1} W(\sigma_b;\sigma_i,\sigma_j,\sigma_k) \,=\, 2\,\cosh(\mathcal{L}_i\sigma_i + \mathcal{L}_2\sigma_j + \mathcal{L}_3\sigma_k). \tag{4.4.4}$$

The value of each spin is ± 1 and using the identity

$$\cosh[\mathcal{L}\sigma] \,=\, \cosh L, \qquad \sinh[\mathcal{L}\sigma] \,=\, \sigma\,\sinh L$$

we have

$$\begin{aligned} w(\sigma_i,\sigma_j,\sigma_k) = {} & c_1c_2c_3 + +\sigma_j\sigma_k c_1 s_2 s_3 \\ & + \sigma_i\sigma_j s_1 s_2 c_3 \sigma_i\sigma_k s_1 c_2 s_3, \end{aligned} \tag{4.4.5}$$

where we have defined

$$c_i \,\equiv\, \cosh\mathcal{L}_i, \qquad s_i \,\equiv\, \sinh\mathcal{L}_i.$$

It is important to note that the quantity $w(\sigma_i,\sigma_j,\sigma_k)$ can be written in such a way to be proportional to the Boltzmann factor of the triangular lattice! This means that

there should exist some parameters $\mathcal{K}_i$ and a constant $\mathcal{D}$ such that

$$w(\sigma_i, \sigma_j, \sigma_k) \;=\; \mathcal{D}\, \exp\left[\mathcal{K}_1 \sigma_j \sigma_k + \mathcal{K}_2 \sigma_i \sigma_k + \mathcal{K}_3 \sigma_j \sigma_k\right]. \tag{4.4.6}$$

These parameters can be determined by expanding the exponential as

$$\exp[x \sigma_a \sigma_b] \;=\; \cosh\, x + \sigma_a \sigma_b \sinh\, x,$$

and comparing with eqn (4.4.5). Doing so, we obtain the important result that the products $\sinh 2\mathcal{L}_i \sinh 2\mathcal{K}_i$ are all equal

$$\sinh 2\mathcal{L}_1 \sinh 2\mathcal{K}_1 \;=\; \sinh 2\mathcal{L}_2 \sinh 2\mathcal{K}_2 \;=\; \sinh 2\mathcal{L}_3 \sinh 2\mathcal{K}_3 \;\equiv\; h^{-1} \tag{4.4.7}$$

with the constant h equal to

$$h \;=\; \frac{(1-v_1^2)(1-v_2^2)(1-v_3^2)}{4\left[(1+v_1v_2v_3)(v_1+v2v_3)(v_2+v1v_3)(v_3+v1v_2)\right]^{1/2}}, \tag{4.4.8}$$

where $v_i = \tanh \mathcal{K}_i$, while the constant $\mathcal{D}$ is expressed by

$$\mathcal{D}^2 \;=\; 2h\, \sinh 2\mathcal{L}_1 \sinh 2\mathcal{L}_2 \sinh 2\mathcal{L}_3.$$

The identity (4.4.6) admits a natural graphical interpretation: as shown in Fig. 4.8, summing over the spin of type B at the center of the hexagonal lattice (the one at the center of the star), a direct interaction is generated between the spins of type A placed at the vertices of a triangle. In this way one can switch between the Boltzmann factor of the star of the hexagonal lattice and the Boltzmann factor of the triangular lattice.

4.5 Critical Temperature of Ising Model in Triangle and Hexagonal Lattices

By using the star–triangle identity, it is now easy to determine the critical temperatures of the Ising model on triangular and hexagonal lattices. In fact, substituting the identity (4.4.6) in (4.4.3), the consequent expression is precisely the partition function of the Ising model on a triangular lattice made of $N/2$. Hence, rescaling $N \to 2N$, one has

$$Z^{\mathrm{H}}_{2N}(\mathcal{L}) \;=\; \mathcal{D}^N\, Z^{\mathrm{T}}_N(\mathcal{K}). \tag{4.5.1}$$

Using this equation, together with the duality relation (4.3.6), we obtain a relation that involves the partiton function alone of the triangular lattice

$$Z^{\mathrm{T}}_N(\mathcal{K}) \;=\; h^{-N/2}\, Z^{\mathrm{T}}_N(\mathcal{K}^*), \tag{4.5.2}$$

with

$$\sinh 2\mathcal{K}_i^* \;=\; h\, \sinh 2\mathcal{K}_i, \quad\; i = 1, 2, 3, \tag{4.5.3}$$

and h given in (4.4.8). Thanks to (4.5.3), there is a one-to-one correspondance between the point $(\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3)$ (relative to the high-temperature phase of the model) and the

point $(\mathcal{K}_1^*, \mathcal{K}_2^*, \mathcal{K}_3^*)$ (relative to the low-temperature phase). If, in the space of the coupling constants, there is a line of fixed points under this mapping, this clearly corresponds to the value $h = 1$. For equal couplings ($\mathcal{K}_1 = \mathcal{K}_2 = \mathcal{K}_3 \equiv \mathcal{K}$), from (4.4.8) we have the equation

$$\frac{(1-v^2)^3}{4\left[(1+v^3)v^3(1+v)^3\right]^{1/2}} = 1, \tag{4.5.4}$$

with $v = \tanh \mathcal{K}$. Taking the square of both terms of this equation and simplifying the expression, one arrives at

$$(1+v)^4(1+v^2)^3(v^2-4v+1) = 0.$$

The only solution that also satisfies (4.5.4) and has a physical meaning is given by

$$v_c = 2 - \sqrt{3}.$$

This root determines the critical temperature of the homogeneous triangular lattice

$$\tanh \frac{K}{kT_c} = 2 - \sqrt{3},$$

or, equivalently

$$\sinh \frac{2K}{kT_c} = \frac{1}{\sqrt{3}}. \tag{4.5.5}$$

Numerically

$$T_c^{tr} = 3.64166...K. \tag{4.5.6}$$

Using eqn (4.3.7) we can obtain the critical temperature of the Ising model on a homogeneous hexagonal lattice

$$\sinh \frac{2L}{kT_c} = \sqrt{3}. \tag{4.5.7}$$

Its numerical value is given by

$$T_c^{hex} = 1.51883...L. \tag{4.5.8}$$

It is interesting to compare the value of the critical temperatures (4.5.6) and (4.5.8) with the critical temperature of the square lattice $T_c^{square} = 2.26922J$, given by eqn (4.2.15). At a given coupling constant, the triangular lattice is the one with the higher critical temperature, followed by the square lattice, and then the hexagonal lattice. The reason is simple: the triangular lattice has the higher coordination number, $z = 6$, the hexagonal lattice has the lower coordination number, $z = 3$, while the square lattice is in between the two, with $z = 4$. The higher number of interactions among the spins of the triangular lattice implies that such a system tends to magnetize at higher temperatures than those of the other lattices.

4.6 Duality in Two Dimensions

In the previous sections we showed that the duality property of the Ising model, both for the square lattice and the hexagonal/triangular lattices, can be established on the basis of a geometrical argument, i.e. counting the closed polygons on the original lattice and its dual. However, the duality properties of a statistical model can be characterized in a purely algebric way by considering a particular transformation of the statistical variables entering the partition function. A particularly instructive example is the following. Consider the expression

$$Z(\beta) \,=\, \beta^{1/4} \sum_{n=-\infty}^{\infty} e^{-\pi\beta n^2}. \tag{4.6.1}$$

This can be interpreted as the partition function of a quantum system with energy levels given by $E_n = \pi n^2$. This expression is obviously useful for determining the numerical value of the partition function in the low-temperature phase ($\beta \gg 1$), since in this regime the sum is dominated by the first terms. In the high-temperature phase ($\beta \ll 1$), the situation is rather different and many terms are actually needed to reach a sufficient degree of accuracy. However, using the Poisson resummation formula discussed in Appendix 4B, it is easy to see that we have

$$\begin{aligned} Z(\beta) \,&=\, \beta^{1/4} \sum_{n=-\infty}^{\infty} e^{-\pi\beta n^2} \\ &= \beta^{1/4} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} dx\, e^{-\pi\beta x^2}\, e^{2\pi i m x} \\ &= \beta^{-1/4} \sum_{m=-\infty}^{\infty} e^{-\pi m^2/\beta}. \end{aligned} \tag{4.6.2}$$

Hence this partition function satisfied the important duality relation

$$Z(\beta) \,=\, Z\left(\frac{1}{\beta}\right). \tag{4.6.3}$$

In view of this identity, the partition function in the high-temperature phase can be efficiently computed by employing its dual expression: for $\beta \ll 1$ a few terms of (4.6.2) are indeed enough to saturate the entire sum. This example shows that, sometimes, simple algebraic transformations permit us to establish important duality relations of the partition functions. In this section we focus our attention on these aspects of the two-dimensional statistical models.

Curl and divergence. In two dimensions, the duality relation is strictly related to the curl and the divergence of a vector field. In fact, a two-dimensional vector field $\vec{v}$ with vanishing line integral along a close loop $\mathcal{C}$

$$\oint_{\mathcal{C}} d\vec{s} \cdot \vec{v} \,=\, 0, \tag{4.6.4}$$

satisfies the equation

$$\nabla \wedge \vec{v} = 0. \tag{4.6.5}$$

In this case, $\vec{v}$ can be expressed as the gradient of a scalar function Φ, i.e. $\vec{v} = \vec{\nabla}\,\Phi$. Going to the components, we have

$$\vec{v} = (v_1, v_2) = \left(\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}\right), \quad \text{if} \quad \vec{\nabla} \wedge \vec{v} = 0. \tag{4.6.6}$$

Vice versa, a vector field $\vec{v}$ with vanishing flux across a close surface $\mathcal{S}$

$$\oint_{\mathcal{S}} d\vec{\Sigma} \cdot \vec{v} = 0, \tag{4.6.7}$$

satisfies the equation

$$\vec{\nabla} \cdot \vec{v} = 0, \tag{4.6.8}$$

and it can always be expressed as $\vec{v} = \vec{\nabla} \wedge \vec{\Psi}$, where in two dimensions $\vec{\Psi} = (\psi, \psi)$ is a vector function of equal components. Explicitly

$$\vec{v} = (v_1, v_2) = \left(\frac{\partial \psi}{\partial y}, -\frac{\partial \psi}{\partial x}\right), \quad \text{if} \quad \vec{\nabla} \cdot \vec{v} = 0. \tag{4.6.9}$$

The comparision between eqn (4.6.6) and eqn (4.6.9) shows that we can swap between them by exchanging $x \longleftrightarrow -y$.

Curl and divergence on a lattice. The above equations have a counterpart for variables that live on a lattice. Consider a square lattice and its dual, where the sites of the first lattice are identified by the coordinates (i, j) while those of the dual by the coordinates $\left(i + \frac{1}{2}, j + \frac{1}{2}\right)$. Suppose that there are some statistical variables defined along the links of the original lattice: denote by $\rho_{i+\frac{1}{2},j}$ the variable defined along the horizontal segment that links the site (i, j) to the site $(i + 1, j)$ and by $\rho_{i,j+\frac{1}{2}}$ the one defined along the vertical segment that links (i, j) to $(i, j + 1)$. If the circulation along the perimeter $\mathcal{S}$ of the elementary cell of the lattice is zero (see Fig. 4.10), we have

$$\rho_{i+\frac{1}{2},j} + \rho_{i+1,j+\frac{1}{2}} - \rho_{i+\frac{1}{2},j+1} - \rho_{i,j+\frac{1}{2}} = 0.$$

This is the discrete version of the curl-free equation on the sites of the dual lattice. It can be identically satisfied in terms of a variable $\phi_{i,j}$ defined on the sites of the original lattice, by imposing

$$\begin{aligned} \rho_{i+\frac{1}{2},j} &= \phi_{i+1,j} - \phi_{i,j}, \\ \rho_{i,j+\frac{1}{2}} &= \phi_{i,j+1} - \phi_{i,j}. \end{aligned}$$

Vice versa, the discrete version on a lattice of the divergence-free condition (4.6.8) is given by

$$\rho_{i+\frac{1}{2},j} - \rho_{i-\frac{1}{2},j} + \rho_{i,j+\frac{1}{2}} - \rho_{i,j-\frac{1}{2}} = 0.$$

This can be satisfied by expressing the variables ρ in terms of a discrete curl of a variable $\psi_{i+\frac{1}{2},j+\frac{1}{2}}$ defined on the dual lattice

$$\begin{aligned} \rho_{i+\frac{1}{2},j} &= \psi_{i+\frac{1}{2},j+\frac{1}{2}} - \psi_{i+\frac{1}{2},j-\frac{1}{2}}, \\ \rho_{i,j+\frac{1}{2}} &= -\psi_{i+\frac{1}{2},j+\frac{1}{2}} + \psi_{i-\frac{1}{2},j+\frac{1}{2}}. \end{aligned} \tag{4.6.10}$$

After these general considerations, let's see two examples.

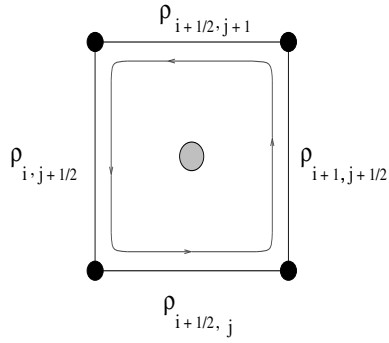


Fig. 4.10 *Circulation along the links of the original lattice. The site at the center belongs to the dual lattice.*

4.6.1 Self-duality of the p-state Model

Consider a statistical model with scalar variables $\phi_{i,j}$ defined on the $N \times N$ sites of a square lattice, with periodic boundary conditions. Assume that these variables take discrete values on the interval (p an integer)

$$1 \leq \phi_{ij} \leq p,$$

and their hamiltonian is a function of the differences of the next neighbor values

$$\mathcal{H} = -\sum_{i,j}^{N} \left[K_1 \left(\phi_{i+1,j} - \phi_{i,j}\right) + K_2 \left(\phi_{i,j+1} - \phi_{i,j}\right)\right]. \tag{4.6.11}$$

Introducing the notation

$$\begin{aligned} \rho_{i+\frac{1}{2},j} &= \phi_{i+1,j} - \phi_{i,j}; \\ \rho_{i,j+\frac{1}{2}} &= \phi_{i,j+1} - \phi_{i,j}, \end{aligned}$$

together with $K_1 = \beta J_1$, $K_2 = \beta J_2$ for the coupling constants along the horizontal and vertical directions, respectively, the partition function is given by

$$Z[K] = \mathrm{Tr}_\phi \exp\left[K_1 \rho_{i+\frac{1}{2},j} + K_2 \rho_{i,j+\frac{1}{2}}\right]. \tag{4.6.12}$$

In this expression we adopt the notation[4]

$$\mathrm{Tr}_\phi \equiv \prod_{i=1}^{N} \prod_{j=1}^{N} \frac{1}{\sqrt{p}} \sum_{\phi_{i,j}=1}^{p}$$

and we have taken into account the periodic boundary conditions

$$\begin{aligned} \phi_{i+N,j} &= \phi_{i,j}, \\ \phi_{i,j+N} &= \phi_{i,j}. \end{aligned}$$

[4]We have inserted the factor $1/\sqrt{p}$ in order to make the final expressions of the partition function symmetric.

There are then N^2 variables $\phi_{i,j}$ over which it is necessary to sum in order to obtain $Z[K]$. However, since the hamiltonian depends on them only through the variables ρ, it would be more convenient to use directly these quantities. Since their number is equal to $2N^2$, we need to implement the N^2 conditions of vanishing circulation

$$R_{i+\frac{1}{2},j+\frac{1}{2}} \equiv \rho_{i+\frac{1}{2},j} + \rho_{i+1,j+\frac{1}{2}} - \rho_{i+\frac{1}{2},j+1} - \rho_{i,j+\frac{1}{2}} = 0, \quad (\mathrm{mod}\ p). \tag{4.6.13}$$

This can be done by introducing N^2 variables $\psi_{i+\frac{1}{2},j+\frac{1}{2}}$, that take p integer values, conjugated to each of $R_{i+\frac{1}{2},j+\frac{1}{2}}$ and defined on the sites of the dual lattice. We can insert in the partition function the N^2 expressions

$$\Delta_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{p} \sum_{\psi_{i+\frac{1}{2},j+\frac{1}{2}}=1}^{p} \exp\left[-\frac{2\pi i}{p}\, \psi_{i+\frac{1}{2},j+\frac{1}{2}}\, R_{i+\frac{1}{2},j+\frac{1}{2}}\right].$$

They are equal to 1 if the condition (4.6.13) is satisfied and 0 otherwise. Hence, the partition function can be equivalently written as

$$Z[K] = \mathrm{Tr}_\rho\, \Delta_{i+\frac{1}{2},j+\frac{1}{2}}\, \exp\left[K_1 \rho_{i+\frac{1}{2},j} + K_2 \rho_{i,j+\frac{1}{2}}\right],$$

namely

$$Z[K] = \mathrm{Tr}_\rho\, \mathrm{Tr}_\psi\, \exp\left[K_1 \rho_{i+\frac{1}{2},j} + K_2 \rho_{i,j+\frac{1}{2}} - \frac{2\pi i}{p}\, \psi_{i+\frac{1}{2},j+\frac{1}{2}}\, R_{i+\frac{1}{2},j+\frac{1}{2}}\right]. \tag{4.6.14}$$

where

$$\mathrm{Tr}_\psi \equiv \frac{1}{p} \sum_{\psi_{i+\frac{1}{2},j+\frac{1}{2}}=1}^{p}.$$

Notice that the sum on the ρ's can be explicitly performed. Each variable ρ appears in three terms: for instance, considering $\rho_{i+\frac{1}{2},j}$, its contribution to the partition function is equal to

$$G = \frac{1}{p} \sum_{\rho_{i+\frac{1}{2},j}=1}^{p} \exp\left[\rho_{i+\frac{1}{2},j}\left(K_1 - \frac{2\pi i}{p}(\psi_{i+\frac{1}{2},j+\frac{1}{2}} - \psi_{i+\frac{1}{2},j-\frac{1}{2}})\right)\right]. \tag{4.6.15}$$

If we now define the dual coupling constant in terms of the Fourier transform of the original coupling constant

$$e^{\tilde{K}\sigma} = \frac{1}{\sqrt{p}} \sum_{b=1}^{p} e^{Kb} \exp\left(\frac{-2\pi i \sigma b}{p}\right), \tag{4.6.16}$$

eqn (4.6.15) can be expressed as

$$G = \frac{1}{\sqrt{p}} \exp\left[\tilde{K}_2\left(\psi_{i+\frac{1}{2},j+\frac{1}{2}} - \psi_{i+\frac{1}{2},j-\frac{1}{2}}\right)\right]. \tag{4.6.17}$$

By summing on all variables ρ in (4.6.14), the partition function can be equivalently expressed in terms of the variables ψ of the dual lattice and it fulfills the important self-duality relation

$$Z[K] \;=\; Z[\tilde{K}] \tag{4.6.18}$$

with

$$\begin{aligned} Z[\tilde{K}] \;=\; \mathrm{Tr}_{\psi} \prod_{i=1}^{N} \prod_{j=1}^{N} \exp\left[\tilde{K}_2 \left(\psi_{i+\frac{1}{2},j+\frac{1}{2}} - \psi_{i+\frac{1}{2},j-\frac{1}{2}}\right)\right] \\ \times \exp\left[\tilde{K}_1 \left(\psi_{i+\frac{1}{2},j+\frac{1}{2}} - \psi_{i-\frac{1}{2},j+\frac{1}{2}}\right)\right]. \end{aligned} \tag{4.6.19}$$

In conclusion, the dual coupling constants are defined by the Fourier transform of the original couplings, eqn (4.6.16). More precisely, the coupling $\tilde{K}_2$ relative to the vertical links is determined by the original coupling K_1 of the horizontal links, while the coupling $\tilde{K}_1$ of the horizontal links depends on the coupling K_2 of the vertical links of the original lattice. This procedure can be clearly implemented also when the couplings are not constant but change along the sites of the lattice.

4.6.2 Duality Relation between XY Model and SOS Model

The application of the duality transformation does not necessarily lead to the same model. Even though in these cases we cannot predict the critical temperature of the model, the duality relation that links two different models can nevertheless be useful for studying the excitations in their high- and low-temperature phases respectively. For instance, this is the case of the XY model that is related by duality to the SOS (Solid on Solid) model. The statistical variables of the XY model are the angles θ_i (with values between $-\pi$ and π) defined on each site of the lattice. The hamiltonian is

$$\mathcal{H} \;=\; -\sum_{\langle r,r'\rangle} \hat{f}(\theta_r - \theta_{r'}), \tag{4.6.20}$$

where $\hat{f}(\theta)$ is a periodic function, with period 2π

$$\hat{f}(\theta + 2\pi) \;=\; \hat{f}(\theta).$$

The usual choice is[5]

$$\hat{f}(\theta_r - \theta_{r'}) \;=\; J\,[1 - \cos(\theta_r - \theta_{r'})]\,. \tag{4.6.21}$$

The partition function is given by

$$Z[K] \;=\; \prod_{r} \int_{-\pi}^{\pi} \frac{d\theta_r}{2\pi} \prod_{\langle r,r'\rangle} e^{f(\theta_r - \theta_{r'})}, \tag{4.6.22}$$

[5]For simplicity in the sequel we only consider the homogeneous case.

with $f = \beta \hat{f}$. Since every term of this sum is a periodic function of the angles, it can be expanded in the Fourier series

$$e^{f(\theta-\theta')} = \sum_{n=-\infty}^{\infty} e^{\tilde{f}(n)} \exp(2\pi i\, n(\theta - \theta')). \tag{4.6.23}$$

For the inverse formula we have

$$e^{\tilde{f}(n)} = \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} e^{f(\theta)} \exp(-2\pi i\, n\theta).$$

Using eqn (4.6.23), the partition function becomes

$$Z[K] = \prod_r \int_{-\pi}^{\pi} \frac{d\theta_r}{2\pi} \prod_{\langle r,r'\rangle} \sum_{\{n_{r,r'}\}} e^{\tilde{f}(n_{r,r'})} \exp(2\pi i\, n_{r,r'}(\theta_r - \theta_{r'})), \tag{4.6.24}$$

where $n_{r,r'}$ are variables with integer values defined on the links between the next neighbor sites r and r'. In two dimensions, every angle θ_r enters the expression for four different terms, i.e. those relative to the segments that link the site r to its four next neighbor sites. By adopting the previous notation for the coordinates of the sites and for the variables defined along the links, the term in which $\theta_{i,j}$ is present is given by (see Fig. 4.11)

$$\exp\left[\theta_{i,j}(n_{i+\frac{1}{2},j} + n_{i,j+\frac{1}{2}} - n_{i-\frac{1}{2},j} - n_{i,j-\frac{1}{2}}\right].$$

Thanks to the identity

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} d\alpha\, e^{i\alpha x} = \delta_{x,0},$$

by integrating over θ_r in (4.6.24), we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta_r e^{i\theta_r \sum_{r'} n_{r,r'}} = \delta_{\sum_{r'} n_{r,r'},0},$$

i.e. the variable $n_{r,r'}$ defined along the links has zero divergence. Referring to the general discussion of the previous section, the variables $n_{r,r'}$ can then be expressed in

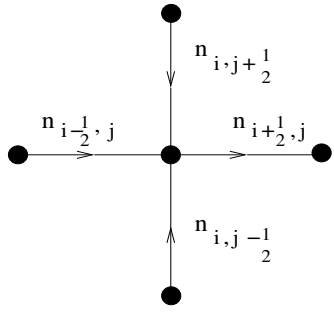


Fig. 4.11 *Condition of the vanishing divergence of the variables* $n_{r,r'}$.

terms of the differences of the integer value variables m_s defined on the sites of the dual lattice. In such a way, the original definition (4.6.22) of the partition function becomes a sum over all possible integer values of the variables m_s, defined on the sites s of the dual lattice

$$Z[K] = \sum_{\{m_s\}} \prod_{\langle s,s' \rangle} e^{\tilde{f}(m_s - m_{s'})}. \tag{4.6.25}$$

Hence, the dual model corresponding to the XY model is the SOS model, so called because the integer variables m_s can be regarded as the heights (either positive or negative) of a surface of a solid.

Appendix 4A. Numerical Series

In this appendix we will briefly discuss a numerical method for extracting useful information on the critical behavior of the thermodynamical quantities by using the first terms of their perturbative series. Let us consider a thermodynamical quantity, the partition function for instance, and suppose that such a quantity is expressed by a series expansion in the parameter x:

$$f(x) = \sum_{n=0} a_n x^n. \tag{4.A.1}$$

The problem consists of obtaining the parameters x_c and γ relative to its behavior close to the critical point x_c

$$f(x) \sim b\,(x_c - x)^{-\gamma} = b\,x_c^{-\gamma} \left(1 - \frac{x}{x_c}\right)^{-\gamma}, \tag{4.A.2}$$

if the only information available is the first k terms of the series (4.A.1).

The solution of this problem is the following. First of all, the estimate of the critical point x_c can be done by means of the convergence radius of the series (4.A.1) by assuming that there is no other singularity (also complex) closer to the origin. Expanding the right-hand side of (4.A.2) in a power series we have

$$f(x) \sim b\,x_c^{-\gamma} \left[1 + \gamma \left(\frac{x}{x_c}\right) + \frac{\gamma(\gamma+1)}{2!} \left(\frac{x}{x_c}\right)^2 + \cdots \frac{\gamma(\gamma+1)\cdots(\gamma+k-1)}{k!} \left(\frac{x}{x_c}\right)^k + \cdots\right]. \tag{4.A.3}$$

Considering the ratio of the next two coefficients of this series and comparing with the corresponding ratio of the series (4.A.1) we have

$$R_n = \frac{a_n}{a_{n-1}} = \frac{1}{x_c}\left[1 + \left(\frac{\gamma - 1}{n}\right)\right]. \tag{4.A.4}$$

Hence, with the hypothesis made above on the singularities of the function $f(x)$, a plot of the ratios R_n versus the variable $1/n$ should show a linear behavior, whose

slope provides an estimate of the quantity $x_c^{-1}(\gamma - 1)$, whereas its value at the origin gives an estimate of x_c^{-1}.

As an example of this method, let us consider the susceptibility of the Ising model on a two-dimensional triangular lattice. The high-temperature series expansion of this quantity is known up to the twelfth term and it is given by ($v = \tanh \beta J$)

$$\begin{aligned}\chi(T) = 1 &+ 6v + 30v^2 + 138v^3 + 606v^4 + 258v^5 \\ &+ 10818v^6 + 44574v^7 + 181542v^8 + 732678v^9 \\ &+ 2.935.218v^{10} + 11.687.202v^{11} + 46.296.210v^{12} + \cdots\end{aligned} \tag{4.A.5}$$

Employing the ratios R_n obtained by these coefficients, we arrive at the following estimates of the critical temperature and the coefficient γ:

$$v_c^{-1} \simeq 3.733 \pm 0.003; \quad \gamma \simeq 1.749 \pm 0.003$$

which are remarkably close to their exact values:

$$v_c^{-1} = 2 + \sqrt{3} = 3.73205...; \quad \gamma = \frac{7}{4} = 1.75.$$

Appendix 4B. Poisson Resummation Formula

Consider the series

$$f(x) = \sum_{m=-\infty}^{\infty} G(x + mT), \tag{4.B.1}$$

where $G(x)$ is a function that admits a Fourier transform. Since $f(x)$ is a periodic function

$$f(x) = f(x + T),$$

it can be expressed in a Fourier series

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \exp\left[\frac{2\pi i n x}{T}\right],$$

with the coefficients given by

$$c_n = \frac{1}{T}\int_0^T dy\, f(y) \exp\left[-\frac{2\pi i n y}{T}\right].$$

Substituting the expression of $f(x)$, we have

$$c_n = \frac{1}{T}\sum_{m=-\infty}^{\infty}\int_0^T dy\, G(y + mT) \exp\left[-\frac{2\pi i n y}{T}\right].$$

By making the change of variable $y + mT \to z$, we obtain

$$\begin{aligned} c_n &= \frac{1}{T} \sum_{m=-\infty}^{\infty} \int_{mT}^{(m+1)T} dz\, G(z)\, \exp\left[-\frac{2\pi i n z}{T}\right] \\ &= \frac{1}{T} \int_{-\infty}^{\infty} dz\, G(z)\, \exp\left[-\frac{2\pi i n z}{T}\right] = \frac{1}{T}\, \hat{G}\left(\frac{2\pi n}{T}\right), \end{aligned}$$

where $\hat{G}(p)$ is the Fourier transform of the function $G(x)$

$$\hat{G}(p) = \int_{-\infty}^{\infty} dz\, G(z)\, e^{-ipz}.$$

In such a way, the original series (4.B.1) can be expressed as

$$\sum_{m=-\infty}^{\infty} G(x + mT) = \frac{1}{T} \sum_{m=-\infty}^{\infty} \hat{G}\left(\frac{2\pi m}{T}\right) \exp\left[\frac{2\pi i m}{T}\right]. \tag{4.B.2}$$

This equation is known as the *Poisson sum formula*. It is also equivalent to the following identity for the $\delta(x)$ function

$$\sum_{m=-\infty}^{\infty} \delta(x - mT) = \frac{1}{T} \sum_{m=-\infty}^{\infty} \exp\left[\frac{2\pi i m x}{T}\right]. \tag{4.B.3}$$

References and Further Reading

The famous articles by Peierls, Kramers and Wannier are:

R.E. Peierls, *On Ising's model of ferromagnetism*, Proc. Camb. Philos. Soc. 32 (1936), 477.

H.A. Kramers, G.H. Wannier, *Statistics of the two-dimensional ferromagnet. Part I*, Phys. Rev. 60 (1941), 252.

H.A. Kramers, G.H. Wannier, *Statistics of the two-dimensional ferromagnet. Part II*, Phys. Rev. 60 (1941), 263.

Two important references on the duality properties in statistical mechanics and field theory are given by

L. Kadanoff, *Lattice Coulomb gas representations of two-dimensional problems*, J. Phys. A 11 (1978), 1399.

R. Savit *Duality in field theory and statistical mechanics*, Rev. Mod. Phys. 52 (1980), 453.

It is worth mentioning the studies on the duality properties of the three-dimensional Ising model. See:

F. Wegner, *Duality in generalized Ising models and phase transitions without local order parameters*, J. Math. Phys. 12 (1971), 2259.

R. Balian, J.M. Drouffe and C. Itzykson, *Gauge fields on a lattice. II Gauge-invariant Ising model*, Phys. Rev. 11 (1975), 2098.

The duality transformation plays an important role also in the theory of the fundamental interactions. For this aspect it is useful to consult the article:

N. Seiberg, E. Witten, *Electric–magnetic duality, monopole condensation and confinement in $N = 2$ supersymmetric Yang–Mills theory*, Nucl. Phys. B 426 (1994), 19.

Problems

1. Three-dimensional lattices

Generalize Peierls's argument to the Ising model on three dimensional lattices and prove that the model admits a phase transition.

2. Low-temperature series in the presence of a magnetic field

Consider the two-dimensional Ising model on a square lattice with equal coupling along the horizontal and vertical links and in the presence of an external magnetic field B. Generalize the discussion on the series expansion of the free energy in the low-temperature phase and show that Z_N can be written as

$$Z_N \;=\; \exp[2NK + N\beta B] \sum_{r,s=1}^{\infty} n(r,s)\, \exp[-2Kr]\, \exp[-s\beta B]$$

where $K = \beta J$ and $n(r,s)$ is the number of closed graphs made of r links on the dual lattice, having in their internal region s points of the original lattice.

3. Free energy

Consider the high-temperature series expansion of a homegeneous Ising model on a two-dimensional square lattice

$$\begin{aligned} Z_N =& (2\cosh \beta J)^{2N} \\ & \times \left[1 + Nv^4 + 2Nv^6 + 2Nv^8 + \frac{1}{2}N(N-9)v^8 + \cdots\right] \end{aligned}$$

($v = \tanh \beta J$). In the thermodynamic limit, one should have $Z_N \simeq (Z_1)^N = e^{-N\beta f}$, where f is the free energy per unit site. Using the formula above to find the high series expansion of Z_1

$$Z_1 \;=\; 2(\cosh \beta J)^2\, (1 + v^4 + 2v^6 - 2v^8 + \cdots).$$

4. Poisson sum rule

a Generalize the Poisson sum rule to the d-dimensional case.

b Using the Poisson sum rule show

$$\sum_{n=0}^{\infty} \frac{1}{x^2+n^2} = \frac{1}{2x^2} - \frac{\pi}{2x} + \frac{\pi}{x}\,\frac{1}{1-e^{-2Pix}}.$$

5. Self-duality

Consider the function

$$Z(K) = K^{1/4} \sum_{n=-\infty}^{\infty} e^{-\pi K n^2}$$

that satisfies

$$Z(K) = Z\left(\frac{1}{K}\right).$$

a Show that in this case the duality relation does not imply a phase transition at $K = 1$.

b How many terms are necessary in the original expression to compute $Z(K)$ with a precision 10^{-4} for $K = 0.01$? How many terms are needed to reach the same precision by using its dual expression?

6. Critical temperature of the three-state model

Using the self-duality of the three–state model to determine its critical temperature.

7. Quadratic model

Let

$$\beta\mathcal{H} = -\frac{J}{2} \sum_{\langle r,r'\rangle} (\Phi_r - \Phi_{r'})^2 + \frac{1}{4} \ln J$$

be the hamiltonian of a two-dimensional system where its variables assume all real values. Show that the model is self-dual under the transformation $J \longleftrightarrow 1/J$.

5
Combinatorial Solutions of the Ising Model

To make a correct conjecture on an event, it seems that it is necessary to calculate the number of all the possible cases exactly and to determine their combinatorics.

Jacob Bernoulli, *Ars Conjectandi*

There are many methods to solve exactly the two-dimensional Ising model at zero magnetic field. Some of these methods have proved to be quite general and they have been employed in the solution of other important models of statistical mechanics. This is the case, for instance, for the method of commuting transfer matrices, based on the solution of the Yang–Baxter equations, which will be discussed in the next chapter. On the contrary, other methods prove to be applicable only to the Ising model, such as the two combinatorial approaches that we are going to discuss in this chapter. Both methods are quite ingenious and original and this alone justifies their detailed analysis. The first method, which starts from the high-temperature series expansion of the Ising model, finally reduces the free energy computation to a problem of a random walk on a lattice. The second method, which also starts from the high-temperature series, transforms the problem of computing the free energy of the Ising model into a counting problem of dimer configurations on a lattice.

5.1 Combinatorial Approach

5.1.1 Partition Function

The combinatorial solution of the Ising model, originally proposed by M. Kac and J.C. Ward, has its starting point in the high-temperature series expansion of the partition function, discussed in Section 4.2.1 of the previous chapter. The elegant solution presented here is due to N.V. Vdovichenko. In the following we consider, for simplicity, only the homogeneous case in which there is only one coupling constant, so that in the partition function only the parameter $v = \tanh \beta J$ enters. The partition function on a square lattice is given by

$$Z_N \;=\; 2^N (1 - v^2)^{-N}\, \Phi(v). \tag{5.1.1}$$

with

$$\Phi(v) \;=\; \sum_r g_r v^r, \tag{5.1.2}$$

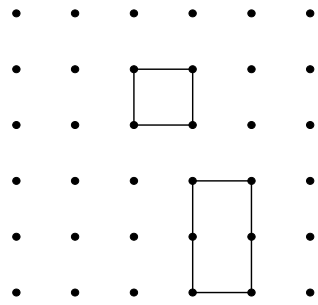


Fig. 5.1 *Graph of order* v^{10}.

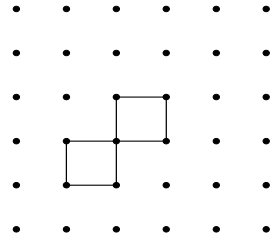


Fig. 5.2 *Self-intersecting graph.*

where g_r is the number of closed graphs, not necessarily connected, given by an even number r of links. The graph shown in Fig. 5.1, for instance, is one of the terms of order v^{10} present in the summation (5.1.2).

There are three steps in Vdovichenko's method of solution: (a) the first step consists of expressing the sum over the polygons as a sum over the closed loops without intersections; (b) the second step in transforming the sum over the closed loops without intersections into a sum over all possible closed loops; (c) in the last step, the problem is reduced to a random walk on a lattice that can be easily solved.

Let's discuss the implementation of the first step, i.e. how to organize the sum over the polygons in terms of their connected parts. Let's observe that each graph consists of one or more connected parts. For non-self-intersecting graphs this statement is obvious: the graph of Fig. 5.1, for instance, consists of two disconnected parts. But for self-intersecting graphs the statement can be ambiguous and there could be different connected parts according to the different decompositions. In order to clarify this issue, consider the graph in Fig. 5.2. This can be decomposed in three different ways, as shown in Fig. 5.3: it can be decomposed into one or two connected parts without intersections or into one connected part but with intersection. It is easy to show that this rule is quite general, namely there are always three possible decompositions for all the self-intersections of a graph.

The sum over the polygons given in eqn (5.1.2) can be organized into a sum over the connected parts of the graphs but one has to be careful to count properly the different terms, in particular to not count the same configuration more than once. This problem can be solved by weighting each graph by a factor $(-1)^n$, where n is

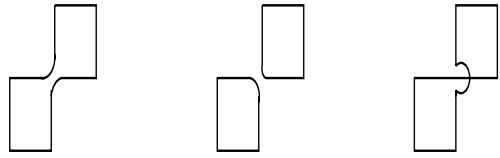


Fig. 5.3 *Three different decompositions in the connected parts for a self-intersecting graph.*

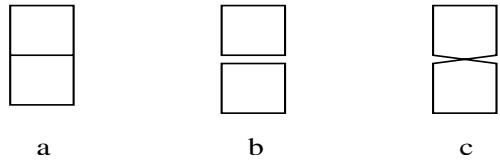


Fig. 5.4 *Graph with repeated bonds.*

the total number of self-intersections of a loop. In this way, all extra terms in the sum disappear. In the example of Fig. 5.3, the first two terms are weighted by $+1$, and the last term by -1, so that in the final expression there is correctly only one term.

Notice that, by adopting the prescription given above to perform the sum over the closed loops, one can include in the sum also the graphs with repeated bonds; the simplest of them is given in Fig. 5.4. These graphs are obviously absent in the original formulation of the high-temperature expansion of the model, since in some of their sites there is an odd number of links. However, with the new weight associated to the diagrams, it is easy to see that these terms are canceled in the sum. In fact, in the connected decomposition part of these graphs, each common link can be passed through in two different ways, one without intersection (as in Fig. 5.4b), the other with self-intersection, as shown in Fig. 5.4c. Hence, the connected parts of this graph have equal and opposite signs and therefore they cancel in the sum.

There is still a disadvantage in the procedure of assigning a weight to the graphs because it depends on a global property of the graph such as the number of its intersections. It would be more convenient to express the weight $(-1)^n$ in a local way. This is possible thanks to the familiar geometrical property that the total angle of rotation spanned by the tangent going around a closed plane loop is $2\pi(l+1)$ where l is an integer (positive or negative), with a parity that coincides with the number ν of the self-intersection of the loop. Hence, we can assign a phase factor $e^{i\alpha/2}$ to each point of the loop, where the angle of rotation α takes values $\alpha = 0, \pm\frac{\pi}{2}$ in correspondence with the angle of the change of direction to the next bond, so that the product of all these factors going around the loop gives $(-1)^{\nu+1}$. For a set of s loops we will have $(-1)^{n+s}$, with $n = \sum \nu$.

In summary, we can automatically take into account the number of self-intersections of a loop by weighting each node by $e^{i\alpha/2}$ and multiplying the graph (given by a set of s loops) by the factor $(-1)^s$, since this term will compensate the same factor present in the previous expression $(-1)^{n+s}$.

Let's now denote by f_r the sum over single loops of r links, each loop weighted according to the prescription above. The sum on all pairs of loops with total number

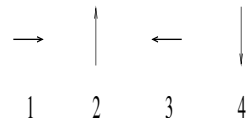


Fig. 5.5 *Possible directions of movement on a square lattice.*

r of links is then given by

$$\frac{1}{2!}\sum_{r_1+r_2=r} f_{r_1} f_{r_2},$$

where the factor 2! in the denominator takes into account that the permutation of the two indices gives rise to the same pair of loops. An analogous factor $n!$ is present in the denominator for the sum on n loops. Therefore, the function Φ can be written as

$$\Phi(v) \;=\; \sum_{s=0}(-1)^s \frac{1}{s!} \sum_{r_1,r_2,\cdots=1}^{\infty} v^{r_1+r_2+\cdots+r_s} f_{r_1}\ldots f_{r_s}. \tag{5.1.3}$$

Since in Φ there are terms corresponding to sets of loops with any possible total length[1] $r = r_1 + r_2 + \cdots$, in the sum (5.1.3) the indices $r_1, r_2, \ldots$ assume independently all values from 1 to ∞, so that

$$\sum_{r_1,r_2,\cdots=1}^{\infty} v^{r_1+r_2+\cdots+r_s} f_{r_1}\ldots f_{r_s} \;=\; \left(\sum_{r=1}^{\infty} v^r f_r\right)^s .$$

Hence Φ is expressed as

$$\Phi(v) \;=\; \exp\left[-\sum_{r=1}^{\infty} v^r f_r\right]. \tag{5.1.4}$$

With this expression we have completed the steps (a) and (b) of Vdovichenko's method. It remains then to evaluate explicitly the quantity f_r.

Since in a square lattice there are four different directions in which one can move, it is convenient to number them by the index $\mu = 1, 2, 3, 4$, as shown in Fig. 5.5. Let's introduce a new function $W_r(i, j, \mu)$: this is defined as the sum over all possible paths of length r that start from a given point of coordinates (i_0, j_0) along a direction μ_0 and arrive at a point of coordinate (i, j) along the direction μ. The paths entering the definition of $W_r(i, j, \mu)$ are weighted with the factors $e^{i\alpha/2}$ previously introduced. If we now choose (i_0, j_0) as the initial point , $W_r(i_0, j_0, \mu_0)$ becomes the sum over all loops leaving and returning to the same point.[2] We then have the identity

$$f_r \;=\; \frac{1}{2r}\sum_{i_0,j_0,\mu} W_r(i_0, j_0, \mu), \tag{5.1.5}$$

where the term $1/(2r)$ takes into account the fact that in the sum on the right-hand side each loop can be crossed in two opposite directions and can have any of its r

[1]The loops with a number of sites larger than the number N of the sites of the lattice do not contribute to the sum, since they necessarily contain repeated bonds.

[2]It is understood that these closed loops cannot pass through the same links in the opposite direction. This means that the last step of these walks cannot be along the opposite direction of μ_0.

nodes as a starting point. Thanks to its definition, the function $W_r(i,j,\mu)$ satisfies the recursive equations

$$\begin{aligned}
W_{r+1}(i,j,1) &= W_r(i-1,j,1) + e^{-i\frac{\pi}{4}} W_r(i,j-1,2) + 0 + e^{i\frac{\pi}{4}} W_r(i,j+1,4),\\
W_{r+1}(i,j,2) &= e^{i\frac{\pi}{4}} W_r(i-1,j,1) + W_r(i,j-1,2) + e^{-i\frac{\pi}{4}} W_r(i+1,j,3) + 0\\
W_{r+1}(i,j,3) &= 0 + e^{i\frac{\pi}{4}} W_r(i,j-1,2) + W_r(i+1,j,3) + e^{-i\frac{\pi}{4}} W_r(i,j+1,4),\\
W_{r+1}(i,j,4) &= e^{-i\frac{\pi}{4}} W_r(i-1,j,1) + 0 + e^{i\frac{\pi}{4}} W_r(i+1,j,3) + W_r(i,j+1,4).
\end{aligned} \tag{5.1.6}$$

Let us consider, for instance, the first of them. One can reach the point $i,j,1$ by taking the last $(r+1)$-th step from the left, from below or from above but not from the right. The coefficients present in the equation come from the phase factors relative to the change of directions. With the same argument one can derive the other equations in (5.1.6). Introducing the matrix Λ of the coefficients, the recursive equations can be written as

$$W_{r+1}(i,j,\mu) \;=\; \sum_{i',j',\mu'} \Lambda(ij\mu \mid i'j'\mu')\, W_r(i',j',\mu') \tag{5.1.7}$$

which admits a suggestive interpretation: this equation can be interpreted as a Markov process associated to a random walk on the lattice, with the transition probability between two next neighbor sites expressed by the relative matrix element of Λ. Since there are four possible directions for this motion, keeping fixed all other parameters, Λ is a 4×4 matrix in the indices μ' and μ, whose graphical interpretation is shown in Fig. 5.6.

In the light of the interpretation given above of the recursive equations, the *transition probability* relative to a path of total length r is expressed by the matrix Λ^r. Notice that the diagonal elements of this matrix express the probability to return to the initial point after traversing a loop of length r, i.e. they coincide with $W_r(i_0,j_0,\mu_0)$. Therefore we have

$$\mathrm{T}r\,\Lambda^r \;=\; \sum_{i_0,j_0,\mu} W_r(i_0,j_0,\mu),$$

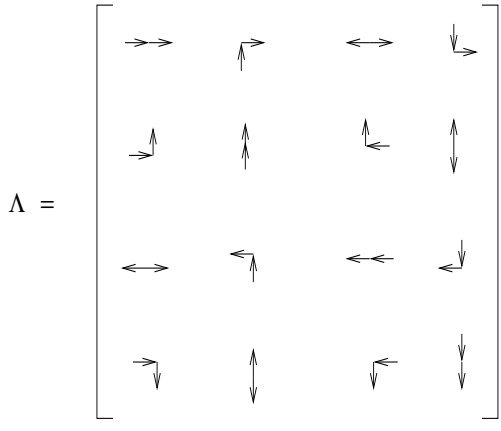


Fig. 5.6 *Matrix elements of* Λ.

and, comparing with eqn (5.1.5), we arrive at

$$f_r \;=\; \frac{1}{2r}\,\mathrm{T}r\,\Lambda^r \;=\; \frac{1}{2r}\sum_a \lambda_a^r, \tag{5.1.8}$$

where λ_a are the eigenvalues of the matrix Λ. Using this expression in (5.1.4) and interchanging the indices of the sum, we have

$$\begin{aligned}\Phi(v) &= \exp\left[-\frac{1}{2}\sum_i\sum_{r=1}^{\infty}\frac{1}{r}v^r\lambda_i^r\right]\\ &= \exp\left[\frac{1}{2}\sum_i \log(1-v\lambda_i)\right] \;=\; \prod_i \sqrt{1-v\lambda_i}.\end{aligned} \tag{5.1.9}$$

The last thing to do is to determine the eigenvalues of Λ. The diagonalization of this matrix with respect the coordinates k and l of the lattice can be easily done by using the Fourier transformation. In fact, defining

$$W_r(p,q,\mu) \;=\; \sum_{k,l=0}^{L} e^{-\frac{2\pi i}{L}(pk+ql)}\,W_r(k,l,\mu),$$

with $N = L^2$, and taking the Fourier transform of (5.1.6), we have

$$\begin{aligned}
W_{r+1}(p,q,1) &= \epsilon^{-p}\,W_r(p,q,1) + \epsilon^{-q}\,\alpha^{-1}\,W_r(p,q,2) + \epsilon^{q}\,\alpha\,W_r(p,q,4),\\
W_{r+1}(p,q,2) &= \epsilon^{-p}\,\alpha\,W_r(p,q,1) + \epsilon^{-q}\,W_r(p,q,2) + \epsilon^{p}\,\alpha^{-1}\,W_r(p,q,3),\\
W_{r+1}(p,q,3) &= \epsilon^{-q}\,\alpha\,W_r(p,q,2) + \epsilon^{p}\,W_r(p,q,3) + \epsilon^{q}\,\alpha^{-1}\,W_r(p,q,4),\\
W_{r+1}(p,q,4) &= \epsilon^{-p}\,\alpha^{-1}\,W_r(p,q,1) + \epsilon^{p}\,\alpha\,W_r(p,q,3) + \epsilon^{q}\,W_r(p,q,4).
\end{aligned} \tag{5.1.10}$$

(5.1.11)

(where $\epsilon = e^{2\pi i/L}$ and $\alpha = e^{i\pi/4}$). Since $W_r(p,q,\mu)$ appears with the same indices p and q both on the left- and right-hand sides of these equations, the Fourier transform of the matrix Λ is diagonal with respect to these indices and we have

$$\Lambda(p,q,\mu \mid p,q,\mu') = \begin{pmatrix} \epsilon^{-p} & \alpha^{-1}\epsilon^{-q} & 0 & \alpha\epsilon^{q}\\ \alpha\epsilon^{-p} & \epsilon^{-q} & \alpha^{-1}\epsilon^{p} & 0\\ 0 & \alpha\epsilon^{-q} & \epsilon^{p} & \alpha^{-1}\epsilon^{q}\\ \alpha^{-1}\epsilon^{-p} & 0 & \alpha\epsilon^{p} & \epsilon^{q}\end{pmatrix}. \tag{5.1.12}$$

An easy computation shows that

$$\begin{aligned}\prod_{i=1}^{4}(1-v\lambda_i) &= \mathrm{Det}(\mathbf{1}-v\Lambda)\\ &= (1+v^2)^2 - 2v(1-v^2)\left(\cos\frac{2\pi p}{L} + \cos\frac{2\pi q}{L}\right).\end{aligned} \tag{5.1.13}$$

Coming back to the original expression (5.1.1), we then have

$$Z_N \;=\; 2^N(1-v^2)^{-N}\prod_{p,q}^{L}\left[(1+v^2)-2v(1-v^2)\left(\cos\frac{2\pi p}{L}+\cos\frac{2\pi q}{L}\right)\right]^{1/2}, \quad (5.1.14)$$

and the free energy of the Ising model is expressed as

$$\begin{aligned}-\frac{F(T)}{kT} &= \log Z_N \\ &= N\log 2 - N\log(1-v^2) \\ &+ \frac{1}{2}\sum_{p,q=0}^{L}\log\left[(1+v^2)^2-2v(1-v^2)\left(\cos\frac{2\pi p}{L}+\cos\frac{2\pi q}{L}\right)\right].\end{aligned} \quad (5.1.15)$$

When $L\to\infty$, the sum becomes an integral

$$\begin{aligned}-\frac{F(T)}{kT} &= \log Z_N \\ &= N\log 2 - N\log(1-v^2) \\ &+ \frac{N}{2(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}\log\left[(1+v^2)^2-2v(1-v^2)\left(\cos\omega_1+\cos\omega_2\right)\right]\,d\omega_1 d\omega_2.\end{aligned} \quad (5.1.16)$$

This expression shows that $F(T)$ is an extensive quantity, since it is proportional to the total number N of the sites of the lattice. Besides the value $v=1$ (which corresponds to $T=0$), $F(T)$ has a singular point at a finite value of T when the argument of the logarithm inside the integral vanishes. As a function of ω_1 and ω_2, the argument of the logarithm has a minimum when $\cos\omega_1=\cos\omega_2=1$ and the corresponding value is

$$(1+v^2)^2-4v(1-v^2)=(v^2+2v-1)^2.$$

It is easy to see that this expression has a minimum, with a null value, only for the positive value

$$v=v_c=\sqrt{2}-1.$$

The corresponding critical temperature T_c, fixed by

$$\tanh\frac{J}{kT_c} \;=\; v_c\,, \quad kT_c \;=\; 2.26922\ldots J, \quad (5.1.17)$$

determines the phase transition point. The expansion of the function $F(T)$ in a power series in $t=k(T-T_c)/J$ around this critical point shows that it has both a singular and a regular part. The regular part is simply obtained by substituting $t=0$ in its expression. In order to determine the singular part, it is sufficient to expand the argument of the logarithm in a power series in t, in ω_1 and ω_2. In this way, the integral in (5.1.16) becomes

$$\int_0^{2\pi}\int_0^{2\pi}\log\left[a_1t^2+a_2(\omega_1^2+\omega_2^2)\right]\,d\omega_1 d\omega_2, \quad (5.1.18)$$

where a_1 and a_2 are two constants expressed by

$$a_1 \,=\, 32(3-2\sqrt{2})\,\left(\frac{J}{kT_c}\right)^4\,, \quad a_2 \,=\, 2(3-2\sqrt{2}).$$

Computing the integral, the behavior of the free energy in the vicinity of the phase transition is given by

$$F(T) \,\simeq\, A - \frac{B}{2}(T-T_c)^2\,\log\mid T-T_c\mid, \tag{5.1.19}$$

where A and B are two other constants, with $B > 0$. The specific heat, expressed by the second derivative of $F(T)$ with respect to T, has in this case a logarithmic singularity rather than a power law behavior

$$C \,\sim\, B\,\log\mid T-T_c\mid .$$

Correspondingly the critical exponent α of the two-dimensional Ising model is

$$\alpha = 0.$$

5.1.2 Correlation Function and Magnetization

In this section we briefly discuss the main steps that lead to the computation of the two-point correlation function of the Ising model in terms of the combinatorial method. Because of the mathematical intricacy of the formulas employed in this method, we will present only the final result. As shown in the following chapters, the computation of the correlation functions can be done in a more efficient way (in the continuum limit) by using the methods of quantum field theory.

In order to simplify the notation, in the following the coordinates of a generic site of the lattice will be denoted by one index alone, i.e. $i \equiv (i_1, i_2)$. Observe that knowledge of the two-point correlation function

$$G(\mid i-j\mid) \,=\, \langle\sigma_i\sigma_j\rangle, \tag{5.1.20}$$

can be used to see whether or not the system possesses a non-zero magnetization

$$M^2 \,=\, \lim_{|i-j|\to\infty}\langle\sigma_i\sigma_j\rangle. \tag{5.1.21}$$

Let's focus attention on the computation of $G(\mid i-j\mid)$, defined by

$$\langle\sigma_i\sigma_j\rangle \,=\, \frac{1}{Z}\sum_{\{\sigma\}}\sigma_i\sigma_j\exp\left[K\sum_{(k,l)}\sigma_k\sigma_l\right], \tag{5.1.22}$$

with $K = \beta\,J$. Using the familiar identity

$$\exp\left[x\sigma_k\sigma_l\right] \,=\, \cosh x\,(1+\sigma_k\sigma_l\tanh x)\,,$$

the numerator of (5.1.22) can be written as

$$\cosh K^{2N}\,\sum_{\{\sigma\}}\sigma_i\sigma_j\,\prod_{(k,l)}(1+v\sigma_k\sigma_l), \tag{5.1.23}$$

where $v = \tanh K$. Expanding the product one obtains 2^{2N} terms. Using the same graphical method discussed in Section 4.2.1, we draw a line along the segment (k,l) if

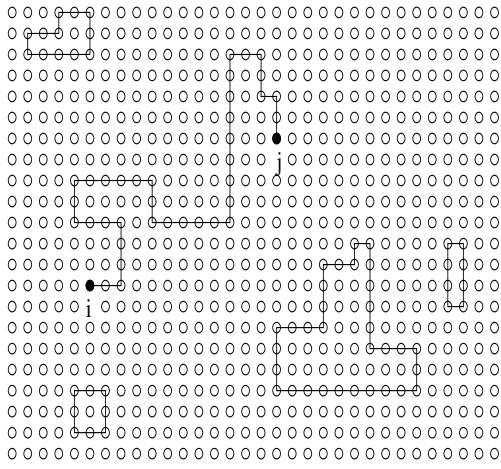


Fig. 5.7 *Graphs that enter the computation of the correlator* $\langle \sigma_i \sigma_j \rangle$.

this enters one of the terms of the expansion. This line has a weight equal to v. Once all the lines are drawn, we need to sum over the values of the spins. The difference with the computation of the partition function in this case consists of the presence of the spins σ_i and σ_j and one has a non-vanishing result only if there is at least one curve that starts from the site i and ends at the site j as shown in Fig. 5.7, where all other contributions are made of closed graphs.

Clearly the closed graphs that appear in the expansion of the numerator are the same as those that enter the expression of the partition function Z_N and therefore they simplify with the term Z_N in the denominator. Hence the correlation function can be expressed by the series

$$\langle \sigma_i \sigma_j \rangle \; = \; \sum_k h_k \, v^k, \tag{5.1.24}$$

where h_k is the number of graphs of length k (also self-intersecting) that connect the two end points. A simple example helps in clarifying the content of such a formula.

Consider the correlator of two nearest neighbor spins. The graphs relative to the lowest orders in v, i.e. $v^1, \ldots, v^5$, are shown in Fig. 5.8. Therefore, in this case, the first terms of the series are

$$v + 2v^3 + 6v^5 + \cdots$$

This example highlights a general and important aspect of the problem. Since the correlation function is nothing else but a conditional probability that the two spins σ_i and σ_j have the same value, for two neighbor spins such a probability is determined by two different effects: (1) the direct interaction between σ_i and σ_j, with a weight v; (2) the sum of all indirect interactions between the two spins, with a weight v^k for those indirect interactions that involving k spins.

Although it is generally difficult to compute the generic coefficient h_k of the series (5.1.24), for their geometrical origin it is, however, easy to determine the first non-vanishing coefficient. Denoting by $s_1 =\mid j_1 - i_1 \mid$ and $s_2 =\mid j_2 - i_2 \mid$ the horizontal and

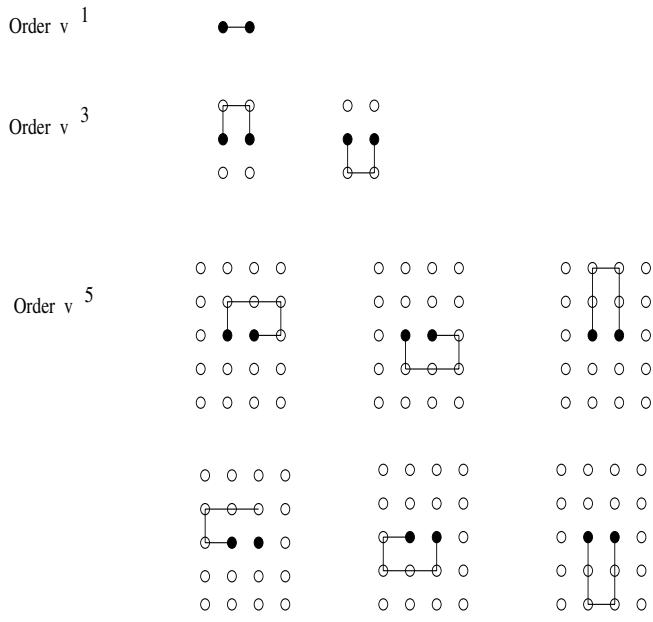


Fig. 5.8 *First-order terms of the correlation function of two nearest neighbor spins.*

vertical distances between the spins σ_i and σ_j, the number of paths of total length $s_1 + s_2$ (made of s_1 horizontal steps and s_2 vertical steps) is given by $(s_1 + s_2)!/s_1!\, s_2!$ and therefore

$$\langle \sigma_i \sigma_j \rangle \simeq \frac{(s_1 + s_2)!}{s_1!\, s_2!}\, v^{s_1+s_2} + \cdots \tag{5.1.25}$$

A further analysis of the series (5.1.24) (which is not discussed here) permits us to reach the following conclusions: for $T \neq T_c$, the two-point correlation function decays exponentially at large distances as

$$\langle \sigma_i \sigma_j \rangle \simeq M^2 + A \exp\left[-\frac{|\, i - j \,|}{\xi} \right], \tag{5.1.26}$$

where $A > 0$ is a constant. Near T_c, the correlation length ξ diverges as

$$\xi \simeq |\, T - T_c \,|^{-1}, \tag{5.1.27}$$

and the critical exponent ν of the two-dimensional Ising model is

$$\nu = 1.$$

The spontaneous magnetization M^2 can be extracted by the limit (5.1.21) and its exact expression, originally obtained by C.N. Yang, is

$$M^2 = \begin{cases} \left[1 - \left(\frac{1 - v^2}{2v} \right)^4 \right]^{1/4}, & T < T_c \\ 0, & T > T_c. \end{cases} \tag{5.1.28}$$

Hence the exact value of the critical exponent β is

$$\beta = \frac{1}{8}.$$

Finally, at $T = T_c$, the correlator decays algebraically as

$$\langle \sigma_i \sigma_j \rangle \simeq \frac{1}{| i - j |^{1/4}}, \tag{5.1.29}$$

and for the critical exponent η we have

$$\eta = \frac{1}{4}.$$

The remaining critical exponents δ and γ can be obtained by the scaling laws (1.1.26)

$$\delta = 15; \quad \gamma = \frac{7}{4}.$$

These are the exact expressions of all the critical exponents of the two-dimensional Ising model.

5.2 Dimer Method

From the geometrical nature of its high-temperature series expansion, the two-dimensional Ising model can be put into correspondence with the problem of counting the number of dimer configurations on a particular lattice. As we will see, this is a problem of a combinatorial nature that can be solved by evaluating the Pfaffian of an antisymmetrical matrix A.

The Pfaffian of an antisymmetric $2N \times 2N$ matrix

$$A = \begin{pmatrix} 0, & a_{1,2}, & \cdots, a_{1,2N} \\ -a_{1,2}, & 0, & \cdots, a_{2,2N} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ -a_{1,2N}, & -a_{2,2N}, & \cdots, 0 \end{pmatrix}$$

is defined as

$$\mathrm{Pf}\, A = \sum_{P}{}' \delta_P \, a_{p_1,p_2} \, a_{p_3,p_4} \cdots a_{p_{2N-1},p_{2N}}, \tag{5.2.1}$$

where $p_1, \ldots, p_{2N}$ is a permutation of the set of numbers $1, 2, \ldots, 2N$, δ_P is the parity of the permutation (± 1 if the permutation P is obtained by an even/odd number

of transpositions), and the sum $\sum_P'$ is over all permutations that satisfy the conditions

$$\begin{aligned} &p_{2m-1} < p_{2m}, \qquad 1 < m < N; \\ &p_{2m-1} < p_{2m+1}, \quad 1 < m < N-1. \end{aligned} \tag{5.2.2}$$

For instance, if $2N = 4$, one has

$$\mathrm{Pf}\, A \;=\; a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}.$$

Notice that, from the antisymmetry of the matrix A, its Pfaffian can also be expressed as

$$\mathrm{Pf}\, A \;=\; \frac{1}{N!\, 2^N} \sum_P \delta_P \, a_{p_1,p_2}\, a_{p_3,p_4} \, \cdots a_{p_{2N-1},p_{2N}}, \tag{5.2.3}$$

where the sum is over all possible permutations.

The computation of the Pfaffian of a matrix is simplified thanks to this important identity:

$$\mathrm{Pf}\, A \;=\; (\det A)^{1/2}. \tag{5.2.4}$$

Unlike Pfaffians, the determinants are in fact easier to compute, in particular by the property that the determinant of a product of matrices is equal to the product of the determinants.

A dimer is an object that can cover the links between nearest neighbor sites, with the condition that a given site cannot be occupied by more than one dimer. The combinatorial nature of the dimer problem consists of determining the number of possible dimers covering a lattice, such that all sites are occupied and none of them are occupied more than once. If the lattice is made of N sites, the number of dimers is $N/2$, hence N must be an even number. Before addressing the study of the Ising model in terms of the dimer formulation, it is convenient to study initially the dimer covering of a square lattice.

5.2.1 Dimers on a Square Lattice

The relationship between the dimer covering of a square lattice and the Pfaffian of a matrix can be highlighted by considering a 4×4 lattice. If we enumerate the sites as shown in Fig. 5.9, the dimer configuration can be identified by the pairs of numbers

$$(1,2)\,,\ (3,7)\,,\ (4,8)\,,\ (5,6)\,,\ (9,13)\,,\ (10,11)\,,\ (12,16)\,,\ (14,15),$$

or, more generally, by

$$(p_1,p_2)\,,\ (p_3,p_4)\,,\ (p_5,p_6)\,,\ \cdots (p_{2N-1},p_{2N}),$$

where $(p_1, p_2, \ldots, p_{2N})$ is a permutation of $(1, 2, \ldots, 2N)$ that satisfies the constraints (5.2.2) relative to the Pfaffian of a matrix. Assigning the matrix elements according

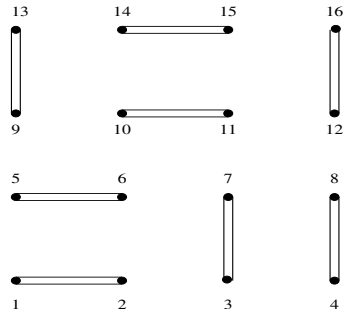


Fig. 5.9 *Dimer configuration of a* 4×4 *square lattice.*

to the rule

$$| A_{p,p'} | = \begin{cases} z_1, & \text{if } p > p', \text{ where } p \text{ and } p' \text{ are horizontal nearest neighbor sites} \\ z_2, & \text{if } p > p', \text{ where } p \text{ and } p' \text{ are vertical nearest neighbor sites} \\ 0, & \text{otherwise} \end{cases} \tag{5.2.5}$$

it is easy to see that there is a one-to-one correspondence between the dimer configurations and the terms present in the definition of the Pfaffian of the matrix A defined above. If we introduce the generating function of the dimers, defined by the formula

$$\Phi(z_1, z_2) \; = \; \sum_{n_1, n_2} g(n_1, n_2)\, z_1^{n_1}\, z_2^{n_2}, \tag{5.2.6}$$

where $g(n_1, n_2)$ is the number of dimers that cover completely the lattice, with n_1 placed horizontally and n_2 placed vertically ($n_1 + n_2 = N/2$), it seems natural to put

$$\Phi(z_1, z_2) \; = \; \mathrm{Pf}\, A. \tag{5.2.7}$$

There is, however, an obstacle: in fact, while $g(n_1, n_2)$, present in the generating function of the dimers, is a positive quantity, the definition of the Pfaffian of A involves also negative terms, i.e. those relative to the odd permutations of the indices. Hence, in order to make eqn (5.2.7) valid, in addition to the modulus (5.2.5) of the matrix elements $A_{p,p'}$, it is also necessary to introduce a phase factor that ensures the positivity of all terms present in Pf A. Thanks to a theorem due to P.W. Kasteleyn, this task can be accomplished for all planar lattices, i.e. for those lattices that do not have crossings of the links. For instance, in the case of a square lattice, an assignment that ensures the validity of eqn (5.2.7) is given by

$$A_{p,p'} \; = \; \begin{cases} z_1, & \text{for the horizontal links that are nearest neighbor} \\ (-1)^p\, z_2, & \text{for the vertical links that are nearest neighbor} \\ 0, & \text{otherwise.} \end{cases} \tag{5.2.8}$$

Notice that the definition of $A_{p,p'}$ given in (5.2.8) is equivalent to assigning a set of arrows along the links of the lattice, as shown in Fig. 5.10. In this way, the original lattice becomes an *oriented lattice*. In the presence of the arrows, the lattice acquires

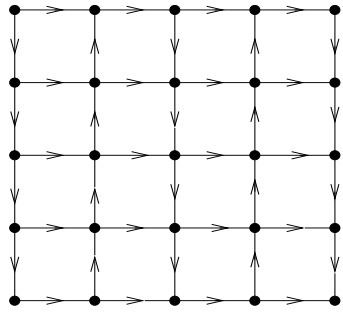


Fig. 5.10 *Assignment of the arrows in the dimer problem on a square lattice. The arrows in the up and the right directions correspond to the positive links, while the others correspond to the negative links.*

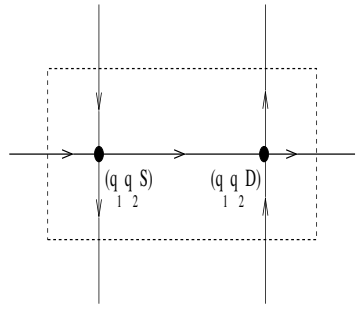


Fig. 5.11 *Elementary cell in the oriented square lattice.*

a periodicity along the horizontal axes under a translation of *two* lattice steps. It is therefore convenient to assume, as an elementary cell, not the one of unit length but the one drawn in Fig. 5.11, identified by its horizontal position q_1 and its vertical position q_2: these coordinates form the vector $\vec{q} = (q_1, q_2)$. Concerning its internal points, the one on the left is identified by (q_1, q_2, S) while the one on the right by (q_1, q_2, D).

Let us consider the matrix elements of the matrix $A_{\vec{q},\vec{p}} = A(\vec{q}, \vec{p})$. They are themselves 2×2 matrices, given by

$$A_{\vec{q},\vec{p}} = A(q_1, q_2; p_1, p_2) = \begin{pmatrix} a(q_1, q_2, S; p_1, p_2, S) & a(q_1, q_2, S; p_1, p_2, D) \\ a(q_1, q_2, D; p_1, p_2, S) & a(q_1, q_2, D; p_1, p_2, D) \end{pmatrix}. \quad (5.2.9)$$

The only non-vanishing matrix elements of $A_{\vec{q},\vec{p}}$ are given by

$$A(q_1, q_2; q_1, q_2) = \begin{pmatrix} 0 & z_1 \\ -z_1 & 0 \end{pmatrix} = \alpha(0, 0),$$
$$A(q_1, q_2; q_1 + 1, q_2) = \begin{pmatrix} 0 & 0 \\ z_1 & 0 \end{pmatrix} = \alpha(1, 0)$$

$$A(q_1,q_2;q_1-1,q_2) = \begin{pmatrix} 0 & -z_1 \\ 0 & 0 \end{pmatrix} = \alpha(-1,0), \tag{5.2.10}$$
$$A(q_1,q_2;q_1,q_2+1) = \begin{pmatrix} -z_2 & 0 \\ 0 & z_2 \end{pmatrix} = \alpha(0,1),$$
$$A(q_1,q_2;q_1,q_2-1) = \begin{pmatrix} z_1 & 0 \\ 0 & -z_2 \end{pmatrix} = \alpha(0,-1).$$

It is important to stress that the matrix A only depends on the difference of the indices

$$A(\vec{q};\vec{p}) = A(\vec{p}-\vec{q}).$$

Imposing periodic boundary conditions along the two directions

$$A(\vec{q}+\vec{N}) = A(\vec{q}),$$

where $\vec{N}=(N_1,N_2)$, A becomes a *cyclic matrix* that can be easily diagonalized with respect to the indices $\vec{q}$ and $\vec{p}$ by a Fourier transform.[3] The matrix elements of $A_{\vec{q},\vec{p}}$ are 2×2 matrices and, consequently, its diagonal form with respect to $\vec{q}$ and $\vec{p}$ consists of 2×2 matrices placed along its main diagonal. Denoting the latter matrices by $\lambda(\beta_1,\beta_2)$ we have

$$\lambda(\beta_1,\beta_2) = \sum_{\vec{q}} A(\vec{q})\, e^{i\vec{q}\cdot\vec{\beta}},$$

where each frequency β_i can have the N_i values $0, 2\pi/N_i, 4\pi/N_1, \ldots, 2\pi(N_i-1)/N_i$. Hence, the determinant of A is expressed by the product of the determinants of the 2×2 matrices λ

$$\frac{1}{N_1N_2}\log \operatorname{Det} A = \frac{1}{N_1N_2}\sum_{k_1=0}^{N_1-1}\sum_{k_2=0}^{N_2-1}\log \operatorname{Det}\lambda\left(\frac{2\pi k_1}{N_1},\frac{2\pi k_2}{N_2}\right). \tag{5.2.11}$$

In the thermodynamic limit $N_i\to\infty$ the sum can be converted to an integral

$$\lim_{N_i\to\infty}\frac{1}{N_1N_2}\log \operatorname{Det} A = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} d\beta_1\, d\beta_2 \log \operatorname{Det}\lambda(\beta_1,\beta_2), \tag{5.2.12}$$

where the matrix $\lambda(\beta_1,\beta_2)$ is explicitly given by

$$\begin{aligned}\lambda(\beta_1,\beta_2) &= \sum_{q_1,q_2}\alpha(q_1,q_2)\, e^{iq_1\beta_1+iq_2\beta_2} \\ &= \alpha(0,0)+\alpha(1,0)\, e^{i\beta_1}+\alpha(-1,0)\, e^{-i\beta_1} \\ &\quad +\alpha(0,1)\, e^{i\beta_2}+\alpha(0,-1)\, e^{-i\beta_2} \\ &= \begin{pmatrix} z_2\, e^{-i\beta_2}-z_2\, e^{i\beta_2} & z_1-z_1\, e^{-i\beta_1} \\ z_1\, e^{i\beta_1}-z_1 & z_2\, e^{i\beta_2}-z_2\, e^{-i\beta_2}\end{pmatrix}.\end{aligned} \tag{5.2.13}$$

[3]The procedure is similar to the one employed in the gaussian and spherical models, discussed in Chapter 3.

Computing the determinant of this matrix and using the important identity (5.2.4), we have

$$\lim_{N_i\to\infty}\frac{2}{N_1N_2}\,\log\,\mathrm{Pf}\,A=\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}d\beta_1\,d\beta_2\,\log\left[4\left(z_1^2\,\sin^2\frac{\beta_1}{2}+z_2^2\,\sin^2\frac{\beta_2}{2}\right)\right]$$
$$=\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi}d\beta_1\,d\beta_2\,\log\,2\left[\left(z_1^2+z_2^2\right)\right)$$
$$-z_1^2\,\cos\beta_1-z_2^2\,\cos\beta_2\Big]\,. \tag{5.2.14}$$

From the relation $\Phi(z_1,z_2)=\mathrm{Pf}\,A$ which links the generating function of the dimers to the Pfaffian of the matrix A, by plugging in (5.2.14) the values $z_1=z_2=1$, we obtain the total number of dimers covering a square lattice. The computation of the integral (proposed as Problem 4 at the end of the chapter), gives

$$\lim_{N_i\to\infty}\frac{2}{N_1N_2}\,\log\,\Phi(1,1)\,=\,\frac{4G}{\pi}, \tag{5.2.15}$$

where G is the *Catalan constant*, whose numerical value is

$$G\,=\,1-\frac{1}{3^2}+\frac{1}{5^2}-\frac{1}{7^2}+\cdots\,=\,0.9159655\ldots$$

In conclusion, the number of dimer coverings of a square lattice of N sites, with periodic boundary conditions on both directions, in the limit $N\to\infty$ is given by,[4]

$$D\,\simeq\,\exp\left[\frac{NG}{\pi}\right],\quad N\to\infty \tag{5.2.16}$$

By using the same method, employing the sum instead of the integral, one can obtain the dimer covering of finite lattices. For instance, for a 8×8 lattice, as that of a chessboard, the number of dimers is 32, and the number of their coverings of the lattice is $D=2^4(901)^2=12088816$, as was originally shown by Michael Fisher.

5.2.2 Dimer Formulation of the Ising Model

For the two-dimensional Ising model on a square lattice there is a one-to-one correspondence between the closed graphs of the high-temperature expansion and the dimer configurations relative to the lattice shown in Fig. 5.12, known as the Fisher lattice. Both lattices have, as a building block, an elementary cell with four external lines, see Fig. 5.13. We can associate to the eight possible configurations of the lines of the Ising model in the elementary cell eight possible dimer configurations on the Fisher lattice, as shown in Fig. 5.14 (by rotation, the configuration (c) gives rise to three other configurations whereas the configuration (d) only one). In such a way, to each closed graph of the high-temperature expansion of the Ising model on a square lattice there corresponds a dimer configuration on the Fisher lattice, and vice versa.

[4]Since the elementary cell of the oriented lattice is double the elementary cell of the ordinary lattice, we have $N=2N_1\,N_2$.

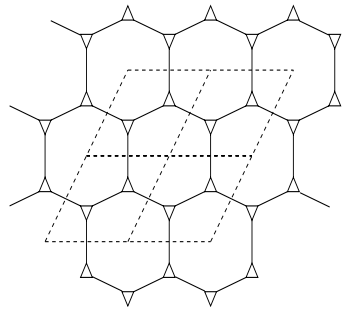


Fig. 5.12 *Fisher lattice.*

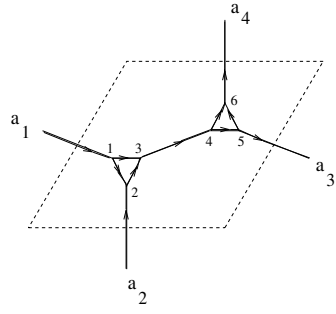


Fig. 5.13 *Elementary cells of the square and Fisher lattices.*

Let us consider the high-temperature expansion of the partition function of the model, given in eqn (4.2.4), here written as

$$(2\cosh K\ \cosh L)^{-N}\, Z_N \;=\; \sum_{r,s=0}^{\infty} n(r,s) v^r\, w^s, \tag{5.2.17}$$

where $n(r,s)$ is the number of closed graphs having r horizontal and s vertical links. Assigning weight v to the dimers along the segments a_1 and a_3, weight w to the dimers placed on the segments a_2 and a_4, and weight 1 to all internal dimers of the cell, it is easy to see that the right-hand side of eqn (5.2.17) may be interpreted as the generating function of the dimer configurations on the Fisher lattice. In turn, this function can be expressed in terms of the Pfaffian of an opportune antisymmetric matrix A. Hence we can follow the same steps for the computation of the dimer covering on the square lattice, with the only difference that, instead of the two internal points of the square lattice, this time the elementary cell has six internal points as shown in Fig. 5.13, with the corresponding orientation of the links. However, as in the previous case, the only matrices different from zero are $\alpha(0,0)$, $\alpha(\pm 1,0)$, and $\alpha(0,\pm 1)$, so that the matrix of

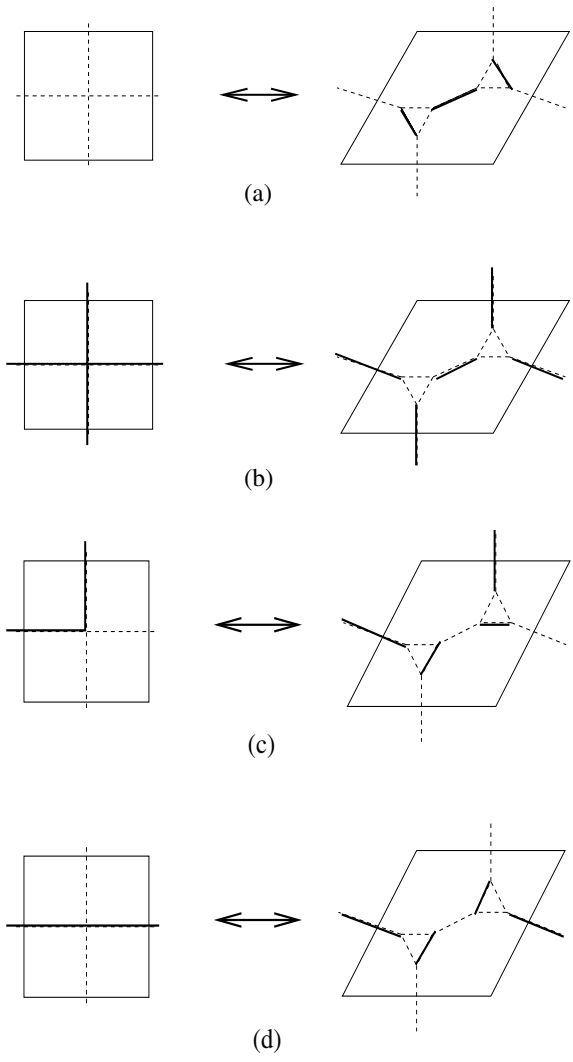


Fig. 5.14 *Correspondence between the lines of the Ising model on a square lattice and the dimers on the Fisher lattice.*

the eigenvalues is given in this case by

$$\lambda(\beta_1,\beta_2) = \begin{pmatrix} 0 & 1 & 1 & 0 & -v\,e^{-i\beta_1} & 0 \\ -1 & 0 & 1 & 0 & 0 & -w\,e^{-i\beta_2} \\ -1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \\ v\,e^{i\beta_1} & 0 & 0 & -1 & 0 & 1 \\ 0 & w\,e^{i\beta_2} & 0 & -1 & -1 & 0 \end{pmatrix}, \tag{5.2.18}$$

and therefore

$$\mathrm{Det}\,\lambda(\beta_1,\beta_2) \;=\; (1+v^2)(1+w^2) - 2v(1-w^2)\,\cos\beta_1 - 2w(1-v^2)\,\cos\beta_2. \quad (5.2.19)$$

Computing the Pfaffian of the matrix A, we obtain the free energy of the Ising model: in the thermodynamic limit and in the homogeneous case $v = w$, it is given by

$$-\frac{F(T)}{kT} \;=\; \lim_{N\to\infty} \log\, Z_N \;=\; -\log 2 + \log(1-v^2) \quad (5.2.20)$$
$$-\frac{1}{2(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} \log\left[(1+v^2)^2 - 2v(1-v^2)\,(\cos\phi_1+\cos\phi_2)\right]\, d\beta_1\, d\beta_2.$$

This expression coincides with eqn (5.1.16).

References and Further Reading

Combinatorial methods in the solution of the two-dimensional Ising model have been proposed in the papers:

M. Kac, J.C. Ward *A combinatorial solution of the two dimensional Ising model*, Phys. Rev. 88 (1952), 13321337.

N.V. Vdovichenko, *A calculation of the partition function for a plane dipole lattice*, Sov. Phys. JETP 20 (1965), 477.

The exact expression for the magnetization of the two-dimensional Ising model can be found in:

C.N. Yang, *The spontaneous magnetization of a two-dimensional Ising model*, Phys. Rev., 85 (1952), 808.

For the dimer solution see:

M. Fisher *On the dimer solution of the Ising models*, J. Math. Phys., 7 (1996), 1776.

P.W. Kasteleyn, *The statistics of dimers on a lattice: the number of dimer arrangments on a quadratic lattice*, Physica 27 (1961), 1664.

For recent developments on the dimer formalism, we draw the attention of the reader to the following papers:

P. Fendley, *Classical dimers on the triangular lattice*, Phys. Rev. B 66 (2002), 214513.

R. Moessner, S.L. Sondhi, *Ising and dimer models in two and three dimensions*, Phys. Rev. B 68 (2003), 054405.

Problems

1. High-temperature series

Determine the first three terms of the high-temperature expansion of the correlation function $\langle \sigma_{i+2,j}\, \sigma_{i,j} \rangle$ of two spins separated by two lattice sites.

2. Pfaffian and determinant

Prove that for a $2N \times 2N$ antisymmetric matrix A we have the identity

$$(\mathrm{Pf}\, A)^2 \;=\; \det A.$$

3. Number of dimers

a Give an argument to justify the exponential growth of eqn (5.2.16) for the dimer coverings in N, where N is the number of sites of a lattice.

b Use eqn (5.2.16) to estimate the number of dimer coverings of a 4×4 square lattice.

4. Generating function of dimers on a square lattice

Consider the function

$$F(x,y) \;=\; \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} d\beta_1\, d\beta_2 \log\left[x + y - x\cos\beta_1 - y\cos\beta_2\right].$$

Its value at $x = y =$, i.e. $F(2,2)$, provides the solution to the problem of the dimer covering of a square lattice, eqn (5.2.14).

a Prove that

$$F(x,0) \;=\; \log\frac{x}{2}.$$

b Show that we have the identity

$$\frac{\partial F}{\partial x} \;=\; \frac{2}{\pi x} \arctan\sqrt{\frac{x}{y}}.$$

c Expanding in power series the term $\arctan\sqrt{\frac{x}{y}}$ and integrating term by term, show that

$$F(2,2) \;=\; \frac{4\,G}{\pi}$$

where G is the Catalan constant.

6
Transfer Matrix of the Two-dimensional Ising Model

I did much of the work in the writing room of the P & O liner Arcadia, in the Atlantic and Indian Oceans. This was good for concentration, but not for communication.

Rodney J. Baxter

In this chapter we study the solution of the two-dimensional Ising model by means of the transfer matrix. Unlike the methods discussed in the previous chapter, the transfer matrix approach has greater generality and can used to solve exactly other two-dimensional models. Even if the general ideas behind this approach have been explained in Chapter 2 by means of the one-dimensional case, their application to the two-dimensional cases requires more powerful and sophisticated mathematical tools: for instance, the study the eigenvalues of the transfer matrix in the Ising model for $T \neq T_c$ needs to employ elliptic functions. The same is also true for other models. In order to present in the simplest possible way the main lines of this method, in the following we focus attention only on the solution of the model at $T = T_c$ because this case can be analyzed in terms of simple trigonometric functions.

An important condition is required for implementing the method efficiently: the commutativity of the transfer matrix for different values of the coupling constants. In the Ising model, for instance, this condition can be satisfied by the transfer matrix $T_D(K, L)$ along the diagonal of the square lattice. If the coupling constants K and L fulfill the condition

$$\sinh 2K \, \sinh 2L \, = \, \sinh 2K' \, \sinh 2L'. \tag{6.0.1}$$

the transfer matrix has the property[1]

$$[T_D(K, L), T_D(K', L')] \, = \, 0. \tag{6.0.2}$$

Equation (6.0.2) implies that the eigenvectors of the transfer matrix do not change if the coupling constants vary along the curve given by eqn (6.0.1). This is a crucial circumstance for the exact diagonalization of $T_D(K, L)$. Equally important is the possibility to implement the commutativity of the transfer matrices by means of a

[1] $[A, B]$ denotes the commutator of the two matrices A and B and it is given by $[A, B] = AB - BA$.

particular conditions (of local nature) satisfied by the Boltzmann weights. These conditions are known as the Yang–Baxter equations and they play an important role in all exactly solvable models: they enter not only the solution of statistical models but also S-matrix theory, the formalism of quantum groups, and the classification of knots.

6.1 Baxter's Approach

There are several ways to define a transfer matrix for the two-dimensional Ising model and each of them shows certain advantages. The transfer matrix that we discuss in this section is associated to the square lattice rotated by 45 degrees, as shown in Fig. 6.1. The coupling constants K and L, originally placed along the horizontal and vertical directions, are now defined along the diagonals. This lattice is particularly useful to establish the commutativity properties of the transfer matrix defined on it. As is evident from Fig. 6.1, the sites of this lattice can be divided into two classes, A and B, identified by the empty and filled circles: each row of type A is followed by one of type B and vice versa. Let m be the total number of rows: assuming periodic boundary conditions along the vertical direction, m is necessarily an even integer. Moreover, imposing periodic boundary conditions also along the horizontal direction, it is easy to see that there is an equal number n of sites both for the rows of type A and type B. For each row, there are 2^n possible spin configurations and in the following they will be simply denoted by μ_r

$$\mu_r \,=\, \{\sigma_1, \sigma_2, \ldots, \sigma_n\}_{\text{row r}}.$$

Since the spins of type A interact only with those of type B and vice versa, it is convenient to introduce two transfer matrices V and W, both of dimension $2^n \times 2^n$ (see Fig. 6.2). Denoting collectively by μ the spins of the lower row and by μ' those of the upper row, the operators $V(K, L)$ and (K, L) are defined by their matrix elements[2]

$$V_{\mu,\mu'}(K, L) \,=\, \exp\left[\sum_{i=1}^{n}(K\,\sigma_{i+1}\,\sigma'_i + L\,\sigma_i\,\sigma'_i)\right], \tag{6.1.1}$$

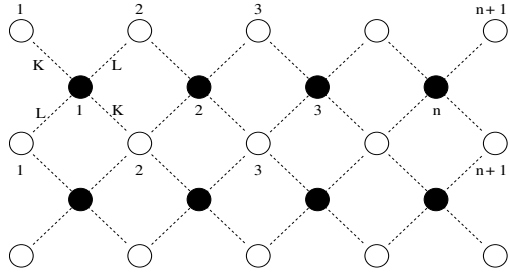


Fig. 6.1 *Square lattice rotated by* 45 *degrees.*

[2]With this choice of matrix elements, the application of two transfer matrices A and B, one after the other, corresponds to their multiplication in the order AB.

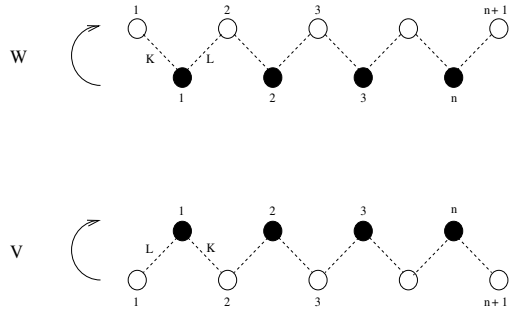


Fig. 6.2 *Transfer matrices V and W.*

$$W_{\mu,\mu'}(K,L) \;=\; \exp\left[\sum_{i=1}^{n}(K\,\sigma_i\,\sigma'_i + L\,\sigma_i\,\sigma'_{i+1})\right]. \tag{6.1.2}$$

In both formulas we have assumed the periodic boundary conditions $\sigma_{n+1} \equiv \sigma_1$ and $\sigma'_{n+1} \equiv \sigma'_1$.

All statistical weights of the model are generated by the iterated application of the operators V and W to the configuration of the first row. The partition function is thus expressed as

$$Z_N(K,L) \;=\; \sum_{\mu_1}\sum_{\mu_2}\dots\sum_{\mu_m} V_{\mu_1,\mu_2}\,W_{\mu_2,\mu_3}\,V_{\mu_3,\mu_4}\dots W_{\mu_m,\mu_1},$$

namely

$$Z_N(K,L) \;=\; \mathrm{Tr}\,(VWVW\dots VW) \;=\; \mathrm{Tr}\,(VW)^{m/2}. \tag{6.1.3}$$

Since the trace of a matrix is independent of its representation, the most convenient way to compute the partition function (6.1.3) consists of diagonalizing the matrix VW, so that

$$Z(K,L) \;=\; \lambda_1^m + \lambda_2^m + \dots + \lambda_{2^n}^m, \tag{6.1.4}$$

where λ_1^2, λ_2^2, ... are the eigenvalues of VW. In the thermodynamic limit (where both m and n go to infinity) it is only the maximum eigenvalue that matters because, taking initially the limit $m \to \infty$, with n finite, we have[3]

$$Z(K,L) \;=\; (\lambda_{max})^m\left[1 + \left(\frac{\lambda_1}{\lambda_{max}}\right)^m + \left(\frac{\lambda_2}{\lambda_{max}}\right)^m + \dots\right] \;\simeq\; (\lambda_{max})^m\,. \tag{6.1.5}$$

So we arrive at a formula that is quite analogous to the one-dimensional Ising model. From an algebraic point of view, though, there is a substantial difference between the two cases: while in the one-dimensional case the problem consists of diagonalizing a

[3]With real coupling constants, the matrix VW has all matrix elements positive. The matrices that share such a property are knows as *positive matrices*. The Perron–Frobenius theorem, whose proof is proposed as a problem at the end of the chapter, states that any finite-dimensional positive matrix has a unique maximum eigenvalue, also positive. The corresponding eigenvector has all its components positive as well.

2×2 matrix, in the two-dimensional case it is necessary to find the eigenvalues of a $2^n \times 2^n$ matrix, in the limit $n \to \infty$. The mathematical difficulty of such a problem can be faced by taking advantage of some important properties of the transfer matrices.

6.1.1 Commutativity of the Transfer Matrices

The operators V and W explicitly depend on the coupling constants of the lattice, as shown by their definition (6.1.1) and (6.1.2). Consider now the product of V with W but with different coupling constants, as shown in Fig. 6.3

$$V(K, L)\, W(K', L'). \tag{6.1.6}$$

Denoting by $\mu = \{\sigma_1, \ldots, \sigma_n\}$ the spins of the lower row, by $\mu' = \{\sigma'_1, \ldots, \sigma'_n\}$ the spins of the upper row and by $\mu'' = \{\sigma''_1, \ldots, \sigma''_n\}$ those of the half-way row, the matrix elements of this operator between the states μ and μ' are obtained according to the usual rule of the product of matrices, namely as a sum over the intermediate states μ''

$$\left(V(K, L)\; W(K', L')\right)_{\mu,\mu'} \;=\; \sum_{\{\sigma''\}} \prod_{j=1}^{n} \exp\left[\sigma''_j (K\sigma_{j+1} + L\sigma_j + K'\sigma'_j + L'\sigma'_{j+1})\right].$$

Since each intermediate spin $\sigma"_j$ appears only in a single term of the expression,[4] the sum over these spins is particularly simple and the matrix elements of the operator (6.1.6) assume the factorized form

$$\left(V(K, L)\; W(K', L')\right)_{\mu,\mu'} \;=\; \prod_{j=1}^{n} X(\sigma_j, \sigma_{j+1}; \sigma'_j, \sigma'_{j+1}), \tag{6.1.7}$$

with the elementary statistical weight $X(a, b, c, d)$ explicitly given by (see Fig. 6.4)

$$\begin{aligned} X(a, b, c, d) \;&=\; \sum_{\sigma''=\pm 1} \exp\left[\sigma''(La + Kb + K'c + L'd)\right] \\ &=\; 2\, \cosh\left[La + Kb + K'c + L'd\right], \quad a, b, c, d = \pm 1. \end{aligned} \tag{6.1.8}$$

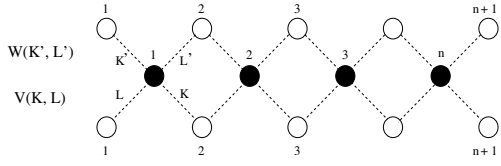


Fig. 6.3 *Product of V and W with different coupling constants.*

[4]This is one of the mathematical advantages of the transfer matrix defined on the diagonal of the lattice.

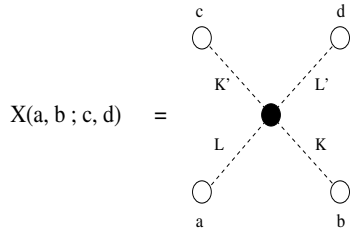


Fig. 6.4 *Elementary statistical weight* $X(a, b; c, d)$.

Exchanging the role of the coupling constans (K, L) and (K', L'), one obtains, in general, a different result for the product VW. There is, however, the identity

$$V(K, L)\, W(K', L') \;=\; V(K', L')\, W(K, L), \tag{6.1.9}$$

if the coupling constants satisfy the equation

$$\sinh 2K \, \sinh 2L \;=\; \sinh 2K' \, \sinh 2L'. \tag{6.1.10}$$

To prove this result, let's observe that for the factorized form (6.1.7) of the product VW, the transformation

$$X(a, b, c, d) \;\longrightarrow\; e^{Mac}\; X(a, b, c, d)\; e^{-Mbd}$$

does not change the expression (6.1.7). This observation permits us to satisfy eqn (6.1.9) by solving a simpler problem, i.e. the problem to find a number M such that

$$e^{Mac}\; X(a, b; c, d) \;=\; X'(a, b; c, d)\; e^{Mbd}, \tag{6.1.11}$$

where X' is the statistical weight obtained by changing $K \to K'$ and $L \to L'$ in the original X. In summary, in order to satisfy the *global* commutativity condition (6.1.9), it is sufficient to find a solution to the *local* condition (6.1.11). This problem can be solved by using the star–triangle identity discussed in Section 4.3.1. Let us consider, in fact, the graphical representation of eqn (6.1.11) given in Fig. 6.5a: both in the right and left diagrams there is a triangle, given by the interaction of the relative spins. Imposing

$$K_1 = L, \quad K_2 = K', \quad K_3 = M$$

and changing each triangle into a star, with the relative coupling constants L_i given by eqn (4.4.7), it is easy to see by looking at Fig. 6.5b that the two expressions are equal if

$$L_1 \;=\; K, \quad L_2 \;=\; L',$$

namely, if the coupling constants satisfy the condition

$$\sinh 2K \, \sinh 2L \;=\; \sinh 2K' \, \sinh 2L'. \tag{6.1.12}$$

Equation (6.1.9) can be further elaborated and entirely expressed in terms of the matrix V. Thanks to the periodic boundary conditions, it is in fact evident that W

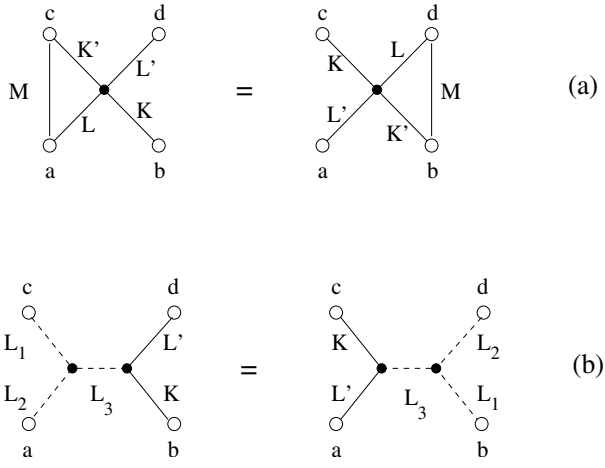


Fig. 6.5 *Star–triangle transformation of eqn (6.1.11), where the sum over the spins is represented by the black circles.*

differs from V simply by a translation of a lattice spacing. With the help of the operator T, with matrix elements

$$T_{\mu,\mu'} \;=\; \delta(\sigma_1,\sigma_2')\,\delta(\sigma_2,\sigma_3')\;\ldots\;\delta(\sigma_n,\sigma_1'), \tag{6.1.13}$$

and whose effect is to move the lattice of a lattice spacing to the right, one can verify that

$$W(K,L) \;=\; V(K,L)\,T. \tag{6.1.14}$$

Moreover

$$V(K,L) \;=\; T^{-1}\,V(K,L)\,T, \qquad W(K,L) \;=\; T^{-1}\,W(K,L)\,T. \tag{6.1.15}$$

Using (6.1.14), eqn (6.1.9) becomes

$$V(K,L)\,V(K',L') \;=\; V(K',L')\,V(K,L), \tag{6.1.16}$$

where the coupling constants satisfy eqn (6.0.1).

6.1.2 Commutativity of the Transfer Matrices: Graphical Proof

The commutativity relation (6.1.16) can be proved in a graphical way. To this end we must first consider the square lattice in its usual orientation and then define two sets of operators $P_i(K)$ and $Q_i(L)$ by means of their matrix elements

$$\begin{aligned}(P_i(K))_{\mu,\mu'} \;&=\; \exp[K\sigma_i\sigma_{i+1}]\,\delta(\sigma_1,\sigma_1')\ldots\delta(\sigma_n,\sigma_n')\\ (Q_i(L))_{\mu,\mu'} \;&=\; \delta(\sigma_1,\sigma_1')\ldots\delta(\sigma_{i-1},\sigma_{i-1}')\exp[L\sigma_i\sigma_i']\\ &\quad\times\delta(\sigma_{i+1},\sigma_{i+1}'\ldots\delta(\sigma_n,\sigma_n').\end{aligned} \tag{6.1.17}$$

$P_i(K)$ creates the statistical weight of the spins σ_i and σ_{i+1} placed on the same horizontal row (without changing their values from the row μ to the next one), while

$V_i(L)$ creates the statistical weight of the spins σ_i and σ'_i, placed on the next neighbor two rows. The result of these operators is visualized in Fig. 6.6. It is possible to adopt a uniform notation by defining the operators $U_i(K, L)$

$$U_i(K,L) \;=\; \begin{cases} P_j(K), & i = 2j \\ Q_j(L), & i = 2j-1. \end{cases} \tag{6.1.18}$$

These operators satisfy

$$U_i(K,L)\, U_j(K',L') \;=\; U_j(K',L')\, U_i(K,L), \qquad |\, i-j \,| \geq 2. \tag{6.1.19}$$

Suppose we are dealing with a set of coupling constants (K_1, K_2, K_3) and (L_1, L_2, L_3) linked to one another by the star–triangle relation (4.4.7):

$$\sinh 2K_1 \,\sinh 2L_1 \;=\; \sinh 2K_2 \,\sinh 2L_2 \;=\; \sinh 2K_3 \,\sinh 2L_3 \;=\; h^{-1}. \tag{6.1.20}$$

Using the explicit expression of the matrix elements of P_i and Q_i, it is easy to show that

$$U_{i+1}\, U'_i\, U''_{i+1} \;=\; U''_i\, U'_{i+1}\, U_i, \tag{6.1.21}$$

where we have introduced the notation $U_i = U_i(K_1, L_1)$, $U'_i = U_i(K_2, L_2)$ and $U''_i = U_i(K_3, L_3)$. The graphical interpretation of this equation is given in Fig. 6.7. Let us

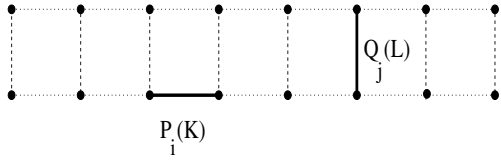


Fig. 6.6 *Action of the operators $P_i(K)$ and $Q_j(L)$ on the square lattice.*

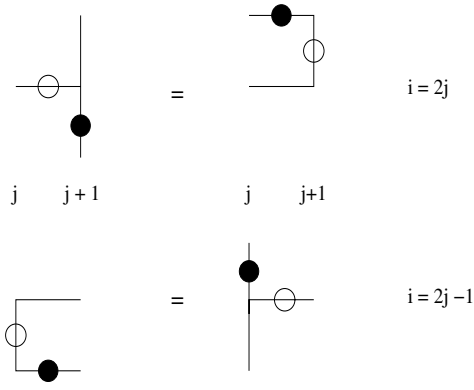


Fig. 6.7 *Graphical form of eqn (6.1.21). The full circle corresponds to the couplings (K_1, L_1), the empty circle to (K_2, L_2), and the line to (K_3, L_3).*

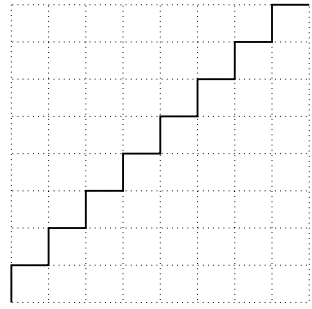


Fig. 6.8 *The operator $V(K,L)$ on the square lattice.*

consider now the operator $V(K,L)$ given by the product

$$V(K,L) \,=\, U_1(K,L)\,U_2(K,L)\,\ldots U_N(K,L), \tag{6.1.22}$$

with $N = 2n$: its action consists of introducing the statistical weights along the main diagonal of the lattice as shown in Fig. 6.8. It is easy to see that $V(K,L)$ coincides with the transfer matrix considered in the previous sections.

Let (K_i, L_i) $(i = 1,2,3)$ be three different pairs of coupling constants that satisfy the star–triangle equation (6.1.20) and let's define

$$V \,=\, U_1\,U_2\,\ldots\,U_N, \quad V' \,=\, U_1'\,U_2'\,\ldots\,U_N'.$$

Using iteratively eqns (6.1.19) and (6.1.21), one can show that these operators satisfy the condition

$$V\,V'(U_N'^{-1}\,U_N''\,U_N) \,=\, (U_1\,U_1''\,U_1'^{-1})\,V'\,V. \tag{6.1.23}$$

The graphical proof is given in Fig. 6.7, where the sequence of diagrams is generated by the repeated application of the graphical identities of Fig. 6.9. The terms within the parentheses of eqn (6.1.23) refer to the spins at the boundary and they disappear if we adopt periodic boundary conditions. In this case we have then the commutativity relation (6.1.16)

$$V(K,L)\,V(K',L') \,=\, V(K',L')\,V(K,L). \tag{6.1.24}$$

6.1.3 Functional Equations and Symmetries

The factorized form (6.1.7) of VW allows us to write down a functional equation for the matrix elements of this operator. Consider the elementary statistical weight $X(a,b\,;\,c,d)$, given by the formula (6.1.8). For the values

$$K' \,=\, L + \frac{i\pi}{2}, \quad L' \,=\, -K, \tag{6.1.25}$$

we have

$$X(a,b\,;\,c,d) \,=\, 2\,\cosh\left[L(a+c) + K(b-d) + \frac{i\pi c}{2}\right] \,=\, ic\,\sinh\left[L(a+c) + K(b-d)\right]. \tag{6.1.26}$$

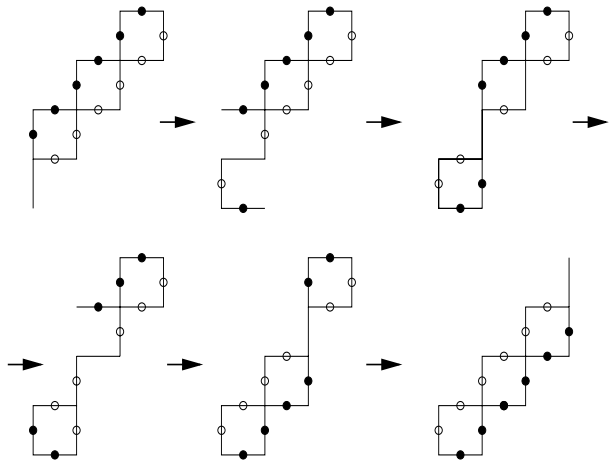


Fig. 6.9 *Graphical proof of the commutativity relation of the transfer matrices along the diagonal of the square lattice.*

Hence, it is different from zero only in two cases:

- $a = c$ and $b = d$, where we have

$$X(a, b; a, b) \;=\; 2ia \,\sinh 2La \;=\; 2i \,\sinh 2L;$$

- or $a \neq c$ and $b \neq d$, and in this case

$$X(a, b; -a, -b) \;=\; -2ia \,\sinh 2Kb \;=\; -2iab \,\sinh 2K.$$

Corresponding to the particular values (6.1.25) of the coupling constants, the matrix elements of VW are expressed as

$$\begin{aligned}\left(V(K, L)\; W\left(L + \frac{i\pi}{2}, -K\right)\right)_{\mu,\mu'} \;&=\; (2i \sinh 2L)^n\, \delta(\sigma_1, \sigma_1')\, \delta(\sigma_2, \sigma_2') \,\ldots \delta(\sigma_n, \sigma_n') \\ &+(-2i \sinh 2K)^n\, \delta(\sigma_1, -\sigma_1')\, \delta(\sigma_2, -\sigma_2') \,\ldots \delta(\sigma_n, -\sigma_n'). \end{aligned} \tag{6.1.27}$$

If we introduce the identity operator $\mathbf{I}$, with matrix elements

$$\mathbf{I}_{\mu,\mu'} \;=\; \delta(\sigma_1, \sigma_1')\, \delta(\sigma_2, \sigma_2') \,\ldots \delta(\sigma_n, \sigma_n'), \tag{6.1.28}$$

and the operator $\mathbf{R}$, with matrix elements

$$\mathbf{R}_{\mu,\mu'} \;=\; \delta(\sigma_1, -\sigma_1')\, \delta(\sigma_2, -\sigma_2') \,\ldots \delta(\sigma_n, -\sigma_n') \tag{6.1.29}$$

(both matrices have dimension $2^n \times 2^n$), eqn (6.1.27) can be written in an operatorial form as

$$V(K,L)\; W\left(L+\frac{i\pi}{2},-K\right) \;=\; (2i\sinh 2L)^n\,\mathbf{I} + (-2i\sinh 2K)^n\,\mathbf{R}. \tag{6.1.30}$$

As we will see in the next section, this formula is extremely useful to determine the eigenvalues of the matrices V and W, and to find the inverse of the matrix $V(K,L)$ (see Problem 2 at the end of the chapter).

Let's discuss the symmetry properties of the matrices V and W. Interchanging K with L and σ_i with σ_i', the matrix W becomes the transpose of V:

$$W(K,L) = V^T(L,K), \tag{6.1.31}$$

$$V(K,L)\,W(K,L) = \left[V(L,K)\,W(L,K)\right]^T. \tag{6.1.32}$$

Since changing the sign of K and L is equivalent to changing the sign of $\sigma_1,\ldots,\sigma_n$ or $\sigma_1',\ldots,\sigma_n'$, we also have

$$V(-K,-L) \;=\; R\,V(K,L) \;=\; V(K,L)\,R, \tag{6.1.33}$$

with a similar relation for the matrix W.

Let p be the number of spin pairs (σ_{j+1},σ_j') with opposite value and q the number of spin pairs (σ_j,σ_j') with opposite values. Hence, $p+q$ counts the total number of changes of signs that we have in the sequence $\sigma_1,\sigma_1',\sigma_2,\sigma_2',\ldots,\sigma_n'$. So $p+q$ is an even number and, from the definition (6.1.1), it follows that

$$V_{\mu,\mu'}(K,L) \;=\; \exp\left[(n-2p)\,K + (n-2q)\,L\right]. \tag{6.1.34}$$

In the thermodynamic limit $n\to\infty$ there is no difference whether n is an even or an odd number and, imposing for simplicity

$$n \;=\; 2s, \tag{6.1.35}$$

where s is an integer, eqn (6.1.34) can be written in terms of two numbers p' and q' that belong to the interval $(0,s)$

$$V_{\mu,\mu'}(K,L) \;=\; \exp\left[\pm 2p'\,K \pm 2q'\,L\right]. \tag{6.1.36}$$

The variables p' and q' are either both even or odd, so the matrix $V(K,L)$ satisfies the relation

$$V\left(K \pm i\frac{\pi}{2}, L \pm i\frac{\pi}{2}\right) \;=\; V(K,L), \tag{6.1.37}$$

with a similar relation for $W(K,L)$.

6.1.4 Functional Equations for the Eigenvalues

Let us proceed to the determination of the eigenvalues of $V(K, L)$ by using the functional equations satisfied by this operator. Suppose that K and L are two complex numbers subject to the condition

$$h^{-1} \,=\, \sinh 2K \, \sinh 2L, \tag{6.1.38}$$

where h is a given real number. In this case, thanks to eqn (6.1.38), there is an infinite number of transfer matrices that commute with each other, see eqn (6.1.16). They also commute with T, eqn (6.1.15), and with R, eqn (6.1.33). These commutation properties imply that, for all the values of K and L that satisfy eqn (6.1.38), the transfer matrices have a common basis of eigenvectors. These eigenvectors can depend neither on K nor on L, but they can be functions of h. Denoting by $y(h)$ one of these eigenvectors and by $v(K, L)$, t, and r the eigenvalues of the matrices $V(K, L)$, T, and R, we have

$$\begin{aligned} V(K, L)\, y(h) &= v(K, L)\, y(h); \\ T\, y(h) &= t\, y(h); \\ R\, y(h) &= r\, y(h). \end{aligned} \tag{6.1.39}$$

The eigenvalues t and c also satisfy

$$t^n \,=\, r^2 \,=\, 1, \tag{6.1.40}$$

so they are complex numbers of unit modulus, independent of K and L. Notice that if K and L satisfy eqn (6.1.38), the same happens with K' and L' defined in (6.1.25). Hence, applying the functional relation (6.1.30) to the vector $y(h)$, we have

$$v(K, L)\, v\left(L + \frac{i\pi}{2}, -K\right)\, t \,=\, (2i \sinh 2L)^n + (-2i \sinh 2K)^n\, r. \tag{6.1.41}$$

Let $\lambda^2(K, L) \equiv \lambda_i^2$ be one of the eigenvalues[5] of the matrix $V(K, L)\, W(K, L)$. Since $y(h)$ is also an eigenvector of this matrix and $W = V\, T$, we have

$$\lambda^2(K, L) \,=\, v^2(K, L)\, t. \tag{6.1.42}$$

With the definition

$$\lambda(K, L) \,=\, v(K, L)\, \sqrt{t}, \tag{6.1.43}$$

eqn (6.1.41) becomes a functional equation that has to be satisfied by the eigenvalues of the transfer matrix

$$\lambda(K, L)\, \lambda\left(L + \frac{i\pi}{2}, -K\right) \,=\, (2i \sinh 2L)^n + (-2i \sinh 2K)^n\, r. \tag{6.1.44}$$

[5] In the following we will omit, for brevity, the index i. The different eigenvalues will be identified by the different solutions of the functional equation (6.1.44).

6.2 Eigenvalue Spectrum at the Critical Point

In this section we show how it is possible to determine the spectrum of the transfer matrix only using the commutativity property and the analytic structure of the eigenvalues, together with the functional equation (6.1.44). A crucial aspect of the solution is the appropriate parameterization of the coupling constants K and L that satisfy eqn (6.1.38): a clever parameterization will allow us to take advantage of the powerful theorems of complex analysis and to extract the analytic properties of the eigenvalues.

The actual implementation of this program presents a different level of complexity according to the value of the parameter h. In order to highlight the main steps of such a method, it is convenient to discuss the simplest case:[6] this corresponds to the value $h = 1$ for which the system is at the critical point

$$\sinh 2K \, \sinh 2L \, = \, 1 \tag{6.2.1}$$

(see Chapter 4 and, in particular, Section 4.2.3). Equation (6.2.1) can be identically satisfied by imposing

$$\begin{array}{l} \sinh 2K \, = \, \tan u, \\ \sinh 2L \, = \, \cot u. \end{array} \tag{6.2.2}$$

The coupling constants K and L are both real and positive for the values u that fall in the range $(0, \frac{\pi}{2})$. The parameterization (6.2.2) allows us to write $\exp(\pm 2K)$ and $\exp(\pm 2L)$ as

$$\begin{array}{r} \exp(2K) = (1 + \sin u)/\cos u, \\ \exp(-2K) = (1 - \sin u)/\cos u, \\ \exp(2L) = (1 + \cos u)/\sin u, \\ \exp(2L) = (1 - \cos u)/\sin u. \end{array} \tag{6.2.3}$$

These expressions have the following important properties:

1. they are periodic functions of u, with period period 2π;
2. they are meromorphic functions[7] of u, with simple poles.

Since the eigenvalues $\lambda(K, L)$ of the transfer matrix can be regarded as functions of u, it is convenient to adopt the notation $\lambda(u)$ and write the functional equation (6.1.44) as

$$\lambda(u) \, \lambda(u + \frac{\pi}{2}) \, = \, (2i \cot u)^n + (-2i \tan u)^n \, r. \tag{6.2.4}$$

Expressing $\exp(\pm 2K)$ and $\exp(\pm 2L)$ in terms of the functions (6.2.3), the matrix elements of $V_{\mu,\mu'}$ assume the form

$$V_{\mu,\mu'} \, = \, \frac{A(u)}{(\sin u \, \cos u)^s}, \tag{6.2.5}$$

[6] In the general case one has to use a parameterization in terms of elliptic functions.

[7] A meromorphic function has only poles as singularities in the complex plane.

where $A(u)$ is a polynomial in $\sin u$ and $\cos u$, of total degree $2s$. Hence its general expression is given by

$$A(u) \;=\; e^{-2isu}\,\left(a_0 + a_1\, e^{iu} + \cdots + a_{2n}\, e^{4isu}\right). \tag{6.2.6}$$

Let's now consider the first equation in (6.1.39), which actually consists of 2^n equations. Using known theorems of linear algebra, the eigenvalues $v(K,L)$ are expressed as linear combinations of the matrix elements of $V(K,L)$, whose coefficients are given by ratios of the components of the eigenvectors $y(h)$. For the commutativity of all matrices involved in the problem, such ratios are functions only of the variable h *but totally independent of* u. This is a crucial property for the considerations that follow because it implies that each eigenvalue $v(K,L)$ is expressed by a linear combination of terms as (6.2.5) and therefore it has the same form. The same is true for $\lambda(u)$, defined in (6.1.43).

Notice that replacing u with $u+\pi$ is equivalent to changing K in $-K \pm i\frac{\pi}{2}$ and L in $-L \pm i\frac{\pi}{2}$, as evident from eqns (6.2.3). However, these substitutions are equivalent to multiplying V by R, as one can see from eqn (6.1.33). Hence, denoting $v(K,L)$ by $v(u)$, the first of the equations (6.1.39) becomes

$$V(K,L)\; R\; y(h) \;=\; v(u+\pi)\, y(h), \tag{6.2.7}$$

where we have taken into account once again the independence of $y(h)$ of the variable u. Using the first and the last equation in (6.1.39), we have

$$v(u+\pi) \;=\; r\, v(u),$$

namely

$$\lambda(u+\pi) \;=\; r\,\lambda(u). \tag{6.2.8}$$

Since the generic form of $\lambda(u)$ is given by (6.2.5) and $r = \pm 1$, for the periodicity (6.2.8) the corresponding polynomial $A(u)$ in (6.2.6) only has the even coefficients c_{2k} different from zero when $r=1$, while it only has the odd coefficients c_{2k+1} different from zero when $r=-1$. Then the eigenvalues $\lambda(u)$ can be expressed as

$$\lambda(u) \;=\; \rho\,(\sin u \cos u)^{-s} \prod_{j=1}^{l} \sin(u-u_j) \tag{6.2.9}$$

where ρ and $u_1, u_2, \ldots, u_l$ are constants to be determined, with

$$l \;=\; \begin{cases} 2s, & \text{if } r=+1 \\ 2s-1, & \text{if } r=-1. \end{cases}$$

Substituting this expression into the functional equation (6.2.4), we have

$$\rho^2 \prod_{j=1}^{l} \sin(u-u_j)\,\cos(u-u_j) \;=\; 2^{2s}\,\left(\cos^{4s} u + r\sin^{4s} u\right). \tag{6.2.10}$$

This identity must be satisfied for all values of u. This expression can be simplified by the substitution

$$x \,=\, e^{2iu}, \quad x_j \,=\, e^{2iu_j}.$$

We then have

$$\rho^2 \left(\frac{i}{4}\right)^l \prod_{j=1}^{l} \frac{(x^2 - x_j^2)}{x_j} \,=\, 2^{-2s}\, x^{l-2s} \left[(x+1)^{4s} + r\,(x-1)^{4s}\right]. \tag{6.2.11}$$

Both polynomials on the right- and on the left-hand sides are of degree l in the variable x^2 and therefore the constants ρ and $x_1, \ldots, x_l$ are determined by the identity of these two polynomials. Since $x_1^2, \ldots, x_l^2$ are the l distinct zeros of the left term, the same should hold for the term on the right-hand side. So, they are fixed by the condition

$$\left[(x+1)^{4s} + r(x-1)^{2s}\right] \,=\, 0,$$

whose solutions are given by

$$x_j^2 \,=\, -\tan^2 \frac{\theta_j}{2}, \quad j = 1, \ldots, l$$

where

$$\theta_j \,=\, \begin{cases} \pi\,(j - \frac{1}{2})/2s, & \text{if } r = +1 \\ \pi\, j/2s, & \text{if } r = -1. \end{cases}$$

All these values of θ_j fall in the range $(0, \pi)$, so that, defining

$$\varphi_j \,=\, \frac{1}{2} \ln \tan \frac{\theta_j}{2}, \quad j = 1, \ldots, l$$

we have

$$u_j \,=\, \mp \frac{\pi}{4} - i\,\varphi_j, \quad j = 1, \ldots, l. \tag{6.2.12}$$

Since the sign ∓ 1 of each solution can be chosen independently, there are 2^l possible solutions. There is, however, an extra condition coming from the limits $u \to \pm i\,\infty$, where $\exp(2K) = \exp(2L) \to \pm i$. Since the matrix elements of the transfer matrix do not change if we alter the sign of $\exp(2K)$ and $\exp(2L)$, we have

$$\lambda(i\,\infty) \,=\, \lambda(-i\,\infty).$$

From the general expression of the eigenvalues, eqn (6.2.9), one can check that this condition is automatically satisfied when $r = -1$, while if $r = 1$, it leads to the condition

$$(u_1 + \cdots + u_{2s})/\pi \,=\, N + \frac{1}{2}s,$$

where N is an integer. This implies that only $2s - 1$ among the possible signs of the solutions (6.2.12) can be chosen in an independent way. Therefore, as expected, in both cases $r = \pm 1$ there are 2^{2s-1} eigenvalues λ.

To summarize, the eigenvalues $\lambda(u)$ are given by

$$\lambda(u) \;=\; \rho\,(\sin u\,\cos u)^{-s}\,\prod_{j=1}^{l}\sin\left(u+i\varphi_j+\frac{1}{4}\eta_j\pi\right), \tag{6.2.13}$$

where $\eta_1,\ldots,\eta_l$ have values ±1 and, for $r=1$ there is the further condition

$$\eta_1+\cdots+\eta_{2s} \;=\; 2s-4M, \tag{6.2.14}$$

where M is an integer.

6.3 Away from the Critical Point

The analysis done for the eigenvalues at the critical point $T=T_c$ can also be performed for generic values of T. As previously mentioned, this requires a parameterization in terms of the elliptic functions and will not be pursued here. We only mention that this analysis leads to the determination of the maximum eigenvalues of the transfer matrix whose final expression is given by

$$\log\lambda_{max} \;=\; \frac{1}{2}\sum_{j=1}^{2s}\mathcal{F}\left[\pi\left(j-\frac{1}{2}\right)/2s\right], \tag{6.3.1}$$

where the function $\mathcal{F}(\theta)$ is

$$\mathcal{F}(\theta) \;=\; \log\left\{2\,\left[\cosh 2K\,\cosh 2L+h^{-1}(1+h^2-2h\cos 2\theta)^{1/2}\right]\right\}. \tag{6.3.2}$$

In the thermodynamic limit, when $s\to\infty$, the free energy is given by

$$-F/k_BT \;=\; \frac{1}{2\pi}\int_0^{\pi}F(\theta)\,d\theta. \tag{6.3.3}$$

The analysis of the singularity that arises in this expression when $h\to1$ is proposed as an exercise.

6.4 Yang–Baxter Equation and R-matrix

At the heart of the solvability of many lattice statistical models there is the commutativity of the transfer matrix that, as a sufficient condition, needs the Yang–Baxter equation satisfied by the Boltzmann weights R. Let's elaborate on this problem in more abstract terms. Consider Fig. 6.10, where each of the lines stands for a vector space spanned by the statistical variables. Let's denote the three vector spaces by $V^{\mu}_{p_1}$, $V^{\nu}_{p_2}$, and $V^{\lambda}_{p_3}$, with μ,ν, and λ that label the different multiplets and the p_i's that denote the spectral parameters of the Boltzmann weights. Two or more adjacent lines, for example those representing the spaces $V^{\mu}_{p_1}$ and $V^{\nu}_{p_2}$, are tensor products of those spaces, $V^{\mu}_{p_1}\otimes V^{\nu}_{p_2}$. The Boltzmann weight R, associated to the operation of crossing

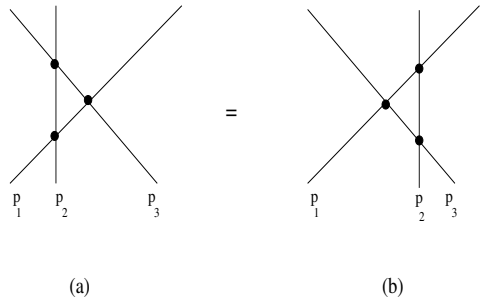


Fig. 6.10 *Yang–Baxter equation satisfied by the Boltzmann weights (here represented by the dots) as functions of the spectral parameter p.*

the lines in the diagram, can be abstractly described as a mapping from a vector space of the initial states to the vector space of the final state,

$$R^{\mu\nu}(p_1 - p_2) \,:\, V^{\mu}_{p_1} \otimes V^{\nu}_{p_2} \,\rightarrow\, V^{\nu}_{p_2} \otimes V^{\mu}_{p_1}. \tag{6.4.1}$$

Here it is assumed that, from the homogeneity of the lattice, the Boltzmann weights depends only on the difference $p_1 - p_2$ of the spectral parameters. This matrix is usually referred to as the R-matrix and satisfies the Yang–Baxter equation of Fig. 6.10

$$\begin{aligned}&(R^{\mu\nu}(p_1 - p_2) \otimes 1)(1 \otimes R^{\mu\lambda}(p_1 - p_3))(R^{\nu\lambda}(p_2 - p_3) \otimes 1)\\ &= (1 \otimes R^{\nu\lambda}(p_2 - p_3))(R^{\mu\lambda}(p_1 - p_3) \otimes 1)(1 \otimes R^{\mu\nu}(p_1 - p_2)).\end{aligned} \tag{6.4.2}$$

The Yang–Baxter equation is nonlinear and usually it it is difficult to solve directly. Nevertheless its solution has been found for many lattice models, leading to the exact determination of their free energy. An essential property is the invariance of R under a quantum group symmetry, a topic that will be discussed in more detail in Section 18.9, whereas further aspects of R-matrices and the Yang–Baxter equation can be found throughout the literature quoted at the end of the chapter. Here we present the main features of this formalism through the study of a significant example.

6.4.1 Six-vertex Model

Consider a square $N \times N$ lattice where the fluctuating variables α are attached to each bond connecting the nearest-neighbor lattice sites. The vertex Boltzmann weight $R^{\gamma\delta}_{\alpha\beta}$ corresponds to each configuration around any lattice site

$$R^{\gamma\delta}_{\alpha\beta} \;=\; \gamma \,-\, \begin{array}{c}\delta\\ |\\ \beta\end{array} \,-\, \alpha$$

Denoting the energy of the vertex by $\epsilon(\alpha, \beta, \gamma, \delta)$, one has $R^{\gamma\delta}_{\alpha\beta} = \exp\left[-\epsilon(\alpha, \beta, \gamma, \delta)/k_B T\right]$. In the six-vertex model each bond can accept one of the two states characterized by an incoming or outgoing arrow associated to the values $\alpha = \pm$. Furthermore, the

only allowed configurations of this model are those in which there are two incoming and two outgoing arrows at each vertex, i.e.

$$R^{++}_{++} = \leftarrow \begin{smallmatrix}\uparrow\\\uparrow\end{smallmatrix} \leftarrow,\ R^{--}_{--} = \rightarrow \begin{smallmatrix}\downarrow\\\downarrow\end{smallmatrix} \rightarrow$$

$$R^{+-}_{+-} = \leftarrow \begin{smallmatrix}\downarrow\\\downarrow\end{smallmatrix} \leftarrow,\ R^{-+}_{-+} = \rightarrow \begin{smallmatrix}\uparrow\\\uparrow\end{smallmatrix} \rightarrow$$

$$R^{+-}_{-+} = \leftarrow \begin{smallmatrix}\downarrow\\\uparrow\end{smallmatrix} \rightarrow,\ S^{-+}_{+-} = \rightarrow \begin{smallmatrix}\uparrow\\\downarrow\end{smallmatrix} \leftarrow$$

A configuration of the system in shown in Fig. 6.11. Assuming invariance under $+ \Leftrightarrow -$, we can parameterize the Boltzmann weights as

$$\begin{aligned} R^{++}_{++} &= R^{--}_{--} = a = \sin(\gamma - p) \\ R^{+-}_{+-} &= R^{-+}_{-+} = b = \sin p \\ R^{+-}_{-+} &= R^{-+}_{+-} = c = \sin \gamma \end{aligned}$$

where p is the spectral parameter whereas γ is the coupling constant. The weights can be arranged as a 4×4 matrix

$$R^{\gamma\delta}_{\alpha\beta} = \begin{pmatrix} \leftarrow \begin{smallmatrix}\uparrow\\\uparrow\end{smallmatrix} \leftarrow & & & \\ & \leftarrow \begin{smallmatrix}\downarrow\\\downarrow\end{smallmatrix} \leftarrow & \leftarrow \begin{smallmatrix}\downarrow\\\uparrow\end{smallmatrix} \rightarrow & \\ & \rightarrow \begin{smallmatrix}\uparrow\\\downarrow\end{smallmatrix} \leftarrow & \rightarrow \begin{smallmatrix}\uparrow\\\uparrow\end{smallmatrix} \rightarrow & \\ & & & \rightarrow \begin{smallmatrix}\downarrow\\\downarrow\end{smallmatrix} \rightarrow \end{pmatrix} = \begin{pmatrix} a & & & \\ & b & c & \\ & c & b & \\ & & & a \end{pmatrix} \tag{6.4.3}$$

It is not difficult to check that this R-matrix satisfies the Yang–Baxter equation (6.4.2). To express the partition function in terms of the matrix R, let's define the *monodromy matrix* (the sum over the repeated indices is implicit):

$$L^{\gamma\{\delta\}}_{\alpha\{\beta\}}(p,\gamma) \equiv R^{\gamma\delta_1}_{\alpha_2\beta_1}(p,\gamma)\, R^{\alpha_2\delta_2}_{\alpha_3\beta_2}(p,\gamma) \cdots R^{\alpha_N\delta_N}_{\alpha\beta_N}(p,\gamma) \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{6.4.4}$$

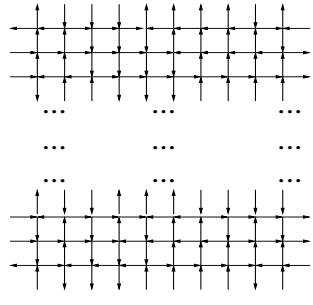
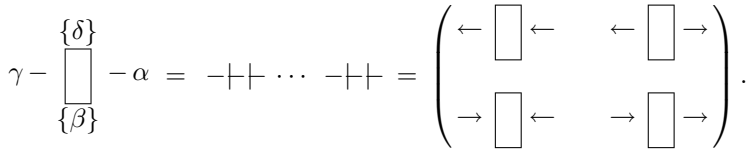


Fig. 6.11 *A configuration of the six-vertex model with periodic boundary conditions.*

In this formula we have a matrix product with respect to the *horizontal space* but a tensor product with respect to the N *vertical space*. Therefore the final result is a 2×2 matrix with entries that are operators in $V_N = \otimes_{k=1}^{N} V_v^{(k)}$ ($V_v^{(k)}$ is the vertical space associated to the k-th column, in our case $V_v^{(k)} = \mathbf{C}^2$). The graphical form of the monodromy matrix is



With periodic boundary conditions along the horizontal and vertical axes, the transfer matrix of the model is

$$T(p,\gamma) \;=\; \mathrm{Tr}_h\, L(p,\gamma), \tag{6.4.5}$$

and the partition function is the trace in the tensor product of the vertical space

$$Z(p,\gamma) \;=\; \mathrm{Tr}_v [T(p,\gamma)]^N. \tag{6.4.6}$$

Since the R-matrix satisfies the Yang–Baxter equation (6.4.2), the monodromy matrix satisfies

$$R_{\alpha'\beta'}^{\alpha''\beta''}(p-p')\, L_{\alpha\{\gamma''\}}^{\alpha'\{\gamma''\}}(p)\, L_{\beta\{\gamma\}}^{\beta'\{\gamma'\}} \;=\; L_{\beta\{\gamma''\}}^{\beta''\{\gamma''\}}(p')\, L_{\alpha'\{\gamma\}}^{\alpha''\{\gamma'\}}(p)\, R_{\alpha\beta}^{\alpha'\beta'}(p-p'). \tag{6.4.7}$$

This implies that the operators A, B, C, and D of the monodromy matrix satisfy the commutation relations

$$\begin{aligned} A(p)\, B(p') \;&=\; \frac{a(p'-p)}{b(p'-p)}\, B(p')\, A(p) - \frac{c(p'-p)}{b(p'-p)}\, B(p)\, A(p') \\ D(p)\, B(p') \;&=\; \frac{a(p-p')}{b(p-p')}\, B(p')\, D(p) - \frac{c(p'-p)}{b(p'-p)}\, B(p)\, D(p') \\ [C(p), B(p')] \;&=\; \frac{c(p-p')}{b(p-p')} - A(p)\, D(p'). \end{aligned} \tag{6.4.8}$$

Equation (6.4.7) also reflects the integrability of the model since it yields the commutativity of the transfer matrix for different spectral parameters

$$[T(p), T(p')] \;=\; 0, \tag{6.4.9}$$

whose proof of (6.4.9) is similar to the one given in Section 6.1.2. Notice that this equation represents an infinite set of conservation laws for the operators t_n:

$$[t_n, t_m] \;=\; 0, \quad \log T(p) \;=\; -\sum_n t_n\, p^n. \tag{6.4.10}$$

The lowest conserved charges can be identified with the momentum and the hamiltonian of the associated quantum system[8]

$$t_0 = iP, \quad t_1 = H.$$

Using the commutativity of the transfer matrices, their maximal eigenvalue can be found, in principle, along the lines discussed for the Ising model in previous sections. Equivalently, the solution of the model can be addressed by the Bethe ansatz approach, as sketched in Problem 4. Here we simply report the final result for the free energy per unit site:

$$\begin{aligned} -F/k_BT &= \log \lambda_{max}(p,\gamma) \\ &= \log \sin(\gamma - p) + \int_{-\infty}^{\infty} \frac{dt}{t} \frac{\sinh[(\pi-\gamma)t]\, \sinh[2pt]}{2\cosh \gamma t\, \sinh \pi t}. \end{aligned} \tag{6.4.11}$$

Let's conclude by outlining the origin of the quantum group symmetry of the model. First, let's write the R-matrix (6.4.3) in terms of the Pauli matrices σ_3 and $\sigma_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$ as

$$R = \begin{pmatrix} a & & & \\ & b & c & \\ & c & b & \\ & & & a \end{pmatrix} = \begin{pmatrix} \sin\left(\frac{1}{2}\gamma - \sigma_3(p - \frac{1}{2}\gamma\right) & & \sigma_- \sin\gamma & \\ & & & \\ & \sigma_+ \sin\gamma & & \sin\left(\frac{1}{2}\gamma + \sigma_3(p - \frac{1}{2}\gamma\right) \end{pmatrix}. \tag{6.4.12}$$

It is easy to see that the Yang–Baxter equation (6.4.2) satisfied by the R-matrix implies the usual $SU(2)$ relations of the Pauli matrix, i.e.

$$[\sigma_3, \sigma_\pm] = \pm 2\,\sigma_\pm, \quad [\sigma_+, \sigma_-] = \sigma_3.$$

Taking the limits of the spectral parameter $p \to \pm i\infty$, let's write the monodromy matrix similarly to eqn (6.4.12)

$$L = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \sin\left(\frac{1}{2}\gamma - J_3(p - \frac{1}{2}\gamma\right) & & J_- \sin\gamma & \\ & & & \\ & J_+ \sin\gamma & & \sin\left(\frac{1}{2}\gamma + J_3(p - \frac{1}{2}\gamma\right) \end{pmatrix}. \tag{6.4.13}$$

From the Yang–Baxter equation (6.4.7) satisfied by the monodromy matrix, we can obtain the commutation relations for the J's:

$$[J_3, J_\pm] = \pm 2 J_\pm, \quad [J_+, J_-] = \frac{\sin(\gamma J_3)}{\sin\gamma}. \tag{6.4.14}$$

These are the commutation relations of the quantum group $SU_q(2)$ that will be discussed in further detail in Section 18.9. Notice that one recovers the usual $SU(2)$ commutation relations when $\gamma \to 0$.

[8] The one-dimensional quantum system associated to the classical two-dimensional six-vertex model is the Heisenberg chain and its continuum limit is described by the Sine–Gordon model.

References and Further Reading

The most important book on the transfer matrix of the two-dimensional system is by Rodney Baxter:

R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, New York, 1982.

Exact solutions of two-dimensional systems can also be obtained by means of the *Bethe ansatz*. A thorough discussion of this method is given in:

B. Sutherland, *Beautiful Models. 70 Years of Exactly Solved Quantum Many-Body Problems*, World Scientific, Singapore, 2004.

M. Gaudin, *La fonction d'onde de Bethe*, Masson, Paris, 1983.

For solutions of the Yang–Baxter equation and the R-matrix see:

L. Faddev, *Integrable models in* (1+1)*-dimensional quantum field theories*, Les Houches, Session XXXIX, 1982, Recent Advances in Field Theory and Statistical Mechanics, Elsevier 1984.

G.E. Andrew, R.J. Baxter and P.J. Forrester, *Eight-vertex SOS and generalized Rogers-Ramanujan-type identities*, J. Stat. Phys. 35 (1984), 193.

V.O. Tarasov, *Irreducible monodromy matrices for the R matrix of the XXZ model and local lattice quantum Hamiltonians*, Theor. Math. Phys. 63 (1985), 440.

V.V. Bazhanov, N.Y. Reshetikin, *Critical RSOS and conformal field theory*, Int. J. Mod. Phys. A 4 (1989), 115.

M. Takahashi, M. Suzuki, *One-dimensional anisotropic Heisenberg model at finite temperatures*, Prog. Theor. Phys. 48 (1972), 2187.

For further studies on the common root of the Yang–Baxter equation in many areas of physics and mathematics (including knot theory), the reader may consult the review:

M. Wadati, T. Deguchi, Y. Akutsu, *Exactly solvable models and knot theory*, Phys. Rep. 180 (1989), 247.

Problems

1. Perron–Frobenius theorem

Consider a finite dimensional positive matrix M, i.e. with all its matrix elements positive, $M_{ij} > 0$. Assume, for simplicity, that M is also a symmetric matrix. Prove that its maximum eigenvalue is positive and non-degenerate. Moreover, prove that

the corresponding eigenvectors have all the components with the same sign (which, therefore, can be chosen to be all positive).

2. Inverse of the matrix V

Consider the operator $\mathbf{R}$ defined in eqn (6.1.29). Using the property $\mathbf{R}^2 = \mathbf{I}$, prove that the inverse of the operator $\mathbf{A} = (2i \sinh 2L)^n \mathbf{I} + (-2i \sinh 2K)^n \mathbf{R}$ is given by

$$A^{-1} = \frac{1}{(2 \sinh 2K)^{2n} - (2 \sinh 2L)^{2n}} \left[(2i \sinh 2L)^n \mathbf{I} - (-2i \sinh 2L)^n \mathbf{R} \right].$$

Use this expression and the functional equation (6.1.30) to determine the inverse of the operator $V(K, L)$.

3. Free energy

Analyze the expression of the free energy of the Ising model, given in eqn (6.3.3), as a function of the parameter h. Show that, with $t = (T - T_c)/T_c = h - 1$, for $t \to 0$ one has

$$F \simeq t^2 \log |t|.$$

4. Bethe ansatz equation

The solution of the six-vertex model consists in finding the eigenvalues of the transfer matrix

$$T(p)\, \psi = (A(p) + D(p))\, \psi = \lambda\, \psi.$$

This problem can be solved by the algebraic Bethe ansatz, whose main steps are as follows. Define the *pseudo-vacuum* ϕ, as the state annihilated by the operator $C(p)$

$$C(p)\, \phi = 0, \qquad \forall p.$$

a Prove that

$$\phi^{\{\beta\}} = \prod_{k=1}^{N} \delta_{\beta_k,+} = \uparrow \cdots \uparrow .$$

b Prove that $\phi^{\{\beta\}}$ is an eigenstate of A and D with eigenvalues

$$A(p)\, \phi = a^N(p)\, \phi$$
$$D(p)\, \phi = b^N(p)\, \phi .$$

However, applying B to ϕ, one gets neither an eigenvector nor zero, $B(p)\phi \neq \phi, 0$. This suggests looking for an eigenstate of the transfer matrix in the form

$$\psi = B(p_1) \ldots B(p_n)\, \phi$$

where the parameters p_i are to be determined.

c Show that, applying $A(p)$ and $D(p)$ to ψ and pushing them through all the B's by the commutation relations (6.4.8), one gets

$$(A(p) + D(p))\psi \;=\; (\lambda_A(p) + \lambda_B(p))\psi + \text{unwanted terms}$$

where

$$\lambda_A(p) \;=\; a^N(p) \prod_{k=1}^{n} \frac{a(p_k - p)}{b(p_k - p)}, \qquad \lambda_B(p) \;=\; b^N(p) \prod_{k=1}^{n} \frac{a(p_k - p)}{b(p_k - p)}.$$

The *unwanted terms*, coming from the second terms in eqn (6.4.8), contain a $B(p)$ and so they can never give a vector proportional to ψ, unless they vanish. Show that this happens if the Bethe ansatz equations hold:

$$\left(\frac{b(p_j)}{a(p_j)}\right)^N \prod_{k=1}^{n} \frac{a(p_j - p_k)}{b(p_j - p_k)} \frac{b(p_k - p_j)}{a(p_k - b_j)} \;=\; -1, \quad j = 1, 2, \ldots n.$$

Notice that the eigenvalue problem of the transfer matrix has been transformed into a set of transcendental equations above for the spectral parameters $p_1, \ldots, p_n$. A further elaboration of the solution of the Bethe ansatz equations leads to the expression (6.4.11) of the free energy of the model.

This page intentionally left blank

Part III

Quantum Field Theory and Conformal Invariance

This page intentionally left blank

7
Quantum Field Theory

Surely you are joking Mr. Feynman!

7.1 Motivations

The statistical models we have analyzed so far are defined on a lattice and they have a microscopic length-scale given by the lattice spacing a. In all these models there is, however, another length-scale provided by the correlation length ξ: this is a function of the coupling constants and can be varied by varying the external parameters of the systems. When the system is sufficiently close to its critical point, the correlation length is much larger than the microscopic scale, $\xi \gg a$. It is then natural to assume that the configurations of the system are sufficiently smooth on many lattice spacings and to adopt a formalism based on continous quantities like a field $\varphi(x)$ (see Fig. 7.1).

As we will show in the sequel of this book, the quantum field theory formulation of statistical models has the important advantage of greatly simplifying the study of critical phenomena: it helps us to select the most important aspects of phase transitions – those related to the symmetries and the dimensionality of the system – and to reach results of great generality. It is worth stressing that the advantage of this method is not only limited to these technical aspects, for the use of quantum field theory in statistical mechanics permits us to achieve a theoretical synthesis of wide scope. Quantum field theory (QFT) was originally developed to describe elementary particles and to reconcile the principles of special relativity with those of quantum mechanics. After the quantization of the electromagnetic field, the subject has witnessed a rapid evolution

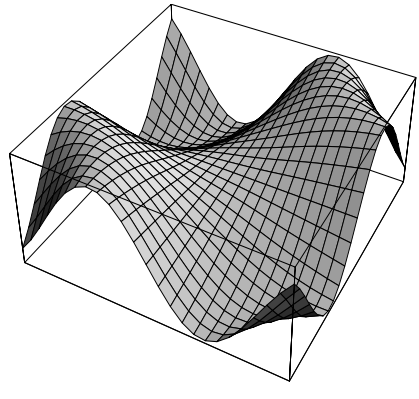


Fig. 7.1 *Continous formulation in terms of a field theory.*

and has been applied to the analysis of weak interactions, responsible for many radioactive decays, and of strong interactions, responsible for the forces of quarks inside hadrons. The degree of refinement reached by this formalism is proved by the incredible precision by which we are able to control nowadays physical effects on a subatomic scale. Moreover, its exceptional theoretical richness has led to extraordinary advances in several fields of physics and mathematics. String theory – a subject developed in recent years in an attempt to unify all fundamental interactions including gravity – can be considered, for instance, as a natural and elegant development of quantum field theory.

The reason why QFT plays a central role both in the context of elementary particles and critical phenomena is due, in a nutshell, to the principle of universality. This is a primary aspect of all local interactions and it is noteworthy that it naturally emerges from the analysis of the renormalization group. Besides, there is a more fundamental reason, for it is possible to show that *any* relativistic quantum theory will look at sufficiently low energy like a quantum field theory.[1] In short, this is the most general theoretical framework to describe a set of excitations above the ground state[2] of a system with infinite degrees of freedom.

Transfer matrix formalism. An obvious question at this point is how can it be possible that a classical statistical system with short-range interactions is equivalent to a relativistic quantum theory. The answer is in the transfer matrix formalism (see Fig. 7.2). Notice that the partition function of a statistical system with short-range interactions can be seen in two equivalent ways: either as a sum over classical variables in a d-dimensional euclidean space with a classical hamiltonian $H(\{s_i\})$, or as the trace of a time evolution operator $T = e^{-\tau\,\mathbf{H}(\{\Phi_i\})}$ associated to a quantum hamiltonian $\mathbf{H}$ in $(d-1)$ dimensions of certain appropriate variables Φ_i. This equivalence is expressed by the identity[3]

$$Z \;=\; \sum_{\{s_i\}} e^{-H(\{s_i\})} \;=\; \mathrm{Tr}_{\Phi_i} \prod_{\tau} e^{-\tau\,\mathbf{H}(\{\Phi_i\})}. \tag{7.1.1}$$

The quantum hamiltonian $\mathbf{H}(\{\mathbf{\Phi_i}\})$ is the first step toward the quantum field theory. Translation and rotation invariance of the quantum theory emerge in fact when the lattice spacing goes to zero. Finally, making a change of the time variable $\tau \to -it$, one arrives at a relativistic theory in $(d-1)$ space dimensions and one time dimension. Vice versa, one can start from a QFT that is relativistically invariant in d space-time dimensions and, with the transformation of the time coordinate $t \to i\tau$, define a euclidean QFT. Once discretized, this theory can be considered for all purposes as a statistical model in d dimensions. In summary, at the root of the equivalence of the formalisms that describe elementary particles and critical phenomena, there is the possibility to adopt either an operatorial or a functional integral approach to a QFT.

[1] See S. Weinberg, *The Quantum Theory of Fields, Vol. I Foundations*, Cambridge University Press, Cambridge, 1995.

[2] The ground state is also called the *vacuum state* of the system.

[3] In the following we will always skip the Planck constant $\hbar$ (considered to be equal to 1) in all formulas.

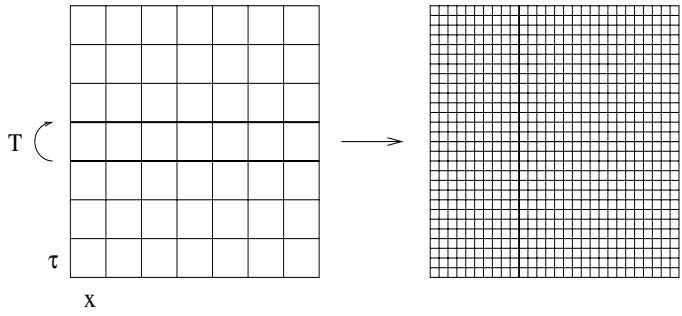


Fig. 7.2 *A classical statistical system in d dimensions and the corresponding quantum system in* $(d-1)$ *dimensions. When the lattice spacing goes to zero one gets a continuous theory both isotropically and translationally invariant.*

This chapter is an introduction to the main concepts of QFT based on the two approaches mentioned above. Since it is impossible to cover in a few pages all aspects of such a large subject, we focus attention only on those aspects that are useful for the comprehension of the following parts of the book and we refer to the references at the end of the chapter for further reading.

7.2 Order Parameters and Lagrangian

Let's start our discussion with the functional formalism of the euclidean QFT that is at the root of the continuous formulation of statistical models. This formalism relies on the possibility to substitute the sum over the classical *discrete* variables $\{s_i\}$ in terms of a functional integral on the *continuous* variables $\varphi(x)$, also classical. This happens near a phase transition point, when the correlation length ξ is much larger than the lattice spacing a:

$$Z = \sum_{\{s_i\}} e^{-H(\{s_i\})} \simeq \int \mathcal{D}\varphi(x)\, e^{-S(\{\varphi\})}, \qquad \xi \gg a. \tag{7.2.1}$$

Let's comment on this expression. The first problem that arises in the functional approach is the identification of the order parameter of the statistical system. As already discussed in Chapter 1, to solve this problem one has to rely on the symmetry of the hamiltonian and on some physical intuition. For instance, in the presence of a Z_2 symmetry, the role of the order parameter can be played by a scalar quantity $\varphi(x)$ that takes values on all of the real axis, odd under the Z_2 transformation, $\varphi(x) \to -\varphi(x)$. For a system that is instead invariant under $O(n)$ symmetry, just to make another example, one can take as order parameter a field with n components $\Phi(x) = [\phi_1(x), \phi_2(x), \ldots, \phi_n(x)]$ that transforms as a vector under the $O(n)$ transformations.

Action and lagrangian. Once the order parameter is identified, one needs next to introduce the Boltzmann weight associated to its different configurations. Only in this way, in fact, can one further proceed to compute statistical averages, correlation functions, and all the other thermodynamic quantities. In analogy with what was done

for the statistical systems defined on a lattice, the probability of the field configuration can be assumed to be proportional to[4]

$$W(\varphi, \{g\}) \,=\, \exp[-\mathcal{S}(\varphi, \{g\}] \,=\, \exp\left[-\int d\mathbf{x}\, \mathcal{L}(x)\right], \tag{7.2.2}$$

where $\mathcal{S}$ is the *action* of the theory, given by an integral on a lagrangian density $\mathcal{L}(x)$. The latter is a local quantity, generically expressed in terms of a polynomial of the fields and their derivatives. To simplify the notation, in the following we focus our attention on a QFT of a scalar field $\varphi(x)$, odd under the Z_2 symmetry. In this case, restricting attention to those terms that are at most of degree 2 in the derivatives,[5] the most general expression of the action is given by

$$\mathcal{S} \,=\, \int d\mathbf{x}\, \left[\frac{1}{2}(\partial_j \varphi)^2 + g_1 \varphi + \frac{g_2}{2}\, \varphi^2(x) + \cdots + \frac{g_n}{n!}\, \varphi^n(x) + \cdots\right]. \tag{7.2.3}$$

In d-dimensional euclidean space, the definition of the derivative term is meant to be a sum over the repeated indices

$$(\partial_j \varphi)^2 \,\equiv\, (\partial_j \varphi)(\partial_j \varphi) \,=\, \sum_{i=1}^{d} \left(\frac{\partial \varphi}{\partial x_i}\right)^2.$$

The lagrangian theory (7.2.3) is also known as the Landau–Ginzburg theory. To cope with the perturbative analysis of such an action, the custom is to isolate firstly its free part, expressed by the quadratic terms

$$\mathcal{S}_0 \,=\, \int d\mathbf{x}\, \mathcal{L}_0 \,=\, \int d\mathbf{x}\, \left[\frac{1}{2}(\partial_j \varphi)^2 + \frac{m^2}{2}\, \varphi^2(x)\right], \tag{7.2.4}$$

and consider the remaining terms in (7.2.3) as the interactive part, denoted by $\mathcal{S}_I = \int d\mathbf{x}\, \mathcal{L}_I$. In the expression above, m is the mass parameter.[6] It is also convenient to introduce the concept of the *manifold of the coupling constants*, defined as the space spanned by the set of all couplings $\{g\} = (g_1, g_3, \ldots, g_n, \ldots)$.

Once the lagrangian is given, the partition function of the system is obtained by summing up all possible configurations of the order parameter

$$Z[\{g\}, a] \,=\, \int \mathcal{D}\varphi\, \exp[-\mathcal{S}[\varphi, \{g\}]]. \tag{7.2.5}$$

In writing this expression we have emphasized that the partition function depends both on the coupling constants g_i and the microscopic cut-off a provided by the lattice spacing of the original theory. Even if we have adopted a continuous formalism to

[4] In the following we will often use the notation $\mathbf{x}$ to denote a vector quantity. Similarly, we will use $d\mathbf{x} = d^d x$.

[5] This can be justified by demanding the causality of the theory.

[6] In the canonical quantization of the theory, m can indeed be identified with the mass of the particle created by the field $\varphi(x)$.

describe a statistical model, it is in fact necessary to take into account the microscopic scales of the systems, and we will see later several effects of such a dependence. Notice that an obvious reason to introduce the microscopic scale a is related to the definition of the measure $\mathcal{D}\varphi$: with this notation we mean a measure on all possible values of the field $\varphi(x)$. Since φ is a continuous quantity defined on *each* point of the space, $\mathcal{D}\varphi$ is not a priori well-defined. In order to make sense of it, one can proceed in two equivalent ways.

The measure. The first approach to define a measure consists of considering the field as a collection of discrete quantities φ_i, defined only on N sites of a lattice with spacing length a, so that $\mathcal{D}\varphi$ can be expressed as a product of the differentials of all these variables, whose number can be enormously large but in any case finite:

$$\mathcal{D}\varphi \,=\, \prod_i^N d\varphi_i. \tag{7.2.6}$$

The second equivalent approach makes use of the translation invariance of the system. This invariance allows us to decompose the field into its Fourier components

$$\varphi(x) \,=\, \frac{1}{\sqrt{N}} \sum_k \varphi(k)\, e^{ikx}.$$

When N is finite, the frequencies are discrete. Furthermore, in the presence of a microscopic scale a, they satisfy the condition

$$|k| \leq \Lambda \simeq \frac{1}{a}.$$

The lattice space a acts then as an ultraviolet cut-off. This turns out to be a very useful quantity, since it permits us also to regularize the divergent terms coming from the perturbative formulation of the theory. In the second approach the measure $\mathcal{D}\varphi$ is also given by the differential of a finite number of variables:

$$\mathcal{D}\varphi \,=\, \prod_{0\leq|k|\leq 1/a} d\varphi(k). \tag{7.2.7}$$

Notice that in both cases the problem to control the behavior of $\mathcal{D}\varphi$ when $N \to \infty$, or, equivalently, $a \to 0$ still remains open. This is a problem not only of the measure but of the entire quantum field theory.

Engineering dimensions. As a matter of fact, the ultraviolet cut-off a also enters other key aspects. Consider, for instance, the engineering dimensions of the coupling constants in the action (7.2.3). To determine such quantities, it is necessary to fix initially the dimension of the scalar field φ. Since $\mathcal{A}$ is a dimensionless quantity, each term of the lagrangian should have dimension a^{-d}. Consider then the kinetic term $(\partial_j\varphi)^2$: imposing the dimension of the field equal to $[\varphi] = a^{x_\varphi}$, we have the condition $a^{-2}\, a^{2\,x_\varphi} = a^{-d}$ and therefore

$$[\varphi] \,=\, a^{1-d/2}. \tag{7.2.8}$$

Once the dimension of $\varphi(x)$ is known, it is easy to obtain the dimensions of the various coupling constants

$$[g_m] = a^{md/2-m-d} \equiv a^{\delta_m}. \tag{7.2.9}$$

It is interesting to observe that each coupling constant has a particular dimension $d_s^{(m)}$ (the so-called upper critical dimension) in which it is dimensionless. For instance g_3 is dimensionless for $d = 6$, g_4 for $d = 4$, and so on. Notice that the quantity δ_m is positive when

$$d \geq d_s^{(m)} = \frac{2m}{m-2}. \tag{7.2.10}$$

Critical behavior. On the basis of the information above, we can already formulate some educated guesses on the critical behavior of the theory – guesses that need however to be refined by further analysis. For a lagrangian with higher coupling constant given by g_n, the corresponding statistical theory is expected to present two different regimes by varying d:

(a) for $d > d_s^{(n)}$, the critical behavior is expected to be described by the mean field theory, with a classical value for the critical exponents;
(b) for $d < d_s$ the system is instead expected to present strong fluctuations with a corresponding significant change of its thermodynamic singularities.

The simplest way to understand these two different critical behaviors is to study the sign of the exponent δ_n: when $\delta_n > 0$ (i.e. $d > d_s^{(n)}$), sending to zero the lattice space a, the corresponding coupling constant becomes smaller, while when $\delta_n < 0$ ($d < d_s^{(n)}$) the coupling constant becomes larger. Consequently, for what concerns the critical behavior, in the first case the microscopic fluctuations are expected to be irrelevant while in the second case to be relevant. Anticipating the results and the terminology of the renormalization group that will be discussed in the next chapter, the coupling constants g_n with $\delta_n > 0$ are called *irrelevant*, those with $\delta_n < 0$ are called *relevant*, and, finally, those with $\delta_n = 0$, *marginal*.

The previous analysis was carried out for a theory invariant under a Z_2 symmetry but the same scenario holds for other theories with different internal symmetry. Namely, each theory has a lower critical dimension d_i, below which there is no longer a phase transition, and an upper critical dimension d_s, beyond which the critical exponents take classical values. The strong fluctuation regime of the order parameters is expected to occur in between, i.e. in the range of dimensions d satisfying

$$d_i \leq d \leq d_s. \tag{7.2.11}$$

For systems with short-range interactions and a discrete symmetry, such as the Ising or the Potts models, the lower critical dimension is always $d_i = 1$, whereas for those with a continuous symmetry, such as the $O(n)$ model, $d_i = 2$. In the range (7.2.11) the critical exponents assume values that are different from their mean field solution and their determination requires more sophisticated theoretical tools.

7.3 Field Theory of the Ising Model

In order to clarify the formulation of a statistical model in terms of a euclidean QFT, it is instructive to study in some detail the case of the Ising model. Consider the partition function of this model, generally expressed as

$$Z \;=\; \sum_{\{s_i\}} \exp\left[\sum_{i,j} J_{ij}\, s_i\, s_j + \sum_i h_i\, s_i\right]. \tag{7.3.1}$$

Let us use an identity valid for the gaussian integral:

$$\int_{-\infty}^{+\infty} \prod_i d\phi_i \, \exp\left[-\frac{1}{4}\sum_{i,j} \phi_i\, J_{ij}^{-1}\, \phi_j + \sum_i \phi_i\, s_i\right] \;=\; A\, \exp\left[\sum_{i,j} J_{ij} s_i\, s_j\right] \tag{7.3.2}$$

(where A is a normalization constant that will be disregarded from now on). This identity allows us to express the partition function (7.3.1) in terms of a lagrangian of a bosonic field ϕ_i, thus swapping from the formulation based on the discrete variables $s_i = (\pm 1)$ to the one based on the continuous variables $\phi_i = (-\infty, +\infty)$. Substituting the identity (7.3.2) in eqn (7.3.1), we have in fact

$$\begin{aligned} Z \;&=\; \sum_{\{s_i\}} \exp\left[\sum_{i,j} J_{ij}\, s_i\, s_j + \sum_i h_i\, s_i\right] \\ &=\; \sum_{\{s_i\}} \int_{-\infty}^{+\infty} \prod_i d\phi_i \, \exp\left[-1/4 \sum_{i,j} \phi_i\, J_{ij}^{-1}\, \phi_j + \sum_i (\phi_i + h_i)\, s_i\right] \\ &=\; \int_{-\infty}^{+\infty} \prod_i d\phi_i \, \exp\left[-\frac{1}{4}\sum_{i,j} (\phi_i - h_i)\, J_{ij}^{-1}\, (\phi_j - h_j)\right] \sum_{\{s_i\}} \exp\left[\sum_i \phi_i\, s_i\right]. \end{aligned} \tag{7.3.3}$$

The sum over the spin configurations in the last term can now be explicitly performed because the spins are decoupled:

$$\sum_{\{s_i\}} \exp\left[\sum_i \phi_i\, s_i\right] \;=\; \prod_i (2\, \cosh\, \phi_i) \;=\; A' \exp\left[\sum_i \log[\cosh\, \phi_i]\right]$$

(where A' is another constant). By means of the linear transformation

$$\phi_i \;\to\; \frac{1}{2} J_{ij}^{-1}\, \phi_j,$$

we arrive (up to multiplicative constants) at the expression

$$\begin{aligned} Z \;&=\; e^{-\frac{1}{4}\sum_{i,j} h_i\, J_{ij}^{-1}\, h_j} \\ &\quad \times \int \mathcal{D}\phi \, \exp\left[-\sum_{i,j} J_{ij} \phi_i\, \phi_j + \sum_i \log[\cosh\,(2 J_{ik}\, \phi_k)]\right]. \end{aligned} \tag{7.3.4}$$

Quadratic part. Notice that the dependence on the magnetic fields is factorized in the prefactor. To understand the nature of the field theory obtained above, it is useful to study its quadratic part. Using the Fourier transform both for the ϕ_i and the coupling constants

$$\phi_i \;=\; \phi(r_i) \;=\; \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \phi(\mathbf{k})\, e^{i\mathbf{k}\cdot\mathbf{r_i}},$$
$$J_{ij} \;=\; J(\mathbf{r_i} - \mathbf{r_j}) \;=\; \frac{1}{N} \sum_{\mathbf{k}} J(\mathbf{k})\, e^{i\mathbf{k}\cdot(\mathbf{r_i}-\mathbf{r_j})},$$

we have

$$\sum_{i,j} J_{ij}\phi_i\,\phi_j \;=\; \sum_{\mathbf{k}} J(\mathbf{k})\,\phi(\mathbf{k})\,\phi(-\mathbf{k}) \;=\; \sum_{\mathbf{k}} J(\mathbf{k})\,|\phi(\mathbf{k})|^2.$$

One should be careful that a quadratic term is also present in the expansion of

$$\log[\cosh x] \;=\; \frac{1}{2}x^2 - \frac{1}{12}x^4 + \cdots$$

explicitly given by

$$2\sum_i (J_{ij}\phi_j)^2 \;=\; 2\sum_{\mathbf{k}} |J(\mathbf{k})|^2\,|\phi(\mathbf{k})|^2.$$

Putting together the two quadratic terms, the free part of the lagrangian reads

$$\int d\mathbf{x}\,\mathcal{L}_0 \;=\; \sum_{\mathbf{k}} \left[J(\mathbf{k}) - 2\,|J(\mathbf{k})|^2\right]\,|\phi(\mathbf{k})|^2. \tag{7.3.5}$$

Let's now expand this expression in powers of k to the second order:[7]

$$J(\mathbf{k}) \;\simeq\; J_0\,(1 - \rho^2\,k^2).$$

If the model has a next neighbor coupling $\tilde{J}$ and the lattice has a coordination number z, we have

$$J_0 \;=\; \sum_r J(r) = \;(z\,\beta\tilde{J})/2, \tag{7.3.6}$$

where $\beta = 1/kT$. The coefficient ρ is of the same order of the lattice spacing a, for it is defined by the average

$$J_0\,\rho^2\,k^2 \;=\; \frac{1}{2}\sum_{\mathbf{r}} J(\mathbf{r})\,(\mathbf{k}\cdot\mathbf{r})^2 \;\simeq\; J_0\,k^2\,a^2.$$

Coming back to eqn (7.3.5), we have

$$\int d\mathbf{x}\,\mathcal{L}_0 \;=\; J_0 \sum_{\mathbf{k}} \left[(1 - 2J_0) + (4J_0 - 1)\,\rho^2\,k^2\right]\,|\phi(\mathbf{k})|^2. \tag{7.3.7}$$

[7] In the inverse Fourier transform, higher orders give rise to higher derivative terms, whose coupling constants are irrelevant.

When the temperature T decreases, J_0 increases and therefore there is a critical value T_c of T for which the term $(1 - 2J_0)$ vanishes:[8]

$$T_c \;=\; z\,\tilde{J}/k, \tag{7.3.8}$$

which coincides with the critical temperature of the mean field solution of the Ising model. At $T = T_c$ the zero mode of the field becomes unstable, because the corresponding integral on this variable in the functional integral (7.3.4) is no longer damped. Hence, T_c signals a phase transition. Imposing

$$\begin{aligned} 1 - 2J_0 &= \frac{T - T_c}{T_c} \\ 4J_0 - 1 &= 1 + \mathcal{O}(T - T_c) \\ J_0 &= \frac{1}{2} + \mathcal{O}(T - T_c) \end{aligned}$$

and substituting in eqn (7.3.7), one has

$$\int d\mathbf{x}\,\mathcal{L}_0 \;=\; \frac{1}{2}\sum_{\mathbf{k}} \left(\frac{T - T_c}{T_c} + \rho^2\,k^2\right) |\phi(\mathbf{k})|^2.$$

Finally, defining

$$\varphi(x) \;=\; \rho\,\phi(x), \quad m^2 \;=\; \frac{1}{\rho^2}\,\frac{T - T_c}{T_c}$$

one arrives at

$$\mathcal{S}_0 \;=\; \int d\mathbf{x}\,\mathcal{L}_0 \;=\; \int d\mathbf{x}\,\frac{1}{2}\left[(\partial_j\varphi)^2 + m^2\,\varphi^2\right] \;=\; \frac{1}{2}\sum_{\mathbf{k}}(k^2 + m^2)|\varphi(\mathbf{k})|^2. \tag{7.3.9}$$

Further interaction terms of the action can be recovered taking into account the higher terms from the expansion of the term $\log[\cosh x]$. They will be discussed later in this chapter.

7.4 Correlation Functions and Propagator

Once the Boltzmann weight of the field configurations is defined, one can proceed to define the correlation functions. They are expressed by the functional integral

$$G^{(n)}(\mathbf{x_1},\ldots,\mathbf{x_n}) \;=\; \langle\varphi(\mathbf{x_1})\ldots\varphi(\mathbf{x_n})\rangle \;=\; \frac{1}{Z}\int \mathcal{D}\varphi\,\varphi(\mathbf{x_1})\ldots\varphi(\mathbf{x_n})\,\exp\left[-\mathcal{S}(\varphi,\{g\})\right]. \tag{7.4.1}$$

[8]Notice that, increasing T, there is another value of the temperature for which the other term $(4J_0 - 1)$ vanishes and then changes sign. This happens because the original matrix J_{ij} is ill-defined since it has negative eigenvalues (all its diagonal terms are zero and correspondingly the sum of its eigenvalues vanishes). Since $s_i^2 = 1$, this drawback can be cured as in the spherical model by adding the identity matrix I to J_{ij} with a proper coefficient in front to ensure the positivity of the eigenvalues. Notice, however, that this operation has the effect of spoiling the simple lattice relation (7.3.6) above.

For a compact expression of these quantities, it is sufficient to couple the field $\varphi(\mathbf{x})$ to an external current $J(\mathbf{x})$, defining a new partition function

$$Z[J] \,=\, \int \mathcal{D}\varphi \, \exp\left[-\mathcal{S}(\varphi,\{g\}) + \int d\mathbf{x}\, J(\mathbf{x})\varphi(\mathbf{x})\right]. \tag{7.4.2}$$

In this way

$$G^{(n)}(\mathbf{x_1},\ldots,\mathbf{x_n}) \,=\, \frac{1}{Z[J]}\,\frac{\delta^n Z[J]}{\delta J(\mathbf{x_1})\ldots\delta J(\mathbf{x_n})}\bigg|_{J=0}. \tag{7.4.3}$$

One can similarly define the correlation functions in momentum space, given by

$$\hat{G}^{(n)}(\mathbf{k_1},\ldots,\mathbf{k_n}) \,=\, \int d\mathbf{x_1}\ldots d\mathbf{x_n}\, e^{-i\mathbf{k_1}\cdot\mathbf{x_1}+\ldots\mathbf{k_n}\cdot\mathbf{x_n}}\, G^{(n)}(\mathbf{x_1},\ldots,\mathbf{x_n}). \tag{7.4.4}$$

Since

$$\int d\mathbf{x}\, J(\mathbf{x})\,\varphi(\mathbf{x}) \,=\, \int \frac{d\mathbf{k}}{(2\pi)^d}\, J(-\mathbf{k})\,\varphi(\mathbf{k}),$$

one has

$$\hat{G}(\mathbf{k_1},\ldots,\mathbf{k_n}) \,=\, (2\pi)^{nd}\,\frac{1}{Z[J]}\,\frac{\delta^n Z}{\delta J(-\mathbf{k_1})\ldots J(-\mathbf{k_n})}. \tag{7.4.5}$$

It is interesting to determine the scale dimensions of the quantities given above: for the correlation functions in real space we have

$$[G^{(n)}(x_1,\ldots,x_n)] \,=\, [\varphi]^n \,=\, a^{n(1-d/2)} \,=\, \Lambda^{n(d/2-1)}, \tag{7.4.6}$$

while for those in momentum space

$$[G^{(n)}(k_i)] \,=\, \Lambda^{-nd}\,[G^{(n)}(x_i)] \,=\, \Lambda^{-n(1/2d+1)}. \tag{7.4.7}$$

For the translation invariance, the Fourier transform (7.4.4) always has a prefactor $\delta^d(\sum_i^n k_i)$. Dividing by this term and denoting by $\bar{G}^{(n)}(k_i)$ the remaining expression, we have

$$[\bar{G}^{(n)}(k_i)] \,=\, \Lambda^{d-n(1/2d+1)}. \tag{7.4.8}$$

The propagator. A special role is played by the two-point correlation function of the free theory

$$G_0^{(2)}(\mathbf{x}_1-\mathbf{x}_2) \,=\, \Delta(\mathbf{x}_1-\mathbf{x}_2) \,=\, \langle\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)\rangle_0. \tag{7.4.9}$$

This is the so-called *propagator* of the theory for reasons that will be immediately clear. Its computation is elementary: expressing the free action as

$$\mathcal{S}_0 \,=\, \int d\mathbf{x}\,\left[\frac{1}{2}(\partial_j\varphi)^2 + \frac{m^2}{2}\varphi^2\right] \,=\, \frac{1}{2}\int d\mathbf{x}\,\varphi(\mathbf{x})\,\left[-\partial^2+m^2\right]\,\varphi(\mathbf{x}), \tag{7.4.10}$$

and computing the gaussian integral in (7.4.2), we arrive at

$$Z_0[J] \,=\, \exp\left[\frac{1}{2}\int d\mathbf{x}\, d\mathbf{y}\, J(\mathbf{x})\,\Delta(\mathbf{x}-\mathbf{y})\,J(\mathbf{y})\right], \tag{7.4.11}$$

where $\Delta(\mathbf{x}-\mathbf{y})$ is, formally, the inverse matrix $(-\partial^2+m^2)$ in coordinate space

$$\Delta(\mathbf{x}-\mathbf{y}) \equiv \left\langle \mathbf{x} \left| \frac{1}{-\partial^2+m^2} \right| \mathbf{y} \right\rangle .$$

A more transparent form is given by its Fourier transform

$$\Delta(\mathbf{x}-\mathbf{y}) = \int_{0\le|k|\le\Lambda} \frac{d\mathbf{k}}{(2\pi)^d} \frac{\exp\left[i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})\right]}{k^2+m^2}, \tag{7.4.12}$$

where $\Lambda = 1/a$ is the ultraviolet cut-off. The euclidean propagator can be computed for any dimension d (and for $\Lambda=\infty$) as follows. Going to radial coordinates and denoting by r and k the modulus of the distance and momentum, we have

$$\Delta(r) = \int \frac{d^dk}{(2\pi)^d} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{k^2+m^2} = \frac{\Omega(d-1)}{(2\pi)^d} \int_0^\infty dk \frac{k^{d-1}}{k^2+m^2} \int_0^\pi d\theta\, \sin^{d-2}\theta\, e^{ikr\cos\theta},$$

where $\Omega(d-1)$ is the solid angle coming from the integration over the $(d-1)$ remaining angles (its explicit expression is given in eqn (2.B.1)). In order to proceed further, we need some integrals involving the Bessel functions

$$\int d\theta \sin^{2\nu}\theta\, e^{ikr\cos\theta} = \frac{\Gamma\left(\nu+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\left(\frac{kr}{2}\right)^\nu} J_\nu(kr),$$

$$\int_0^\infty dk\, k^{\nu+1} \frac{J_\nu(ak)}{k^2+m^2} = m^\nu K_\nu(ma).$$

Using these formulas and simplifying the expressions coming from the Γ functions, the final result is

$$\Delta(r) = (2\pi)^{-d/2} \left(\frac{m}{r}\right)^{d-2} K_{\frac{d-2}{2}}(mr). \tag{7.4.13}$$

Substituting in this formula the relevant values of d (using for $d=1$ and $d=3$ the known expressions for $K_{\pm\frac{1}{2}}(x)$) one easily recovers the results shown in Table 7.1.

Table 7.1: *Propagator, by varying the dimension d, in the limit* $\Lambda \gg m$ *and for* $x=|\mathbf{x}| \gg \Lambda^{-1}$. *In the third column there is the value at the origin when* $\Lambda \gg m$. $K_0(r)$ *and* $K_1(r)$ *are the modified Bessel functions.*

d	$\Delta(x)$ $(\Lambda=\infty)$	$\Delta(0)$ $(\Lambda \gg m)$
1	$\frac{1}{2m}\, e^{-m\,x}$	$\frac{1}{2m}$
2	$\frac{1}{2\pi}\, K_0(m\,x)$	$\frac{1}{2\pi}\, \log\left(\frac{\Lambda}{m}\right)$
3	$\frac{1}{4\pi x}\, e^{-m\,x}$	$\frac{\Lambda}{2\pi^2}$
4	$\frac{m}{2\pi^2 x}\, K_1(m\,x)$	$\frac{\Lambda^2}{16\pi^2}$

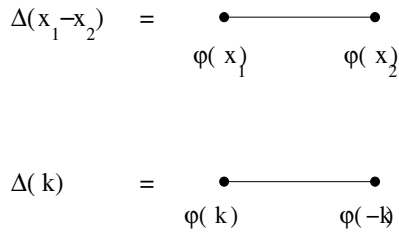


Fig. 7.3 *Propagator of the free theory and its graphical representation.*

Let's comment on other properties of the propagator. It is easy to see that, for any dimension d, $\Delta(x)$ decreases exponentially for $x \to \infty$ as e^{-mx}. Hence, for distance separations of a few units of m^{-1}, the fluctuations of the order parameter are essentially uncorrelated. This means that the correlation length of the system can be identified with the inverse of the mass parameter m

$$\xi = \frac{1}{m}. \tag{7.4.14}$$

When m decreases the correlation length ξ increases and for $m \to 0$ its divergence can be interpreted as the onset of a phase transition.

Notice that the value of $\Delta(x)$ at the origin depends on the dimensionality of the system and on the cut-off. If for $d = 1$ the dependence is rather weak, for $d \geq 2$ there is instead a divergence when $\Lambda \to \infty$. This makes, once more, evident the crucial role played by the ultraviolet cut-off a and by the dimension d of the system.

Finally, since $\Delta(x)$ satisfies the differential equation

$$(-\partial^2_{\mathbf{x}_1} + m^2)\, \Delta(\mathbf{x}_1 - \mathbf{x}_2) = \delta^d(\mathbf{x}_1 - \mathbf{x}_2), \tag{7.4.15}$$

this quantity is also the Green function of the system. From a physical point of view, it describes the propagation of a fluctuation of the field $\varphi(x)$ from position x_1 to x_2. It is convenient to assign to it a graphical representation in terms of a line that connects the two points x_1 and x_2, as shown in Fig. 7.3. An analogous representation is also associated to its Fourier transform

$$\Delta(\mathbf{k}) = \langle \varphi(\mathbf{k})\varphi(-\mathbf{k}) \rangle = \frac{1}{\mathbf{k}^2 + m^2}. \tag{7.4.16}$$

7.5 Perturbation Theory and Feynman Diagrams

In the presence of interactions, it is often impossible to compute exactly the functional integral (7.4.2). For this reason it is important to develop a perturbative formalism based on a power expansion in the coupling constants. It should be stressed that such an approach has some limitations: the most obvious one is that it is restricted to small values of the coupling constants and therefore is unable to catch the strong coupling behavior of the theory. Unfortunately, this is not the only limitation: in most cases, the perturbative series have zero radius of convergence and, at best, they can

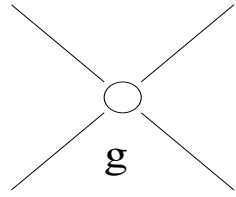


Fig. 7.4 *Vertex of the interaction corresponding to* $\frac{g}{4!}\varphi^4$.

be asymptotic series (see Problem 2). Furthermore, in some quantum field theories there are non-perturbative aspects associated for instance to topological excitations, such as solitons or vortices, that are totally inaccessible to the perturbative approach (see Problem 7). Despite all these drawbacks, it is nevertheless important to study the perturbative formulation since it provides useful information on the analytic nature of the various amplitudes and on the corrections to the free theory behavior.

For the sake of simplicity, we focus our attention on a lagrangian that has only one interaction term, given by φ^4. Isolating the free part, the action can be written as

$$\mathcal{S} \,=\, \mathcal{S}_0 + \frac{g}{4!}\int d\mathbf{x}\,\varphi^4(\mathbf{x}) \,=\, \mathcal{A}_0 + \mathcal{A}_I. \tag{7.5.1}$$

As for the propagator, we can also associate a graphical representation to the interaction term $\frac{g}{4!}\,\varphi^4$: this is given by a vertex with four external lines, as shown in Fig. 7.4.

The perturbative definition of the theory is obtained by expanding the Boltzmann weight in powers of g:

$$e^{-\mathcal{S}_0-\mathcal{S}_I} \,=\, e^{-\mathcal{S}_0}\left[1-\mathcal{S}_I+\frac{1}{2}\mathcal{S}_I^2+\cdots\right].$$

Consider for instance the perturbative definition of the partition function

$$Z[g] \,=\, \int \mathcal{D}\varphi e^{-\mathcal{S}_0}\left[1-\mathcal{S}_I+\frac{1}{2}\mathcal{S}_I^2-\cdots\right]. \tag{7.5.2}$$

Wick's theorem. Order by order in g, all integrals that enter the expression above are of gaussian nature and can be explicitly computed by a generalization of the following gaussian integral in n variables

$$\begin{aligned}\langle x_{k_1}\ldots x_{k_m}\rangle \,&\equiv\, \mathcal{N}\int\prod_i dx_i\, x_{k_1}\ldots x_{k_m}\, e^{-\frac{1}{2}\sum_{i,j}x_i\,A_{ij}\,x_j} \\ &=\, \sum_P A^{-1}_{k_{p_1}k_{p_2}}\,\cdots\,A^{-1}_{p_{k_{m-1}}k_{p_m}}\end{aligned} \tag{7.5.3}$$

where $\mathcal{N}$ is a constant that ensures the correct normalization of the integral, whereas the last sum is over all possible ways of pairing the indices $k_1,\,\ldots\,,k_m$. This expression expresses the content of Wick's theorem in field theory.

Partition function. The partition function (7.5.2) can be written in a compact way as

$$Z[g,J] \,=\, \exp\left[-\frac{g}{4!}\int d\mathbf{x}\frac{\delta^4}{\delta J^4(\mathbf{x})}\right]\,\exp\left[\frac{1}{2}\int d\mathbf{x}\,d\mathbf{y}J(\mathbf{x})\,\Delta(\mathbf{x}-\mathbf{y})\,J(\mathbf{y})\right]. \qquad (7.5.4)$$

an expression that, in the more general case of interaction term $\mathcal{L}_I$, generalizes to

$$Z[\{g\},J] \,=\, \exp\left[-\int d\mathbf{x}\,\mathcal{L}_I\left[\frac{\delta}{\delta J(\mathbf{x})}\right]\right]\,Z_0[J]. \qquad (7.5.5)$$

Let's come back to the analysis of eqn (7.5.4). For the presence of the fourth derivative with respect to the current $J(x)$, the first correction is obtained by expanding $Z_0[J]$ up to second order and then taking the functional derivative with respect to the external currents $J(x_1),\ldots,J(x_4)$ by using the functional relation

$$\frac{\delta^4}{\delta J^4(z)}\,[J(z_1)J(z_2)J(z_3)J(z_4)] \,=\, 4!\,\delta^d(z-z_1)\,\delta^d(z-z_2)\,\delta^d(z-z_3)\,\delta^d(z-z_4).$$

The result is

$$\begin{aligned}\delta Z/Z_0 = -g\,\Bigg(&\frac{1}{8}\int d\mathbf{z}_1\,\Delta(0)+\frac{1}{4}\int d\mathbf{z}_1\,d\mathbf{z}_2\,d\mathbf{z}_3\,\Delta(0)\,\Delta(z_1-z_2)J(z_2)\,\Delta(z_1-z_3)\,J(z_3)\\ &+\frac{1}{4!}\int d\mathbf{z}_1\ldots d\mathbf{z}_5\,\Delta(z_1-z_2)\,J(z_2)\,\Delta(z_1-z_3)\,J(z_3)\\ &\times\,\Delta(z_1-z_4)\,J(z_4)\,\Delta(z_1-z_5)J(z_5)\Bigg).\end{aligned}$$

This expression, as well as all the others relative to higher perturbative orders, can be easily put in graphical form, as shown in Fig. 7.5: in this figure each empty circle (having four external legs) is associated to an integration variable and to the coupling constant g, each line that connects the points x and y is associated to $\Delta(x-y)$ and each black circle relative to the point z_i corresponds to the insertion of a current $J(z_i)$. From Wick's theorem, all currents must be contracted among them: in the first diagram of Fig. 7.5, for instance, this is realized by contracting pairwise the four currents present at the vertex, in the second diagram by contracting two of the currents of the vertex

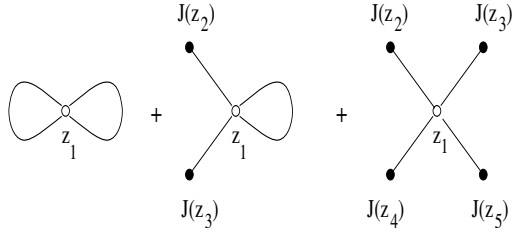


Fig. 7.5 *First perturbative terms of the partition function.*

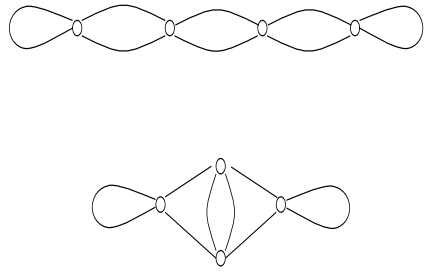


Fig. 7.6 *Two different corrections of order g^4 to the partition function.*

with two external currents, and the remaining ones among themselves and, finally, in the last diagram, by contracting all four currents of the vertex with the external currents. In this procedure, there are certain combinatorial terms that it is necessary to take into account on which we shall comment soon.

Sending to zero all the currents, the only term that survives is the first one. In the Fourier transform, the first diagram in Fig. 7.5 corresponds to

$$\delta Z \;=\; -\frac{g}{8}\,V\left[\int_{0<|k|<1/a}\frac{d^dk}{(2\pi)^d}\,\frac{1}{k^2+m^2}\right]^2, \tag{7.5.6}$$

where V is the volume of the system. Note that, in the absence of the cut-off $1/a$, this gives rise to a divergent correction for $d \geq 2$.

The graphical representation that we used, originally proposed by Feynman, is extremely useful for systematic bookkeeping of the perturbative expansion. The perturbative order is given by the number of vertices of the graph: the graphs in Fig. 7.6, for instance, correspond to two different corrections of order g^4 to the partition function.

Correlation functions. As for the partition function, a similar expansion also exists for the correlation functions. Since their definition is

$$G(\mathbf{x}_1,\ldots,\mathbf{x}_n) \;=\; \frac{1}{Z[J]}\,\frac{\delta^n Z}{\delta J(\mathbf{x}_1)\ldots\delta J(\mathbf{x}_n)}\bigg|_{J=0},$$

in addition to the functional derivatives with respect to the currents of the interaction term, in this case we have also the derivatives with respect to the currents $J(x_1),\ldots,J(x_n)$ coupled to the fields. Note that, at each perturbative order there are also the disconnected graphs coming from the expansion of the partition function of the denominator. For instance, at first order, the two-point correlation function is expressed by the graphs in Fig. 7.7.

Connected correlation functions. In order to eliminate the disconnected graphs coming from the partition function and define instead the connected correlation functions $G_c(x_1,\ldots,x_n)$, it is convenient to introduce the functional $F[J]$ defined as

$$Z[J] \;=\; \exp[F[J]]. \tag{7.5.7}$$

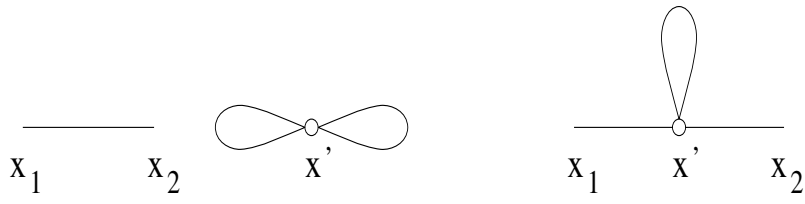


Fig. 7.7 *First-order correction to the two-point correlation function.*

This functional corresponds, up to a sign, to the free energy of the statistical systems and it is easy to show that

$$G_c(\mathbf{x}_1,\ldots,\mathbf{x}_n) \;=\; \left.\frac{\delta^n F[J]}{\delta J(\mathbf{x}_1)\ldots J(\mathbf{x}_n)}\right|_{J=0}. \tag{7.5.8}$$

Using this formula, for the first representatives of the connected correlation functions one has

$$\begin{aligned} G_c^{(1)}(\mathbf{x}) &= \langle\varphi(\mathbf{x})\rangle \\ G_c^{(2)}(\mathbf{x}_1,\mathbf{x}_2) &= \langle\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)\rangle - \langle\varphi(\mathbf{x}_1)\rangle\langle\varphi(\mathbf{x}_2)\rangle \\ G_c^{(3)}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) &= \langle\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)\varphi(\mathbf{x}_3)\rangle - \langle\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_2)\rangle\langle\varphi(\mathbf{x}_3)\rangle \\ &\quad \langle\varphi(\mathbf{x}_2)\varphi(\mathbf{x}_3)\rangle\langle\varphi(\mathbf{x}_1)\rangle - \langle\varphi(\mathbf{x}_1)\varphi(\mathbf{x}_3)\rangle\langle\varphi(\mathbf{x}_2)\rangle \\ &\quad +2\langle\varphi(\mathbf{x}_1)\rangle\langle\varphi(\mathbf{x}_2)\rangle\langle\varphi(\mathbf{x}_3)\rangle. \end{aligned} \tag{7.5.9}$$

Similar expressions are obtained for the connected correlation function in momentum space defined by

$$G_c(\mathbf{k}_1,\ldots,\mathbf{k}_n) \;=\; \left.\frac{\delta^n F[J]}{\delta J(-\mathbf{k}_1)\ldots J(-\mathbf{k}_n)}\right|_{J=0}. \tag{7.5.10}$$

There are of course divergent terms also in the perturbative expansion of the connected correlators. For instance, at first order in g, the correction of the connected two-point function $\langle\varphi(k)\varphi(-k)\rangle$ is given by the diagram on the right side in Fig. 7.7, whose explicit form reads

$$-\frac{g}{2}\,[\Delta(k)]^2\int_{0\le|k|\le\Lambda}\frac{d^dk}{(2\pi)^d}\,\frac{1}{k^2+m^2}.$$

This expression is divergent for $d\ge 2$. More complicated terms, which may also present a nested structure of divergences of the internal loops, appear to higher order. Such a graph (of order g^2) is shown in Fig. 7.8 and its analytic expression apart from the two external propagators $\Delta(k)$ is given by

$$\frac{g_4^2}{6}\int_{0\le|q_i|\le 1/a}\frac{d^dq_1}{(2\pi)^d}\,\frac{d^dq_2}{(2\pi)^d}\,\frac{1}{(q_1^2+g_2)(q_2^2+g_2)[(k-q_1-q_2)^2+g_2]}.$$

Neglecting the possible divergences coming from each loop and focusing, instead, on a simple counting of the power of all the momenta, it is easy to see that, in the absence of the cut-off $1/a$, this integral diverges when $d\ge 3$.

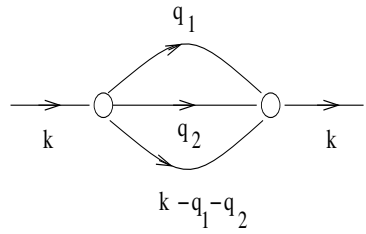


Fig. 7.8 *Correction of order g^2 to the propagator.*

Combinatorial factors. Let's now stop for a moment and discuss briefly the combinatorial factors of the perturbative terms. They come from the number of equivalent ways of contracting the currents. Let's see, for instance, how to arrive at the factor $\frac{1}{8}$ in the first correction to the partition function, eqn (7.5.6): expanding $Z_0[J]$ to the second term

$$Z_0^{(2)} = \frac{1}{2}\left(\frac{1}{2}\right)^2 \int d\mathbf{x}_1 \dots d\mathbf{x}_4 J(\mathbf{x}_1)\,\Delta(\mathbf{x}_1-\mathbf{x}_2)J(\mathbf{x}_2)\,J(\mathbf{x}_3)\Delta(\mathbf{x}_3-\mathbf{x}_4)\,J(\mathbf{x}_4).$$

To this expression we have to apply the differential operator

$$-\frac{g}{4!}\int d\mathbf{x}\left[\frac{\delta}{\delta J(\mathbf{x})}\right]^4.$$

The first derivative of this operator can act on any of the four possible currents $J(\mathbf{x}_1), ..., J(\mathbf{x}_4)$ of the third diagram of Fig. 7.5: hence, there are four possibilities. The second derivative can act on any of the three remaining terms, the third derivative on any of the two remaining terms, and finally, the last one has only one choice. Putting together all these factors, we have: $4 \times 3 \times 2$ equivalent ways of contracting the currents; there is in addition a coefficient $\left(\frac{1}{2}\right)^3$ coming from $Z_0^{(2)}$ and, finally, a factor 1/4! coming from the interaction vertex. Hence the final coefficient is given by

$$4 \times 3 \times 2 \times \left(\frac{1}{2}\right)^3 \times \frac{1}{4!} = \frac{1}{8}. \tag{7.5.11}$$

After this brief excursion in the perturbative series expansion, it is worth formulating more precisely the Feynman rules that enable us to compute the correlation functions either in coordinate or momentum space. Below we state the Feynman rules for the φ^4 theory since their generalization to theories with other vertices φ^r is rather straighforward.

Correlation functions in coordinate space

In order to compute the g^n correction of the connected correlation function $G_c(\mathbf{x}_1, \dots, \mathbf{x}_m)$ one must:

- Draw all connected and distinct graphs with n vertices (of four legs) and m external legs. The latter are identified by the coordinates $\mathbf{x}_1, \dots, \mathbf{x}_m$ of the external points. Connect all the legs by the lines relative to the propagators. Two

diagrams are equivalent if they can be deformed one into the other by moving the vertices of the external points but without cutting any line. For the computation of each diagram, the steps are as follows:

1. Assign to each graph a factor $(-g)^n$.
2. Associate to each vertex a coordinate $\mathbf{z}_i$.
3. Assign to each line that connects the points $\mathbf{x}$ and $\mathbf{y}$ the propagator $\Delta(\mathbf{x}-\mathbf{y})$.
4. Integrate over all coordinates of the internal points.
5. Multiply by $1/2$ for each line whose final points arrive at the same vertex.
6. Multiply by $(p!)^{-1}$ for each set of p lines that link two vertices.
7. Divide by r if the internal points of the graph can be transformed in r ways, leaving the graph invariant.

- Finally, sum over all diagrams.

Correlation functions in momentum space

To compute the g^n corrections to the connected correlation function $G_c(\mathbf{k}_1, \ldots, \mathbf{k}_m)$ the rules are similar:

- Draw all connected and distinct graphs with n vertices (with four legs) and m external legs, the last ones identified by the coordinates $\mathbf{k}_1, \ldots, \mathbf{k}_m$ of the external momenta, and connect all these legs by means of the lines of the propagators. Two diagrams are equivalent if they can be deformed one to the other by moving the vertices or the external points without cutting any lines. To compute explicitly the contribution of the each diagram, the steps are the following:
  1. Assign to each of them a factor $(-g)^n$.
  2. Substitute each line with the expression of the propagator $1/(q_i^2+m^2)$, where $\mathbf{q}_i$ is the momentum of the line, making sure that the conservation of the momenta is enforced at each vertex where several lines meet. Assign the momenta $\mathbf{k}_i$ to the external lines.
  3. Integrate over all the momenta that have not been fixed by the conservation laws, with a factor $(2\pi)^{-d}$ for each integral. The integrals are computed with a cut-off Λ.
  4. Multiply by $1/2$ for each line whose final points arrive at the vertex.
  5. Multiply by $(p!)^{-1}$ for each set of p lines that link two of the same vertices.
  6. Finally, divide by r if the internal points of the diagrams can be transformed in r ways by leaving invariant the graph.
- Finally, sum over all the diagrams.

7.6 Legendre Transformation and Vertex Functions

The expectation value of the field $\varphi(x)$, in the presence of an external current, is given by

$$\langle\varphi(\mathbf{x})\rangle \equiv \bar{\varphi}(\mathbf{x}) = \frac{\delta F[J]}{\delta J(\mathbf{x})}. \tag{7.6.1}$$

We can define the Legendre transform $\Gamma[\bar{\varphi}]$ of the functional $F[J]$ by

$$\Gamma[\bar{\varphi}] + F[J] \;=\; \int d\mathbf{x}\, \bar{\varphi}(\mathbf{x})\, J(x). \tag{7.6.2}$$

Γ is a function of $\bar{\varphi}$ after the elimination of $J(\mathbf{x})$ by means of eqn (7.6.1). It is easy to prove that

$$\frac{\delta\Gamma[\bar{\varphi}]}{\delta\bar{\varphi}(\mathbf{x})} \;=\; J(\mathbf{x}). \tag{7.6.3}$$

This expression can be regarded as the inverse formula of (7.6.1). The formulas above have the following interesting property. If the system is in its ordered phase, one has

$$\bar{\varphi}(\mathbf{x}) \;=\; \left.\frac{\delta F[J]}{\delta J(\mathbf{x})}\right|_{J=0} \;=\; v \neq 0 \tag{7.6.4}$$

and this implies

$$\left.\frac{\delta\Gamma}{\delta\bar{\varphi}(\mathbf{x})}\right|_{\bar{\varphi}(\mathbf{x})=v} \;=\; 0, \tag{7.6.5}$$

i.e. $\Gamma[\bar{\varphi}]$ has a stationary point for a non-vanishing constant value of the field.

Differentiating with respect to $\bar{\varphi}(y)$ the terms of both sides in (7.6.1), one finds

$$\begin{aligned}\delta(\mathbf{x}-\mathbf{y}) \;&=\; \frac{\delta^2 F[J]}{\delta J(\mathbf{x})\delta\bar{\varphi}(\mathbf{y})} \;=\; \int dz\, \frac{\delta^2 F[J]}{\delta J(\mathbf{x})\delta J(\mathbf{z})}\, \frac{\delta J(\mathbf{z})}{\delta\bar{\varphi}(\mathbf{y})} \\ &=\; \int d\mathbf{z}\, \frac{\delta^2 F[J]}{\delta J(\mathbf{x})\delta J(\mathbf{z})}\, \frac{\delta^2\Gamma}{\delta\bar{\varphi}(\mathbf{z})\,\delta\bar{\varphi}(\mathbf{y})}.\end{aligned} \tag{7.6.6}$$

When $J \to 0$, $\frac{\delta^2 F[J]}{\delta J(\mathbf{x})\delta J(\mathbf{z})} = G_c^{(2)}(\mathbf{x},\mathbf{z})$, and therefore the expression

$$\frac{\delta^2\Gamma[\bar{\varphi}]}{\delta\bar{\varphi}(\mathbf{x})\,\delta\bar{\varphi}(\mathbf{y})} \;\equiv\; \Gamma^2(\mathbf{x},\mathbf{y}) \tag{7.6.7}$$

is the "inverse matrix" of the connected two-point function. This statement becomes more evident going to a Fourier transform, because eqn (7.6.6) becomes

$$\Gamma^{(2)}(\mathbf{k})\, G_c^{(2)}(\mathbf{k}) \;=\; 1. \tag{7.6.8}$$

The function $\Gamma^{(2)}$ is also called the *vertex function* of the two-point correlation function: it corresponds to the sum of all Feynman graphs of the two-point correlation function that cannot be separated by removing one line. These graphs are called *one-particle irreducible* (see Fig. 7.9).

There is an interesting graphical representation of (7.6.8): denoting by $X(k)$ the sum of all one-particle irreducible graphs that remain once the external legs are removed, the perturbative expression of the connected two-point function is expressed by the geometrical series

$$G_c^{(2)}(\mathbf{k}) \;=\; \Delta(\mathbf{k}) + \Delta(\mathbf{k})\, X(\mathbf{k})\, \Delta(\mathbf{k}) + \cdots \;=\; \frac{1}{\Delta^{-1}(\mathbf{k}) - X(\mathbf{k})} \;=\; \frac{1}{\Gamma^{(2)}(\mathbf{k})} \tag{7.6.9}$$

whose graphical representation is shown in Fig. 7.10.

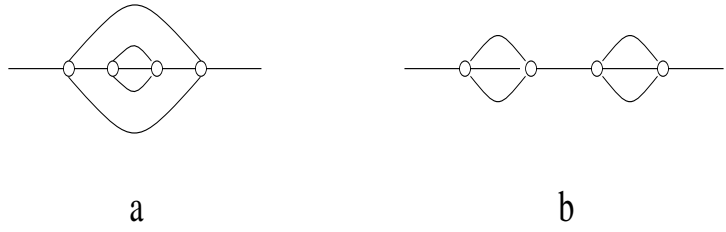


Fig. 7.9 *The graph (a) is one-particle irreducible whereas the graph (b) is not one-particle irreducible: the latter breaks in two by cutting the intermediate line.*

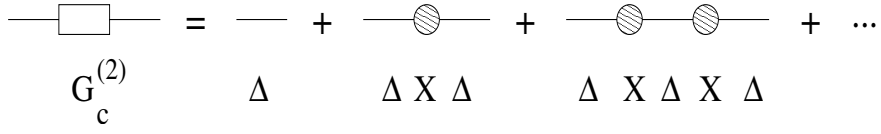


Fig. 7.10 *The connected two-point function is expressed by the infinite series made of the one-particle irreducible diagrams $X(k)$ without the external legs.*

We can proceed in a similar way to define the higher vertex functions, given in coordinate and momentum space by

$$\Gamma^{(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n) \;=\; \left.\frac{\delta^n\Gamma[\bar\varphi]}{\delta\bar\varphi(\mathbf{x}_1)\;\ldots\;\delta\bar\varphi(\mathbf{x}_n)}\right|_{J=0},$$
$$\Gamma^{(n)}(\mathbf{k}_1,\ldots,\mathbf{k}_n) \;=\; \left.\frac{\delta^n\Gamma[\bar\varphi]}{\delta\bar\varphi(-\mathbf{k}_1)\;\ldots\;\delta\bar\varphi(-\mathbf{k}_n)}\right|_{J=0}.$$

For translation invariance, $\Gamma^{(n)}(\mathbf{k}_1,\ldots,\mathbf{k}_n)$ always has a prefactor $\delta^d(\sum_{i=1}^n \mathbf{k}_i)$. The vertex functions where we have factorized this δ function are denoted by $\bar\Gamma^{(n)}(\mathbf{k}_i)$. These functions can be seen as the most fundamental objects of the theory since their knowledge enables to obtain all other correlation functions. They are related to the connected correlation functions $\bar G_c^{(n)}(\mathbf{k}_1,\ldots,\mathbf{k}_n)$ by the relationship

$$\bar G_c^{(n)}(\mathbf{k}_1,\ldots,\mathbf{k}_n) \;=\; -G_c^{(2)}(\mathbf{k}_1)\ldots G_c^{(2)}(\mathbf{k}_n)\,\bar\Gamma^{(n)}(\mathbf{k}_1,\ldots,\mathbf{k}_n) + Q^{(n)}, \qquad (7.6.10)$$

where the $Q^{(n)}$ are the terms coming from the graphs that are reducible by a cut of a line. The graphical representation of the first representatives of the vertex function is shown in Fig. 7.11. The dimension of these quantities is

$$[\Gamma^{(n)}(x_i)] \;=\; [G^{(n)}(x_i)]\,[V]^{-n}\,[G^{(2)}]^{-n} \;=\; \Lambda^{n(d/2+1)}, \qquad (7.6.11)$$

where V is the volume of the system. For their Fourier transform we have

$$[\Gamma^{(n)}(k_i)] \;=\; \Lambda^{-n(d/2-1)}, \qquad (7.6.12)$$
$$[\bar\Gamma^{(n)}(k_i)] \;=\; \Lambda^{n+d-nd/2}.$$

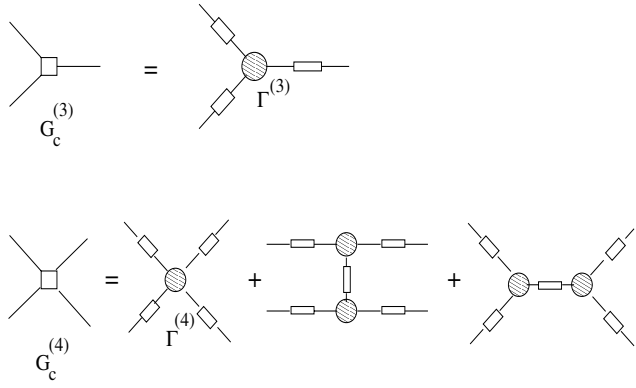


Fig. 7.11 *Vertex functions* $\Gamma^{(3)}$ *and* $\Gamma^{(4)}$*, together with their relation to the corresponding connected correlation functions.*

7.7 Spontaneous Symmetry Breaking and Multicriticality

The vertex functions enter the expansion of the functional $\Gamma[\bar{\varphi}]$

$$\Gamma[\bar{\varphi}] \,=\, \sum_n \frac{1}{n!} \int d\mathbf{x}_1 \,\ldots d\mathbf{x}_n \,\Gamma^{(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\,\bar{\varphi}(\mathbf{x}_1)\ldots\bar{\varphi}(\mathbf{x}_n). \tag{7.7.1}$$

In order to determine whether there is or is not spontaneous symmetry breaking in the system, one needs to check if there is a non-vanishing uniform solution v of the field for eqn (7.6.5). Substituting this value for $\bar{\varphi}$ in (7.7.1), this equation becomes

$$\Gamma[v] \,=\, \sum_n \frac{1}{n!} \left[\int d\mathbf{x}_1 \,\ldots d\mathbf{x}_n \,\Gamma^{(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\right] v^n. \tag{7.7.2}$$

Taking the Fourier transform of $\Gamma^{(n)}(\mathbf{x}_1,\ldots\mathbf{x}_n)$

$$\begin{aligned}\Gamma^{(n)}(\mathbf{k}_1,\ldots\mathbf{k}_n) \,&=\, \int \prod_i^n \frac{d\mathbf{k}_i}{(2\pi)^d}\, e^{i\sum_i \mathbf{k}_i\cdot\mathbf{x}_i}\,\Gamma^{(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\\ &=\, (2\pi)^d\,\delta^d\left(\sum \mathbf{k}_i\right)\,\bar{\Gamma}^{(n)}(\mathbf{k}_1,\ldots,\mathbf{k}_n)\end{aligned}$$

(where the $\delta^d(\sum \mathbf{k}_i)$ comes from the translation invariance) and substituting in (7.7.2) one has

$$\Gamma[v] \,=\, (2\pi)^d\,\delta^d(0)\,\sum_n \frac{1}{n!}\bar{\Gamma}^{(n)}(0,\ldots,0)v^n \,\equiv\, (2\pi)^d\,\delta^d(0)\,U(v). \tag{7.7.3}$$

The term $(2\pi)^d\,\delta^d(0)$ denotes the proportionality of this quantity to the volume of the system. The quantity $U(v)$ is called the *effective potential* of the theory.

At zero order in perturbation theory, the effective potential $U(v)$ exactly coincides with the polynomial terms of the action, with the substitution $\varphi \to v$. Let us analyze the physical scenario that comes from the expression of $U(v)$ at this order, keeping in mind that higher perturbative corrections may alter the values of the coupling constants through the renomalization procedure that will be briefly discussed in the next section.

7.7.1 Universality Class of the Ising Model

Consider the effective potential at zero order for the φ^4 theory:

$$U_0(v) \,=\, h\, v + \frac{m^2}{2} v^2 + \frac{g}{4!} v^4. \tag{7.7.4}$$

Consider the case $h = 0$. For $m^2 > 0$, this function has a unique minimum at the origin: in this case the Z_2 symmetry of the theory is not broken for the expectation value of the field is zero. However, when $m^2 < 0$, the effective potential has two minima at $v_c = \pm\sqrt{-6m^2/g}$. In this case there is an expectation value for the field different from zero (see Fig. 7.12). The choice of any of the two values implies a spontaneous symmetry breaking of the Z_2 invariance. This can be seen explicitly by choosing, for instance, $\bar{\varphi} = v_c$ and making the substitution $\varphi(x) \to \varphi'(x) = \varphi(x) - v_c$ in the lagrangian. Disregarding an additive constant, the lagrangian of the new field is given by

$$\mathcal{L} \,=\, \frac{1}{2}(\partial_j \varphi')^2 - m^2(\varphi')^2(x) + \frac{g\, v_c}{3!}(\varphi')^3(x) + \frac{g}{4!}(\varphi')^4(x). \tag{7.7.5}$$

The mass parameter of the new field is twice the value of the original expression (m^2 is negative in this case) and moreover there is the term $(\varphi')^3$ that is no longer invariant under the transformation $\varphi' \to -\varphi'$. Hence, the fluctuations around any of the two vacua v_c do not respect any longer the original Z_2 symmetry of the model.

One may erroneously think that the Z_2 symmetry of the theory can be re-established by a tunneling effect, as happens in quantum mechanics when a particle is in a double well symmetric potential. However, in field theory, this is impossible because the effective potential given in eqn (7.7.3) contains a term that is proportional to the volume V of the system. In the thermodynamic limit $V \to \infty$, the potential barrier

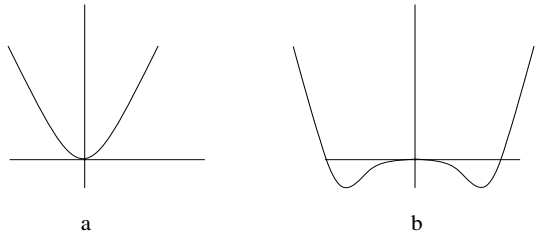


Fig. 7.12 *(a): $m^2 > 0$, there is no spontaneous symmetry breaking; (b) $m^2 < 0$, there is spontaneous symmetry breaking, with an expectation value of the field different from zero.*

between the two minima is therefore infinite and consequently the symmetry cannot be restored by a tunneling effect between the vacua.

Notice that the field theory with an interactive term φ^4 has all the essential features of the class of universality of the Ising model. More specifically: (i) a Z_2 symmetry, under which the order parameter is odd, and (ii) the possibility to have a non-zero vacuum expectation value of the order parameter when the mass term changes its sign. The identification between the two theories become more evident if we make the assumption that the mass parameter depends on the temperature as

$$m^2(T) \simeq (T - T_c). \tag{7.7.6}$$

The upper critical dimension of the φ^4 theory is $d = 4$ and, indeed, beyond this dimension the Ising model has critical exponents that coincide with their mean field values. For $1 < d < 4$, on the contrary, the Ising model has non-trivial values of the critical exponents.

It is important to anticipate that in $d = 2$, in addition to the φ^4 bosonic theory, the Ising model also admits a formulation in term of a fermionic theory. Such a fermionic formulation of the model will be discussed in detail in Chapter 9.

7.7.2 Universality Class of the Tricritical Ising Model

A different class of universality from the Ising model is described by the so-called Blume–Capel model. It involves two statistical variables defined on each site of a lattice:

- a spin variable s_k, with values ± 1;
- a vacancy variable t_k, with values 0 and 1. This variable specifies whether the site is empty (0) or occupied (1).

The more general lattice hamiltonian for these variables (with only next neighbor interactions) is given by

$$\begin{aligned}\mathcal{H} = &-J \sum_{\langle i,j \rangle}^{N} s_i s_j t_i t_j + \Delta \sum_{i=1}^{N} t_i - H \sum_{i=1}^{N} s_i t_i \\ &- H_3 \sum_{\langle i,j \rangle}^{N} (s_i t_i t_j + s_j t_j t_i) - K \sum_{\langle i,j \rangle}^{N} t_i t_j .\end{aligned} \tag{7.7.7}$$

In this expression H is an external magnetic field, H_3 is an additional staggered magnetic field, J is the coupling constant between two next neighbor spins of occupied sites and, finally, Δ is the chemical potential of the vacancies. When $H = H_3 = 0$ the solution of the Blume–Capel model on the lattice shows that there is a tricritical point at (J_c, Δ_c). At a tricritical point a line of first-order phase transition meets the line of a second-order phase transition. Let us see how these physical aspects are captured by a bosonic lagrangian theory with the higher power of interaction given by φ^6. The most general action of this theory is

$$\mathcal{S} = \int d^d x \left[\frac{1}{2} (\partial_j \varphi)^2 + g_1 \varphi + g_2 \varphi^2 + g_3 \varphi^3 + g_4 \varphi^4 + \varphi^6 \right], \tag{7.7.8}$$

where the tricritical point is identified by the conditions $g_1 = g_2 = g_3 = g_4 = 0$. Comparing with the Blume–Capel model, the statistical interpretation of the coupling constants is as follows: g_1 plays the role of an external magnetic field h (the equivalent of H), g_2 measures the displacement of the temperature from its critical value $(T - T_c)$ (the equivalent of $J - J_c$), g_3 plays the role of the staggered magnetic field (the equivalent of H_3) and, finally, g_4 corresponds to $(\Delta - \Delta_c)$.

From the study of the effective potential, it is easy to see that this theory presents a tricritical point. Putting equal to zero all coupling constants of the odd powers of the field, in the remaining even sector we have

$$U_0(\Phi) = g_2\, v^2 + g_4\, v^4 + v^6. \tag{7.7.9}$$

The critical line of the second-order phase transition is identified by the condition of zero mass (i.e. infinite correlation length) – see Fig. 7.13:

$$g_2 = 0, \quad g_4 > 0. \tag{7.7.10}$$

At a line of first-order phase transition there is an abrupt collapse of the vacua. To identify such a line, let's look at the sequence of the potentials (d) and (e) of Fig. 7.14. This sequence shows that, moving with continuity the parameters of the model, the two farthest external vacua become suddenly degenerate with the central one. Hence, the line of the first-order phase transition is characterized by the presence of three degenerate vacua and therefore is identified by the condition

$$g_2 > 0, \quad g_4 = -2\sqrt{g_2}. \tag{7.7.11}$$

In conclusion, the point $g_1 = g_2 = g_3 = g_4 = 0$ is indeed a tricritical point.

By varying the parameters in eqn (7.7.8), the effective potential of this model can take different shapes and consequently its phenomenology can be rather rich. A dimensional analysis shows that the upper critical dimension of the lagrangian theory (7.7.8) is $d = 3$. At this dimension and beyond, the critical exponents take their classical mean field values, while for $1 < d < 3$ they change significantly their values for the strong fluctuations of the order parameters. The exact solution of this model for $d = 2$ will be discussed in detail in Chapter 14.

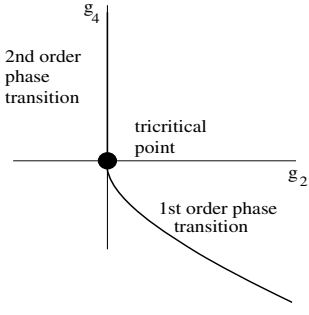


Fig. 7.13 *Phase diagram of the tricritical Ising model in the sub-space of the even coupling constants.*

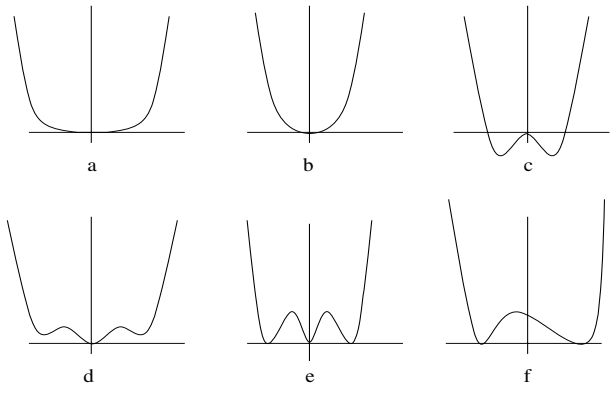


Fig. 7.14 *Some examples of the effective potential of the tricritical Ising model by varying its couplings: (a) critical point; (b) high-temperature phase; (c) low-temperature phase; (d) metastable states; (e) first-order phase transition; (f) asymmetric vacua in the presence of magnetic fields.*

7.7.3 Multicritical Points

Statistical systems that are invariant under a Z_2 symmetry and with multicritical behavior can be described by bosonic field theory with interaction φ^{2n} $(n > 3)$. The criticality of these models is reached by fine tuning $2(n-1)$ parameters: in the lagrangian description this procedure corresponds to putting equal to zero all coupling constants of the powers of the field less than φ^{2n} (except that of φ^{2n-1} that can always be eliminated by a shift of the field φ, as suggested in Problem 5). The detailed description of these classes of universality in $d = 2$ will be presented in Chapter 11.

7.8 Renormalization

In the previous sections we have seen that the perturbative expansion gives rise to expressions that typically diverge when the lattice spacing $\mathbf{a}$ is sent to zero. This is a well-known problem in quantum field theory. Even though its complete analysis goes beyond the scope of this book, we would like nevertheless to draw attention to the main aspects of this topic, using as a guide the Landau–Ginzburg lagrangians. The renormalization of a theory consists of the possibility to eliminate the physical effects coming from the lattice spacing $\mathbf{a}$ – after all, an arbitrary parameter – by an appropriate choice of the coupling constants. For a given dimensionality d of the system, this procedure can be implemented only for certain lagrangians but not for others. To present the main results of this analysis, it is sufficient to focus our attention on the vertex functions $\bar{\Gamma}^{(E)}(k_i)$. It is useful to introduce initially the following concept.

Degree of superficial divergence. The Feynman diagrams that enter the vertex functions $\bar{\Gamma}^{(E)}(k_i)$ are generally expressed by multiple integrals. The degree of superficial divergence D of these expressions is defined as the difference between the number of momenta of the numerator, coming from the differentials $d^d k_i$, and the number of momenta of the denominator, the latter coming from the powers k^2 of the propagators.

Denoting by L the number of integration variables and by I the number of internal lines of the graph, the superficial divergence D is given by

$$D \;=\; L\,d - 2I. \tag{7.8.1}$$

If $D = 0$ the diagram is logarithmically divergent, if $D = 1$ it is linearly divergent, and so on, while if $D < 0$ the diagram is superficially convergent. The reason to distinguish betwen the actual divergent nature of the integral and its superfical divergence comes from the possibility of having nested divergencies: when this happens, the integral can have an actual divergence that is different from the one indicated by its index D. An example is provided by the last diagram in Fig. 7.15: for $d = 4$ this diagram has a degree of superficial divergence $D = -2$ but it actually has an internal loop that is logarithmically divergent.

The key point to introduce such a concept is that the superficial divergence D of an amplitude can be fixed only by using considerations of graph theory. Let us denote by E the number of external lines and by n_r the number of vertices corresponding to the interaction φ^r. There is an elementary relationship between these two quantities: since a vertex of type r has r lines that start from it and each external line has only one ending point, we have

$$E + 2I \;=\; \sum_r r\, n_r,$$

namely

$$I \;=\; \frac{1}{2}\Big(\sum_r r\, n_r - E\Big). \tag{7.8.2}$$

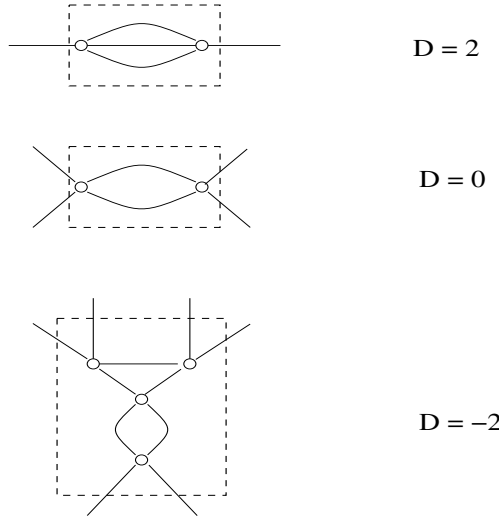


Fig. 7.15 *Degree of superficial divergence of some graphs of the vertex functions (in the dashed box) with $E = 2$, $E = 4$ and $E = 6$, for the φ^4 theory in $d = 4$.*

The number L of integrals coincides with the number of loops of the graph. In turn, this is equal to the number of internal lines I minus the number of conservation laws of the momenta. Each interaction carries a δ function but we must be careful in considering the one that corresponds to the prefactor $\delta(\sum_j^E k_j)$ associated to the total conservation law of the momenta of the E external lines. Hence

$$L = I - (n_r - 1). \tag{7.8.3}$$

Substituting this expression and eqn (7.8.2) in (7.8.1) we have

$$\begin{aligned} D &= \left(d + E - \frac{1}{2}Ed\right) + \sum_r n_r \left(\frac{1}{2}rd - d - r\right) \\ &= \left(d + E - \frac{1}{2}Ed\right) + \sum_r n_r \delta_r, \end{aligned} \tag{7.8.4}$$

where the exponent δ_r is the one defined in eqn (7.2.9). In conclusion, the degree of superficial divergence of an amplitude is given by the sum of two terms: the first is independent of the perturbative order while the second, on the contrary, depends on the type of interaction and on the perturbative order. It is worth noting that the origin of the two terms in (7.8.4) can be traced back by a dimensional analysis: the first term, in fact, simply expresses the dimensionality of the vertex function $\bar{\Gamma}^{(E)}(k_i)$ while the second term takes into proper account the dimensionality of the coupling constants and the perturbative order in which they are involved.

Renormalizable lagrangian. Fixing the dimensionality d of the system, if we require that *independently of the perturbative order* only a finite number of vertex functions is divergent, the coupling constant has to be dimensionless, i.e. $\delta_r = 0$. This condition determines which of the lagrangians is renormalizable in d dimensions: this lagrangian corresponds to a Landau–Ginzburg one with the highest interaction power φ^r equal to

$$r = \frac{2d}{d-2}. \tag{7.8.5}$$

Vice versa, if we start with a lagrangian with φ^r as its highest interaction term, there is a critical dimension, identified by the upper critical dimension d_s given in eqn (7.2.10), in which this lagrangian is renormalizable. Obviously the presence of terms with $\delta_r < 0$ can only decrease the superficial divergence of the amplitudes. For this reason we can focus our attention only on the case in which $\delta_r = 0$.

Consider, for instance, the lagrangian theory

$$\mathcal{L} = \frac{1}{2}(\partial_j \varphi)^2 + \frac{m^2}{2}\varphi^2 + \frac{g_4}{4!}\varphi^4. \tag{7.8.6}$$

Such a theory has $d_s = 4$. If we choose the dimension d of the system exactly equal to d_s, its divergent amplitudes (with $D \geq 0$) correspond to diagrams with external lines $E \leq 4$, as can be seen by eqn (7.8.4). Since the amplitudes with an odd number of external legs vanish[9] for the symmetry $\varphi \to -\varphi$, it remains to consider only those

[9]This is certainly true in the symmetric phase of the theory. In the broken symmetry phase of the model the argument has to be modified accordingly but it still remains true that the model is renormalizable.

with $E = 2$ and $E = 4$. Note that the divergent vertex functions are those coming from the terms φ^2 and φ^4 already present in the lagrangian! Such a theory is therefore renormalizable since it is possible to cure all the divergences of the vertex functions $\bar{\Gamma}^{(2)}$ and $\Gamma^{(4)}$ by adjusting a set of counterterms that have exactly the same form of the original lagrangian

$$\mathcal{L} \to \mathcal{L} + \frac{A}{2}(\partial_j \varphi)^2 + \frac{B}{2}\varphi^2 + \frac{C}{4!}\varphi^4. \tag{7.8.7}$$

Bare quantities. The coefficients A, B, C are (divergent) functions of the cut-off a, chosen in such a way to cancel order by order the divergences of the perturbative series. Observe that, defining

$$\begin{aligned} \varphi_0 &= (1+B)^{1/2}\,\varphi, \\ m_0^2 &= (m^2 + A)(1+B)^{-1}, \\ g_0 &= (g_4 + C)(1+B)^{-2}, \end{aligned} \tag{7.8.8}$$

the modified lagrangian (7.8.7) can be written as

$$\mathcal{L} = \frac{1}{2}(\partial\varphi_0)^2 + \frac{m_0^2}{2}\varphi_0^2 + \frac{g_0^4}{4!}\varphi_0^4, \tag{7.8.9}$$

which is similar to the initial one. However, this transformation changes radically the meaning of the parameters. All quantities, including the field itself, depend now on the cut-off and are non-universal. For these reasons they are called *bare quantities*. They only serve to remove the infinities. In order to link the bare quantities to the physical parameters of the theory, such as the physical value of the mass or the coupling constant, it is necessary to determine (say, experimentally) the latter quantities at a given value of the momenta of the vertex functions (for instance, at zero momenta) and then use eqn (7.8.8) for inverting these relations. It is only after are know the experimental values m^2_{exp} and λ_{exp} that the theory acquires its predictive power, since it is only then that the formalism is able to determine uniquely all other amplitudes. These quantities become finite functions of m^2_{sp} and λ_{sp} and, of course, of the external momenta.

From what was said above, it should be clear that not all the lagrangians are renormalizable. For instance, adding an interaction term φ^5 to the φ^4 theory in $d = 4$, with $\delta_5 = 1$, this term produces an infinite sequence of divergent vertex functions. The perturbative cure of these terms relentlessly leads to the addition of counterterms with arbitrary powers of φ^n in the lagrangian, i.e. we arrive at a theory with an infinite number of parameters. In this case we lose any predictive power of the theory defined in the limit $a \to 0$.

Effective theories. On the other hand, it should be said that if there are reasons to consider the lattice spacing as a *finite physical quantity* that plays an important role in the problem under consideration, a priori there is no reason to exclude non-renormalizable lagrangians. This is, in particular, the modern view about the renormalization problem in quantum field theory and it can be perfectly justified by the

renormalization group approach. In conclusion, the final meaning of quantum field theories is that of *effective theories*, i.e. theories that present a dependence on the length-scale L or, equivalently, on the energy-scale E at which we are analyzing the physical systems. From this point of view, the important point is the possibility to control how the physical properties vary by varying the length or the energy scales. As we will see in the next chapter, in the (infinite-dimensional) manifold of the couplings a change of these scales has the effect of inducing a motion of the point that represents the system. The properties of this motion will be the object of the renormalization group analysis.

7.9 Field Theory in Minkowski Space

Quantum field theories describe the excitations of a physical system. These excitations share the same properties of the elementary particles: they can be created at a given point of the system and annihilated at another, or they can propagate for a given time interval causing scattering processes in the meantime. In the next two sections we highlight these aspects closely related to elementary particles. For doing so, it is necessary first to define the quantum field theory in Minkowski space and, secondly, to adopt an operatorial formalism. We choose to illustrate these features using the Landau–Ginzburg lagrangians as an example, in particular the φ^4 theory. Let's start our discussion from the measure with which we have weighted the configurations of the field φ in the d-dimensional euclidean space

$$W(\{\varphi\}) \,=\, \exp[-\mathcal{S}] \,=\, \exp\left[-\int d^d x\, \mathcal{L}(x)\right], \tag{7.9.1}$$

with

$$\begin{aligned} \mathcal{L} \,&=\, \frac{1}{2}(\partial_j \varphi)^2 + U(\varphi), \\ U(\varphi) \,&=\, \frac{m^2}{2}\varphi^2 + \frac{g}{4!}\varphi^4. \end{aligned} \tag{7.9.2}$$

Let us now select one of the d coordinates, say $x_0 = \tau$, and promote it to the role of a euclidean time variable. Finally, let's make the transformation $\tau \to -it$. As discussed below, this innocent transformation changes completely the meaning of the theory.

Making the same transformation $\tau \to -it$ in the derivative term $(\partial_j \varphi)^2$ of the lagrangian, we get a new expression of $W(\{\varphi\})$, that we denote by $\tilde{W}(\{\varphi\})$

$$\tilde{W}(\{\varphi\}) \,=\, \exp[\, i\, \tilde{\mathcal{S}}\,] \,\equiv\, \exp\left[i \int d^{d-1} x\, dt\, \tilde{\mathcal{L}}\right], \tag{7.9.3}$$

where

$$\tilde{\mathcal{L}} \,=\, \frac{1}{2}\left[\left(\frac{\partial \varphi}{\partial t}\right)^2 - (\nabla \varphi)^2\right] - U(\varphi). \tag{7.9.4}$$

Comparing $\tilde{\mathcal{L}}$ with the quantity in (7.9.2) we note two differences: the first is that there is a relative sign between the derivatives concerning the spatial coordinates and the

one relative to the time variable; the second is that all polynomial terms have changed sign. However, the most important effect is in the quantity $\tilde{W}$, which is now a complex quantity. Hence this quantity has lost the original meaning of probability, acquiring instead the meaning of *amplitude*, in the usual meaning of quantum mechanics. To clarify this point, we will briefly recall the quantization of a particle that moves in an n-dimensional space.

Quantum mechanics of a particle. Let

$$L(q) \,=\, \frac{1}{2}\sum_{i=1}^{n}(\dot{q}_i)^2 - V(q), \tag{7.9.5}$$

be the lagrangian of a particle ($\dot{q}_i = dq_i/dt$),

$$A \,=\, \int_0^t dt\, L(q), \tag{7.9.6}$$

its action, and H the hamiltonian, defined by the Legrendre transformation

$$H(q,p) \,=\, \sum_{i=1}^{n} p_i\, q_i - L \,=\, \sum_{i=1}^{n} \frac{p_i^2}{2} + V(q). \tag{7.9.7}$$

The components of the momentum

$$p_i \,=\, \frac{\delta L}{\delta \dot{q}_i} \,=\, \dot{q}_i,$$

together with the coordinates q_i, are now operators that satisfy the commutation relations

$$[q_k, p_l] \,=\, i\,\hbar\,\delta_{k,l}, \qquad [q_k, q_l] \,=\, 0, \qquad [p_k, p_l] \,=\, 0. \tag{7.9.8}$$

Denoting by E_n the eigenvalues of the hamiltonian and $|E_n\,\rangle$ its eigenvectors, the amplitude that such a particle moves in a time interval t from the point q_0 (where it is localized at the time $t = 0$) to the point q_f, is given by the time evolution of the unitary operator $e^{-itH/\hbar}$

$$\langle q_f, t\,|\,q_0, 0\,\rangle \,=\, \langle q_f\,|\,e^{-itH/\hbar}\,|\,q_0\,\rangle \,=\, \sum_{n=0}^{\infty}\langle q_f\,|\,E_n\,\rangle\langle\,E_n\,|\,q_0\,\rangle\, e^{-it\,E_n/\hbar}, \tag{7.9.9}$$

where we have used the completeness relation

$$\sum_{i=1}^{\infty} |\,E_n\,\rangle\,\langle\,E_n\,| \,=\, 1.$$

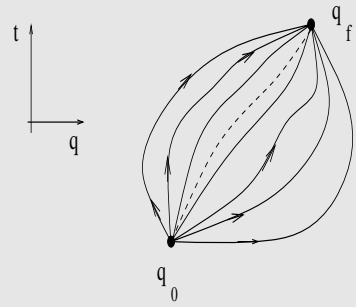


Fig. 7.16 *The Feynman integral, namely a sum over the classical trajectories that link the initial and the final points, each trajectory weighted by* e^{iA} *where* A *is the action of each trajectory. The dashed line corresponds to the classical trajectory, a solution of the classical equation of motion.*

However this is not the only way to compute such an amplitude: as shown by Feynman (see Appendix 7A), it can also be obtained by means of a path integral over all the classical trajectories that connect the points $(q_0, 0)$ and (q_f, t) (see Fig. 7.16). In this approach each path is weighted by $\exp(iA/\hbar)$, namely[10]

$$\langle q_f, t \,|\, q_0, 0 \,\rangle \;=\; \int_{\substack{q(0)\,=\,q_0 \\ q(t)\,=\,q_f}} \mathcal{D}q \, \exp(iA/\hbar). \tag{7.9.10}$$

In the semiclassical limit $\hbar \to 0$, the integral can be estimated by the saddle point method: the most important contribution comes from the trajectory for which the action is stationary, $\delta A = 0$, i.e. the trajectory that satisfies the classical equation of motion

$$\frac{d}{dt}\left(\frac{\delta L}{\delta \dot{q}_i}\right) - \frac{\delta L}{\delta q_i} \;=\; 0. \tag{7.9.11}$$

As shown in Appendix 7A, by means of the path integral we can also compute the time-ordered correlation function of the operators

$$\langle q_f t | T[Q(t_1) \dots Q(t_k)] \, | q_0, 0 \rangle \;=\; \int_{\substack{q(0)\,=\,q_0 \\ q(t)\,=\,q_f}} \mathcal{D}q \, q(t_1) \dots q_k(t) \, \exp(iA/\hbar), \tag{7.9.12}$$

with $t_1 > t_2 > \ldots > t_k$.

Coming back to the field theory, and in particular to eqn (7.9.3), we see then that $\tilde{W}(\{\varphi\})$ can be interpreted as the weight of a *classical configuration* of the field $\varphi(x, t)$

[10] Also in this case, to define the measure $\mathcal{D}q$ it is necessary to make the variable q discrete on the slices $t_k = k\epsilon$ $(k = 0, 1, \ldots, N)$ of the time interval t, with $\epsilon = t/N$, so that $\mathcal{D}q = \prod_{k=0}^{N} dq_k/\sqrt{2\pi\epsilon}$.

in the computation of a *quantum amplitude* (we have imposed $\hbar = 1$)

$$\langle \varphi_f(x,t) \,|\, \varphi_0(x,0) \,\rangle \;=\; \int_{\substack{\varphi(x,t)\,=\,\varphi_f(x)\\ \varphi(x,0)\,=\,\varphi_0(x)}} \mathcal{D}\varphi \, \exp\left[i\, \tilde{\mathcal{S}} \right]. \tag{7.9.13}$$

With this interpretation of $\tilde{W}(\{\varphi\})$, we can now proceed as in the quantum mechanics of a particle but, this time, back-to-front: instead of using the path integral, we will adopt the operatorial approach to describe the dynamics associated to the lagrangian (7.9.4). In QFT the role of the operators $q_i(t)$ is played by the field $\varphi(x,t)$, regarded as an operator that acts at each point (x,t) of space-time. The operator formalism that we have just defined is relativistically invariant, as discussed in Appendix 7B. In this appendix one can also find the relevant definitions used in the following.

The field $\varphi(x,t)$ satisfies the operator differential equation coming from the Euler–Lagrange equation of motion

$$\partial_\mu \left(\frac{\partial \tilde{\mathcal{L}}}{\partial(\partial_\mu \varphi)} \right) - \frac{\partial \tilde{\mathcal{L}}}{\partial \varphi} \;=\; 0 \tag{7.9.14}$$

which, for the φ^4 theory, reads

$$\left(\Box + m^2 \right) \, \varphi(x,t) \;=\; -\frac{g}{3!} \varphi^3(x,t), \tag{7.9.15}$$

where

$$\Box \;=\; \frac{\partial^2}{\partial t^2} - \nabla^2.$$

The conjugate momentum is defined by

$$\pi(x,t) \;=\; \frac{\delta \tilde{\mathcal{L}}}{\delta \dot{\varphi}(x,t)} \;=\; \frac{\partial \varphi}{\partial t}. \tag{7.9.16}$$

As $\varphi(x,t)$, also $\pi(x,t)$ is an operator. In analogy with quantum mechanics, we postulate that these operators satisfy the equal-time commutation relation

$$\begin{aligned} [\varphi(x,t), \pi(y,t)] \;&=\; i\, \delta^d(x-y), \\ [\varphi(x,t), \varphi(y,t)] \;&=\; 0, \\ [\pi(x,t), \pi(y,t)] \;&=\; 0. \end{aligned} \tag{7.9.17}$$

In terms of $\pi(x)$ we can define the hamiltonian density by the Legendre transform

$$\mathcal{H}(x,t) \;=\; \pi(x,t)\, \dot{\varphi}(x,t) - \tilde{\mathcal{L}} \;=\; \frac{1}{2} \left[\left(\frac{\partial \varphi}{\partial t} \right)^2 + (\nabla \varphi)^2 \right] + U(\varphi). \tag{7.9.18}$$

The hamiltonian and the momentum are given by

$$\begin{aligned} H \;&=\; \int d^{d-1}x \, \mathcal{H}(x,t), \\ \mathbf{P} \;&=\; -\int d^{d-1}x \, \pi(x,t)\, \nabla \, \varphi(x,t). \end{aligned} \tag{7.9.19}$$

As a consequence of the equation of motion (7.9.15), both are conserved quantities

$$\frac{dH}{dt} \;=\; \frac{d\mathbf{P}}{dt} \;=\; 0 \tag{7.9.20}$$

and can be expressed in terms of the *stress–energy tensor* $T^{\mu\nu}(x)$. This quantity is defined by (see Appendix 7C)

$$T^{\mu\nu}(x) \;=\; \frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_\mu\varphi)}\,\partial^\nu\varphi - g^{\mu\nu}\,\tilde{\mathcal{L}}. \tag{7.9.21}$$

$T^{\mu\nu}$ is a conserved quantity: it satisfies

$$\partial_\mu\, T^{\mu\nu}(x) \;=\; 0, \tag{7.9.22}$$

and therefore

$$H \;=\; \int T^{00}(x)\, d^{d-1}x, \qquad P^{(i)} \;=\; \int T^{0i}(x)\, d^{d-1}x. \tag{7.9.23}$$

7.10 Particles

To understand the nature of the excitations $\varphi(x,t)$ it is sufficient to consider the free theory ($g = 0$). In this case the operatorial equation satisfied by the field is

$$\left(\Box + m^2\right)\,\varphi(x,t) \;=\; 0. \tag{7.10.1}$$

A plane wave $e^{i\,(\mathbf{k}\dot{\mathbf{x}} - Et)}$ is a solution of this equation if

$$E^2 \;=\; \mathbf{k}^2 + m^2. \tag{7.10.2}$$

One easily recognizes that this is the dispersion relation of a relativistic particle. Taking into account the two roots of this equation, the most general solution of (7.10.1) is given by a linear superposition of plane waves

$$\varphi(x,t) \;=\; \int d\Omega_k\,\left[A_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x} - iE_k t} \;+\; A^\dagger_{\mathbf{k}}\, e^{-i\mathbf{k}\cdot\mathbf{x} + iE_k t}\right]. \tag{7.10.3}$$

In this expression and in the next ones that follow, $E_k \;=\; \sqrt{\mathbf{k}^2 + m^2}$. The coefficients $A_{\mathbf{k}}$ and $A^\dagger_{\mathbf{k}}$ are a set of operators, called *annihilation and creation operators*, respectively. In writing this solution we have adopted a relativistically invariant differential measure

$$d\Omega_k \;\equiv\; \frac{d^{d-1}k}{(2\pi)^{d-1}2E_k}.$$

From the quadratic nature of the relativistic dispersion relation, there are both positive and negative frequencies in the mode expansion of the field. The negative frequencies can be interpreted as the propagation, back in time, of an antiparticle, a statement

that becomes evident if one considers a complex scalar field[11] (see Problem 8). Using (7.9.16), we obtain the conjugate momentum

$$\pi(x,t) \;=\; -i \int d\Omega_k \, E_k \left[A_{\mathbf{k}} \, e^{i\mathbf{k}\cdot\mathbf{x} - iE_k t} - \; A^\dagger_{\mathbf{k}} \, e^{-i\mathbf{k}\cdot\mathbf{x} + iE_k t} \right]. \tag{7.10.4}$$

The commutation relations of $A_{\mathbf{k}}$ and $A^\dagger_{\mathbf{k}}$ can be recovered by imposing the validity of eqn (7.9.17)

$$\begin{aligned} [A_{\mathbf{k}}, A^\dagger_{\mathbf{p}}] \;&=\; (2\pi)^{d-1} \, 2E_k \, \delta^{d-1}(\mathbf{k}-\mathbf{p}), \\ [A_{\mathbf{k}}, A_{\mathbf{p}}] \;&=\; [A^\dagger_{\mathbf{k}}, A^\dagger_{\mathbf{p}}] \;=\; 0. \end{aligned} \tag{7.10.5}$$

Besides the relativistic normalization of these operators, they are the exact analogs of the annihilation and creation operators of the harmonic oscillator.

Substituting the expressions for $\varphi(x,t)$ and $\pi(x,t)$ in H we have

$$H \;=\; \frac{1}{2} \int d\Omega_k \, \left(A^\dagger_{\mathbf{k}} \, A_{\mathbf{k}} + A_{\mathbf{k}} \, A^\dagger_{\mathbf{k}} \right) E_k \;=\; \int d\Omega_k \, \left(A^\dagger_{\mathbf{k}} \, A_{\mathbf{k}} + \frac{1}{2} \right) E_k, \tag{7.10.6}$$

where we have used the commutation relation (7.10.5). The term

$$E_0 \;=\; \frac{1}{2} \int d\Omega_k \, E_k$$

is infinite and corresponds to the *vacuum energy*. Since this quantity is a constant, it can be safely subtracted, so that the new definition of the hamiltonian is

$$H \;=\; \int d\Omega_k \, A^\dagger_{\mathbf{k}} \, A_{\mathbf{k}} \, E_k. \tag{7.10.7}$$

This redefinition employs the *normal product* of the operators: a product of operators is normally ordered if all the creation operators are on the left side of the annihilation operators. Denoting the normal order by : :, the new hamiltonian can be expressed as

$$H \;=\; \frac{1}{2} \int d^d x \, : \left(\pi^2 + (\nabla\varphi)^2 + m^2 \varphi^2 \right) : \tag{7.10.8}$$

Note that the annihilation and creation operators are associated to plane waves with positive and negative time frequency, respectively. Indicating by $\varphi^{(+)}(x)$ and $\varphi^{(-)}(x)$ these two terms in the decomposition of $\varphi(x)$

$$\varphi(x) \;=\; \varphi^{(+)}(x) + \varphi^{(-)}(x),$$

one has, for instance

$$: \varphi(x) \, \varphi(y) \; := \; \varphi^{(+)}(x) \, \varphi^{(+)}(y) + \varphi^{(-)}(x) \, \varphi^{(+)}(y) + \varphi^{(-)}(x) \, \varphi^{(-)}(y) + \varphi^{(-)}(y) \, \varphi^{(+)}(x).$$

Substituting the expressions for $\varphi(x)$ and $\pi(x)$ in the momentum operator, we have

$$P \;=\; \int d\Omega_k \left(A^\dagger_{\mathbf{k}} \, A_{\mathbf{k}} + \frac{1}{2} \right) \mathbf{k} \;=\; \int d\Omega_k \, A^\dagger_{\mathbf{k}} \, A_{\mathbf{k}} \, \mathbf{k}. \tag{7.10.9}$$

Notice that in this case the zero point of the momentum is absent since, in the integration, it is cancelled by the equal and opposite contributions coming from $\pm\mathbf{k}$.

[11] For a real scalar field, a particle coincides with its antiparticle.

In the expression for both the energy and the momentum there is the operator

$$N_{\mathbf{k}} \,=\, A^{\dagger}_{\mathbf{k}}\, A_{\mathbf{k}}. \tag{7.10.10}$$

This is the key observation that supports an interpretation of quantum field theory in terms of particles. In the following we prove that the operators $N_{\mathbf{k}}$ are simultaneously diagonalizable and their eigenvalues are the integer numbers

$$n_{\mathbf{k}} \,=\, 0, 1, 2, \ldots \tag{7.10.11}$$

In this way, the energy and the momentum associated to the field φ can be written as

$$E \,=\, \int d\Omega_k \, n_{\mathbf{k}}\, E_k, \qquad \mathbf{P} \,=\, \int d\Omega_k \, n_{\mathbf{k}}\, \mathbf{k}. \tag{7.10.12}$$

From this expression it is clear that these quantities coincide with the energy and momentum of a set of scalar particles of mass m, with a relativistic dispersion relation. This set contains $n_{\mathbf{k}_1}$ particles of momentum $\mathbf{k}_1$, $n_{\mathbf{k}_2}$ particles of momentum $\mathbf{k}_2$, etc.

The statement that all $N_{\mathbf{k}}$ commute is a simple consequence of the commutation relation (7.9.17)

$$\begin{aligned} [N_{\mathbf{k}_1}, N_{\mathbf{k}_2}] \,&=\, A^{\dagger}_{\mathbf{k}_1}\, [A_{\mathbf{k}_1}, A^{\dagger}_{\mathbf{k}_2}]\, A_{\mathbf{k}_2} + A^{\dagger}_{\mathbf{k}_2}\, [A^{\dagger}_{\mathbf{k}_1}, A_{\mathbf{k}_2}]\, A_{\mathbf{k}_1} \\ &=\, \left(A^{\dagger}_{\mathbf{k}_1}\, A_{\mathbf{k}_1} - A^{\dagger}_{\mathbf{k}_1}\, A_{\mathbf{k}_1} \right) 2E_k\, \delta^{d-1}(\mathbf{k}_1 - \mathbf{k}_2) \,=\, 0. \end{aligned} \tag{7.10.13}$$

As in the familiar harmonic oscillator, the spectrum (7.10.11) derives from the commutation relations

$$[N_{\mathbf{k}}, A^{\dagger}_{\mathbf{k}}] \,=\, A^{\dagger}_{\mathbf{k}}, \qquad [N_{\mathbf{k}}, A_{\mathbf{k}}] \,=\, -A_{\mathbf{k}}. \tag{7.10.14}$$

These expressions say that $A^{\dagger}_{\mathbf{k}}$ creates a particle of momentum $\mathbf{k}$ while $A_{\mathbf{k}}$ annihilates such a particle. The state with the minimum energy is the vacuum state, in which there are no particles

$$N_{\mathbf{k}}\, |\, 0\, \rangle \,=\, 0. \tag{7.10.15}$$

This implies $A_{\mathbf{k}}\, |\, 0\, \rangle \,=\, 0$ and the multiparticle states with momenta $\mathbf{k}_1, \ldots, \mathbf{k}_n$ are given by

$$|\, n_{\mathbf{k}_1}\, , n_{\mathbf{k}_2}, \ldots \rangle \,=\, \frac{1}{(n_{\mathbf{k}_1}! n_{\mathbf{k}_2}! \ldots)^{1/2}} \left(A^{\dagger}_{\mathbf{k}_1} \right)^{n_{\mathbf{k}_1}} \left(A^{\dagger}_{\mathbf{k}_2} \right)^{n_{\mathbf{k}_2}} \ldots\, |\, 0\, \rangle. \tag{7.10.16}$$

Since the operators $A^{\dagger}_{\mathbf{k}_i}$ commute with each other, these states are symmetric under an exchange of the indices and therefore satisfy the Bose–Einstein statistics. The Hilbert space constructed in this way is called the *Fock space* of the theory.

In the light of this discussion, let us see what is the interpretation of the state $\varphi(x)\, |\, 0\, \rangle$. From the field expansion and the action of the operators A and $A^{\dagger}$, one has

$$\varphi(x, 0)\, |\, 0\, \rangle \,=\, \int d\Omega_k \, e^{-i\mathbf{k}\cdot\mathbf{x}}\, |\, \mathbf{k}\, \rangle, \tag{7.10.17}$$

where we have indicated by $|\, \mathbf{k}\, \rangle \,=\, A^{\dagger}_{\mathbf{k}}\, |\, 0\, \rangle$ the one-particle state of momentum $\mathbf{k}$. Therefore the state $\varphi(x) |\, 0\, \rangle$ is given by a linear superposition of one-particle states

of various moment. In other words, applying the field to the vacuum state *we have created a particle at the point* $\mathbf{x}$. This interpretation is further supported by computing the matrix element

$$\langle\, 0 \,|\, \varphi(\mathbf{x}) \,|\, \mathbf{k} \,\rangle \;=\; e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{7.10.18}$$

This is the coordinate representation of the wavefunction of a one-particle state, just as in quantum mechanics $\langle x|p\rangle = e^{ipx}$ is the wavefunction of the state $|\, p \,\rangle$.

7.11 Correlation Functions and Scattering Processes

In defining the correlation functions in Minkowski space we shall take into account that the fields are operators and therefore they do not generally commute. Quantities of interest are the vacuum expectation values of the T-ordered product of operators.[12] In the free case, the only non-zero correlation function of the field φ is the two-point correlators. It can be computed by using the commutation relations (7.10.5) and the relations $A\,|\,0\rangle = 0$ and $\langle 0\,|\,A^\dagger = 0$

$$\begin{aligned}
\Delta_F(x-y) \;&=\; \langle 0\,|T[\varphi(x)\varphi(y)]\,|\,0\,\rangle \\
&=\; \langle 0\,|\varphi^{(+)}(x)\varphi^{(-)}(y)]\,|\,0\,\rangle\,\theta(x_0-y_0) + \langle 0\,|\varphi^{(+)}(y)\varphi^{(-)}(x)]\,|\,0\,\rangle\,\theta(y_0-x_0) \\
&=\; \int d\Omega_k \left[e^{-ik\cdot(x-y)}\,\theta(x_0-y_0) + e^{ik\cdot(x-y)}\,\theta(y_0-x_0)\right].
\end{aligned} \tag{7.11.1}$$

This quantity is the so-called *Feynman propagator* that can be written in a relativistic invariant way as

$$\Delta_F(x-y) \;=\; \int \frac{d^dk}{(2\pi)^d}\,\frac{i}{k^2-m^2+i\epsilon}\,e^{-ik\cdot(x-y)}. \tag{7.11.2}$$

In this formula $k^2 = k_0^2 - \mathbf{k}^2$ and the $i\,\epsilon$ term in the denominator is equivalent to a prescription in computing the integral over the time component of the momentum: using the residue theorem for the integral on dk_0 it is easy to see that one obtains the previous formula (see Fig. 7.17). Note that using the analytic continuation $k^0 = ik_E^0$, the so-called Wick rotation, the Feynman propagator becomes (up to a factor i) the propagator of the euclidean quantum field theory, previously analyzed.

The Feynman propagator can also be obtained by generalizing the formula (7.9.12) in the limit $T \to \infty$, where T is the time separation between the two vacuum states on the right- and on left-hand sides

$$\langle 0\,|T[\varphi(x)\varphi(y)]|\,0\,\rangle \;=\; \frac{1}{Z_0}\int \mathcal{D}\varphi\,\varphi(x)\,\varphi(y)\,e^{i\int d^{d-1}x dt\,\tilde{\mathcal{L}}_0}. \tag{7.11.3}$$

As usual, Z_0 gives the proper normalization

$$Z_0 \;=\; \int \mathcal{D}\varphi\,e^{i\int d^{d-1}x dt\tilde{\mathcal{L}}_0}.$$

[12] In the following formulas, all vectors are d-dimensional with the Minkowski metric. Hence x denotes $(x^0, \mathbf{x})$ and $p\cdot x = p^0\,x^0 - \mathbf{p}\cdot\mathbf{x}$.

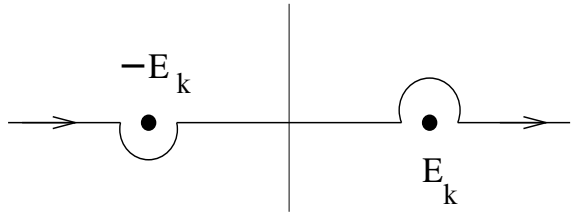


Fig. 7.17 *Integration contour of the variable k_0, which is equivalent to the $i\epsilon$ prescription in the denominator of (7.11.2).*

Analogously to the euclidean case, we can couple the field to an external current and define ($d^d x \equiv d^{d-1}x dt$)

$$Z_0[J] \,=\, \int \mathcal{D}\varphi \, \exp\left\{ i \int \left[\tilde{\mathcal{L}}_0 + J(x)\,\varphi(x) \right] d^d x \right\}. \tag{7.11.4}$$

The integral is gaussian

$$Z_0[J] \,=\, \exp\left\{ \frac{i}{2} \int\int J(x)\,\Delta_F(x-y)\,J(y)\,d^d x\, d^d y \right\},$$

and then

$$\Delta_F(x-y) \,=\, (-i)^2\, \frac{\delta^2 Z[J]}{\delta J(x) \delta J(y)}. \tag{7.11.5}$$

In the interactive case, the partition function is given by

$$Z[J] \,=\, \exp\left\{ i \int \tilde{\mathcal{L}}_I \left[-i \frac{\delta}{\delta J(x)} \right] d^d x \right\}\, Z_0[J], \tag{7.11.6}$$

and the correlation functions are defined by

$$G(x_1, \ldots, x_n) \,=\, \langle 0 \,|T[\varphi(x_1) \ldots \varphi(x_n)]\,|\, 0 \rangle \,=\, (-i)^n\, \frac{\delta^n\, Z[J]}{\delta J(x_1) \ldots \delta J(x_n)}. \tag{7.11.7}$$

They admit an expansion in terms of Feynman graphs, analogously to the one previously analyzed. One should take into account, though, an extra factor i for each vertex and a different expression for the propagator. The perturbative properties of the correlation functions are similar to those previously discussed.

Finally, we would like to comment on a different interpretation of the Feyman graphs in Minkowski space. Since the lines are now associated to the propagation of the particles, the various interaction vertices can be considered as the points of the scattering processes. For instance, the connected four-point function shown in Fig. 7.18 can be employed to compute the probability of the elastic scattering of two in-going particle of momenta k_1 and k_2 and out-coming particles with the same momenta.

Analogously, the connected n-point correlation functions can be used to compute the production processes of $(n-2)$ particles that originate from the collision of two

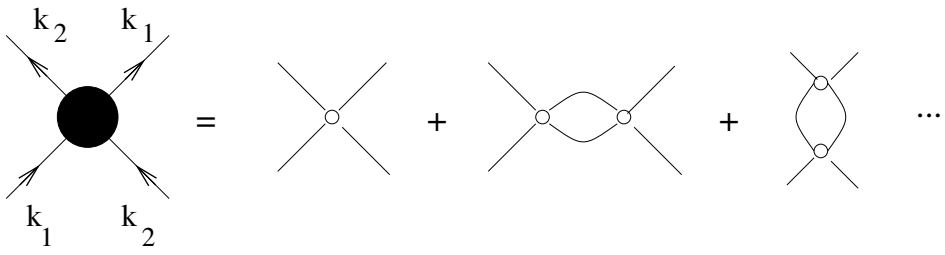


Fig. 7.18 *Elastic scattering amplitude of two particles, given by the infinite sum of all elementary interaction processes ruled by the interaction vertices.*

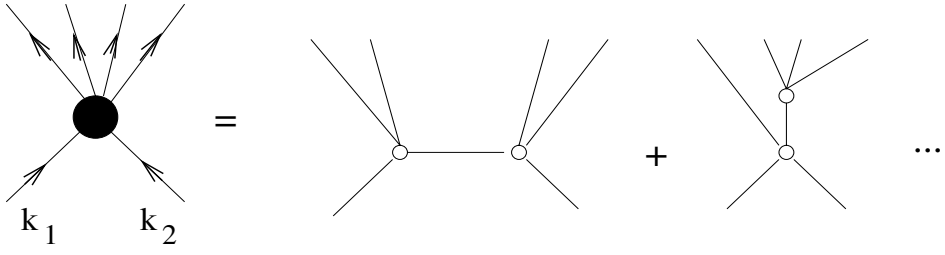


Fig. 7.19 *Production amplitude of a multiparticle state following a collision of two initial particles.*

initial particles, if they have enough energy in their center of mass (equal or larger than the sum of the mass of the $(n-2)$ particles, see Fig. 7.19). In the absence of conservation laws, all these processes are allowed by the relativistic laws. They will be studied in detail in Chapter 17.

Appendix 7A. Feynman Path Integral Formulation

Let $Q(t)$ be the coordinate operator of a quantum particle in the Heisenberg representation and $|q,t\rangle$ its eigenstates

$$Q(t)\,|q,t\rangle \,=\, q\,|q,t\rangle.$$

In the Schrödinger representation Q_S is a time-independent operator, related to $Q(t)$ by the unitary relation $Q(t) \,=\, e^{itH/\hbar}\,Q_S\,e^{-itH/\hbar}$. Q_S has time-independent eigenstates, $Q_S\,|q\rangle = q\,|q\rangle$, and their relation to the previous one is given by $|q\rangle \,=\, e^{-itH/\hbar}\,|q,t\rangle$. These states satisfy the completeness relation

$$\int dq\,|q\rangle\,\langle q| \,=\, 1.$$

It is also useful to introduce the eigenstates of the momentum operator in the Schrödinger representation

$$P\,|p\rangle \,=\, p\,|p\rangle.$$

They satisfy the completeness relation

$$\int \frac{dp}{2\pi} \, |p\rangle\langle p| \, = \, 1,$$

and their scalar product matrix elements with the states $|q\rangle$ is $\langle q|p\rangle = e^{ipq/\hbar}$. Let's compute the amplitude

$$F(q',t';q,t) \, = \, \langle q',t'|q,t\rangle \, = \, \langle q'|e^{-i(t'-t)H/\hbar} \, |q\rangle \tag{7.A.1}$$

dividing the interval $T = (t'-t)$ in $(n+1)$ time slices

$$t = t_0, \, t_1, \ldots, t_{n+1} = t', \qquad t_k = t_0 + k\epsilon.$$

In the limit $n \to \infty$, we have

$$e^{i(t'-t)H/\hbar} \simeq e^{-i\epsilon H/\hbar} \, e^{-i\epsilon H/\hbar} \ldots e^{-i\epsilon H/\hbar}.$$

Inserting n times the completeness relation of the eigenstates $|q\rangle$ into eqn (7.A.1) we get

$$F(q',t';q,t) \, = \, \int \prod_{k=1}^{n} dq_k \, \langle q'|e^{-i\epsilon H/\hbar}|q_n\rangle\langle q_n|e^{-i\epsilon H/\hbar}|q_{n-1}\rangle \ldots \langle q_1|e^{-i\epsilon H/\hbar}|q\rangle. \tag{7.A.2}$$

These matrix elements can be computed exactly in the limit $\epsilon \to 0$. With the hamiltonian given by $H \, = \, \frac{p^2}{2m} + V(q)$, inserting the completeness relation of the $|p\rangle$ state we have

$$\begin{aligned}
&\langle q_k|e^{-i\epsilon H/\hbar}|q_{k-1}\rangle \, = \, \int \frac{dp}{2\pi} \, \frac{dp'}{2\pi} \langle q_k|p\rangle\langle p|e^{-i\epsilon H/\hbar}|p'\rangle\langle p'|q_{k-1}\rangle \\
&= \int \frac{dp}{2\pi} \, \frac{dp'}{2\pi} e^{i(pq_k - p'q_{k-1})/\hbar} \, e^{-i\epsilon/\hbar\left(\frac{p^2}{2m} + V\left(\frac{q_k+q_{k-1}}{2}\right)\right)} \, \delta(p-p') \qquad (7.A.3) \\
&= \int \frac{dp}{2\pi} e^{ip(q_k - q_{k-1})} \, e^{-i\epsilon/\hbar\left(\frac{p^2}{2m} + V\left(\frac{q_k+q_{k-1}}{2}\right)\right)} \, = \, \frac{1}{\sqrt{2\pi\epsilon}} \, e^{i\epsilon/\hbar\left[\frac{(q_k-q_{k-1})^2}{2\epsilon^2} - V\left(\frac{q_k+q_{k-1}}{2}\right)\right]}.
\end{aligned}$$

Making the hypothesis that in the limit $\epsilon \to 0$, q_{k-1} tends to q_k, we have

$$\frac{(q_k - q_{k-1})^2}{\epsilon^2} \to (\dot{q})^2$$

and therefore the matrix element is expressed by the lagrangian associated to this part of the trajectory

$$\langle q_k|e^{-i\epsilon H/\hbar}|q_{k_1}\rangle \, \simeq \, \frac{1}{\sqrt{2\pi\epsilon}} \, e^{i\epsilon/\hbar L(\dot{q}_k, q_k)}. \tag{7.A.4}$$

Coming back to (7.A.2), one thus has

$$\begin{aligned}
F(q',t';q,t) \, &= \, \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{dq_k}{\sqrt{2\pi\epsilon}} \, e^{i\epsilon/\hbar \, \sum_{k=1}^{n}\left[\frac{1}{2}\dot{q}_k^2 - V(q_k)\right]} \\
&\equiv \int \mathcal{D}q \, e^{i/\hbar \int_t^{t'} dt \, L(q,\dot{q})} \, = \, \int \mathcal{D}q \, e^{i/\hbar A}.
\end{aligned} \tag{7.A.5}$$

Let us now consider the correlation function of two operators $Q(t_1)Q(t_2)$, with $t_1 > t_2$. Repeating the same argument given above, one arrives at a representation of this quantity in terms of a path integral

$$\langle q', t'|Q(t_1)Q(t_2)|q, t\rangle \;=\; \int \mathcal{D}q\, q(t_1)\, q(t_2)\, e^{i/\hbar A}. \tag{7.A.6}$$

However, notice that on the right-hand side the order of the two variables is irrelevant. The path integral expression is then equal to the matrix elements on the left-hand side with the established order, in which the only important thing is that $t_1 > t_2$. If t_1 was less than t_2, the right-hand side would be equal to the matrix element of the two operators but in reversed order. This leads to the definition of the time ordering of the operators

$$T[Q(t_1)Q(t_2)] \;=\; \begin{cases} Q(t_1)Q(t_2), & t_1 > t_2 \\ Q(t_2)Q(t_1), & t_2 > t_1 \end{cases} \tag{7.A.7}$$

with an obvious generalization for an arbitrary number of them. In such a way, we arrive at the formula

$$\langle q', t'|T[Q(t_1)\dots Q(t_k)]|q, t\rangle \;=\; \int \mathcal{D}q\, q(t_1)\dots q(t_k)\, e^{i/\hbar A}. \tag{7.A.8}$$

Appendix 7B. Relativistic Invariance

The Lorentz transformations in $(d+1)$ dimensions leave invariant the front line of the light, defined by

$$s^2 \;=\; t^2 - x_1^2 - \dots - x_{d-1}^2.$$

The speed of light c has been imposed equal to 1 and we have also used $t = x_0$ to make the notation uniform. More generally, with the definition of the metric tensor

$$g^{\mu\nu} \;=\; g_{\mu\nu} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0\dots & 0 \\ 0 & -1 & 0 & 0\dots & 0 \\ 0 & 0 & -1 & 0\dots & 0 \\ \dots & \dots & \dots & \dots\dots & \dots \\ 0 & 0 & 0 & 0\dots & -1 \end{pmatrix}$$

the Lorentz transformations Λ^μ_ν are defined by the condition to leave invariant the metric, i.e.

$$g_{\mu\nu}\, \Lambda^\mu_\rho\, \Lambda^\nu_\sigma \;=\; g_{\rho\sigma} \tag{7.B.1}$$

(with a sum over the repeated indices). Thanks to the metric tensor we can rise or low the indices of a vector or of a tensor. We have

$$x^\mu \;=\; (t, \mathbf{x}), \qquad x_\mu \;=\; g_{\mu\nu} x^\nu \;=\; (t, -\mathbf{x}).$$

For the derivative we have

$$\partial_\mu \,=\, \frac{\partial}{\partial x^\mu} \,=\, \left(\frac{\partial}{\partial x^0}, \nabla\right).$$

The space with such a metric is called *Minkowski space*. In order to characterize the infinitesimal form of these transformations, let's impose

$$\Lambda^\mu_\nu \,\simeq\, \delta^\mu_\nu + \omega^\mu_\nu.$$

Substituting into (7.B.1), one has

$$\omega_{\mu\nu} + \omega_{\nu\mu} \,=\, 0. \tag{7.B.2}$$

In d dimensions the number of free parameters of an antisymmetric matrix is equal to $d(d-1)/2$. If we add to these transformations also the translations $x^\mu \to x^\mu + a^\mu$, we arrive at the Poincarè group.

Invariant expressions under the Poincarè group are generically given by scalar products with respect to the metric tensors, such as

$$p \cdot x \,=\, g_{\mu\nu}\, p^\mu\, x^\nu \,=\, p^\mu\, x_\mu \,=\, p^0\, x^0 - \mathbf{p}\cdot\mathbf{x}.$$

Another invariant quantity is given by

$$\partial^\mu\, \partial_\mu \,=\, \Box \,=\, \frac{\partial^2}{\partial (x^0)^2} - \nabla^2.$$

The momentum of a massive particle satisfies

$$p^2 \,=\, p^\mu\, p_\mu \,=\, E^2 - \mathbf{p}^2 \,=\, m^2, \tag{7.B.3}$$

where m is the mass.

Since the norm of the vectors (with respect to the metric g) is an invariant, the distance between two points can be classified as follows: (a) if $(x_1 - x_2)^2 > 0$, this is a *time-like* separation; if $(x_1 - x_2)^2 < 0$ this is a *space-like* separation; and if $(x_1 - x_2)^2 = 0$ we have a *light-like* separation (see Fig. 7.20). Time-like points are related to each other by a causality relation while the space-like points are not. In the latter case, in fact, to have a causal relation between them, a signal should travel faster than the speed of light. For light-like events, the temporality of two events is given by the time that is necessary to the light to travel from x_1 to x_2.

It is easy to prove that the volume elements

$$d^d x \,\equiv\, dx_0\, dx_1\, \ldots dx_{d-1}, \tag{7.B.4}$$

and the momentum volume elements

$$d^d p \,\equiv\, dp_0\, dp_1\, \ldots dp_{d-1} \tag{7.B.5}$$

are both invariant: under a Lorentz transformation, to a dilatation of the time component there corresponds a contraction of the space component, and the two terms

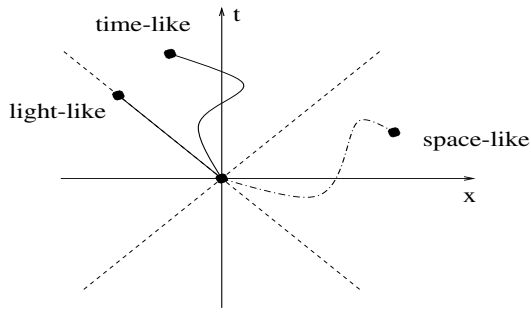


Fig. 7.20 *With x_2 placed at the origin, the point x_1 can be in one of the three positions shown in the figure. For time-like distances, the event x_2 can be in a causal relation with the event x_1. For space-like distances, the two events cannot be linked by a causal relation, since their time separation is larger than the time that the light would spend to cover their spatial distance.*

compensate each other. Using the invariance of the momentum infinitesimal volume, one can prove the invariance of the measure $d\Omega_k$. In fact, it can be written as

$$d\Omega_k \;=\; \frac{d^{d-1}k}{(2\pi)^{d-1}2E_p} \;=\; \frac{d^d p}{(2\pi)^d}\,(2\pi)\,\delta(k^2-m^2)\Big|_{k^0>0}\,. \tag{7.B.6}$$

The lagrangian that appears in (7.9.3) is a scalar density. It gives rise to the equation of motion thanks to the principle of minimum action

$$\begin{aligned} 0 \;=\; \delta\tilde{\mathcal{S}} \;&=\; \int d^d x \left\{ \frac{\partial\tilde{\mathcal{L}}}{\partial\varphi}\,\delta\varphi + \frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_\mu\varphi)}\,\delta(\partial_\mu\varphi) \right\} \\ &=\; \int d^d x \left\{ \left[\frac{\partial\tilde{\mathcal{L}}}{\partial\varphi} - \partial_\mu \frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_\mu\varphi)} \right] \delta\varphi + \partial_\mu \left(\frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_\mu\varphi)}\,\delta\varphi \right) \right\}. \end{aligned} \tag{7.B.7}$$

The last term is a total divergence and it gives rise to a surface integral. This vanishes if we assume that the variation of the field is zero at the boundary. In this way, we arrive to the Euler–Lagrange equation of the field

$$\frac{\partial\tilde{\mathcal{L}}}{\partial\varphi} - \partial_\mu \frac{\partial\tilde{\mathcal{L}}}{\partial(\partial_\mu\varphi)} \;=\; 0. \tag{7.B.8}$$

Appendix 7C. Noether's Theorem

There is a deep relation between the symmetries and the conservation laws of a system. This is the content of Noether's theorem. Suppose we change infinitesimally the field

$$\varphi(x) \to \varphi'(x) + \alpha\,\delta\varphi, \tag{7.C.1}$$

where α is an infinitesimal parameter and $\delta\varphi$ is a deformation of the field. Such a transformation is a symmetry of the system if it leaves invariant the equations of motion. To guarantee this condition it is sufficient that the action remains invariant

under the transformations (7.C.1). More generally, the action is allowed to change up to a surface term, since the latter does not effect the equation of motion. Hence, under (7.C.1), the lagrangian can change at most by a total divergence

$$\tilde{\mathcal{L}} \to \tilde{\mathcal{L}} + \alpha\, \partial_\mu \mathcal{J}^\mu(x).$$

Comparing this expression with the expression that is explicitly obtained by varying the field in the lagrangian according to (7.C.1) one has

$$\begin{aligned} \alpha\, \partial_\mu \mathcal{J}^\mu \;&=\; \frac{\partial \tilde{\mathcal{L}}}{\partial \varphi}(\alpha\delta\varphi) + \left(\frac{\partial \tilde{\mathcal{L}}}{\partial(\partial_\mu \varphi)}\right) \partial_\mu(\alpha\delta\varphi) \\ &=\; \alpha\, \partial_\mu \left(\frac{\partial \tilde{\mathcal{L}}}{\partial(\partial_\mu \varphi)}\, \delta\varphi\right) + \alpha \left[\frac{\partial \tilde{\mathcal{L}}}{\partial \varphi} - \partial_\mu \frac{\partial \tilde{\mathcal{L}}}{\partial(\partial_\mu \varphi)}\right] \delta\varphi. \end{aligned} \tag{7.C.2}$$

The last term vanishes for the equation of motion and therefore we arrive at the conservation law

$$\partial_\mu\, j^\mu(x) \;=\; 0, \qquad j^\mu(x) \;\equiv\; \frac{\partial \tilde{\mathcal{L}}}{\partial(\partial_\mu \varphi)}\, \delta\varphi - \mathcal{J}^\mu. \tag{7.C.3}$$

Let's see the consequence of this result if the system is invariant under the translations $x^\mu \to x^\mu - a^\mu$. The field changes as

$$\varphi(x) \to \varphi(x+a) \;=\; \varphi(x) + a^\mu \partial_\mu\, \varphi(x). \tag{7.C.4}$$

Since the lagrangian is a scalar quantity, it transforms in the same way:

$$\tilde{\mathcal{L}} \to \tilde{\mathcal{L}} + a^\mu \partial_\mu \tilde{\mathcal{L}} \;=\; \tilde{\mathcal{L}} + a^\nu\, \partial_\mu(\delta^\mu_\nu\, \tilde{\mathcal{L}}).$$

Using (7.C.3), we obtain the so-called *stress–energy tensor*

$$T^\mu_\nu \;\equiv\; \frac{\partial \tilde{\mathcal{L}}}{\partial(\partial_\mu \varphi)}\, \partial_\nu \varphi - \tilde{\mathcal{L}}\, \delta^\mu_\nu \tag{7.C.5}$$

that satisfies

$$\partial_\mu\, T^\mu_\nu \;=\; 0. \tag{7.C.6}$$

The energy and the momentum of the system is given by

$$E \;=\; \int d^d x\, T^{00}(x,t), \qquad P^\nu \;=\; \int d^d x\, T^{0\nu}(x,t) \tag{7.C.7}$$

References and Further Reading

The path integral formulation of quantum mechanics is due to Richard Feynman. For a detailed discussion consult the book:

R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.

There are many superb texts on quantum field theory. The reader can consult, for instance:

A. Zee, *Quantum Field Theory in a Nutshell*, Princeton University Press, Panceton, 2003.

R. Barton, *Introduction to Advanced Field Theory*, John Wiley, New York, 1963.

The functional formalism and its relation with statistical mechanics can be found in:

D. Amit, *Field Theory, the Renormalization Group and Critical Phenomena*, Mc-Graw Hill, New York, 1978.

J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Oxford University Press, Oxford, 1971.

G. Parisi, *Statistical Field Theory*, Benjamin/Cummings, New York, 1988.

J.J. Binney, N.J. Dowrick, A.J. Fisher and M.E.J. Newman, *The Theory of Critical Phenomena*, Oxford University Press, Oxford, 1992.

The operatorial approach is discussed in the books:

N.N. Bogoliubov, D.V. Shirkov, *Introduction to the Theory of Quantized Fields*, John Wiley, New York, 1976.

S. Weinberg, *The Quantum Theory of Fields*, Cambridge University Press, Cambridge, 1995.

The application of quantum field theory to elementary particles is discussed in:

M.E. Peskin, D.V. Schroeder, *An Introduction to Quantum Field Theory*, Adison Wesley, New York, 1995.

Problems

1. Lagrangian theory with Z_3 symmetry

Consider a lagrangian of a complex field $\Phi(x)$ and its conjugate $\Phi^\dagger(x)$ which under a Z_3 transformation transform as

$$\Phi(x) \,\to\, e^{2\pi i/3}\,\Phi(x), \qquad \Phi^\dagger(x) \,\to\, e^{-2\pi i/3}\,\Phi^\dagger(x).$$

Write down the most general lagrangian that is invariant under these transfomations.

2. Perturbative series

Consider the one-dimensional integral

$$I(\lambda) \,=\, \int_{-\infty}^{+\infty} dx\, e^{-\alpha x^2+\lambda x^4}.$$

Write the perturbative series of this expression expanding the term $e^{-\lambda x^4}$ in a power series of λ. Compute the perturbative coefficients and show that the series has zero radius of convergence. Give a simple argument of this fact.

3. Correlation functions and Feynman graphs

Draw the Feynman diagrams relative to the g^2 correction of the four-point correlation function $G(x_1, \ldots, x_4)$ for the φ^4 theory. Discuss the convergence of the integrals as functions of the dimensionality d of the system.

4. $\boldsymbol{\varphi^3}$ lagrangian theory

Calculate the first non-vanishing perturbative order of the partition function for the lagrangian theory with interaction $\frac{g}{3!}\,\varphi^3$. Determine the upper critical dimension d_s and discuss the renormalization of this theory.

5. Dimensional regularization

An alternative way to regularize the integrals encountered in perturbative series of quantum field theory consists of the dimensional regularization. The main idea behind this approach is to consider the integrals as functions of the dimensionality d of the system, regarded as a continuous variable. Once they are evaluated in the region of the complex plane d where they converge, their values in other domains are obtained by analytic continuation. Prove the validity of the formula

$$\int \frac{d^d p}{(2\pi)^d} \frac{1}{(p^2+\Delta)^n} \;=\; \frac{1}{(4\pi)^{d/2}} \frac{\Gamma\left(n-\frac{d}{2}\right)}{\Gamma(n)} \left(\frac{1}{\Delta}\right)^{n-d/2}.$$

Discuss the analytic structure of this expression as a function of d.

6. Invariant functions

Consider the functions

$$\Delta^{(\pm)}(x) \;=\; \frac{i}{(2\pi)^d} \int d^{d-1}k \int_{C^{(\pm)}} dk_0 \, \frac{e^{ik\cdot x}}{k^2 - m^2},$$

where the contours of integration are shown in Fig. 7.21.

a Show that the correlation function of the commutator of the field is given by

$$\langle 0 \,|\, [\varphi(x), \varphi(y)\,] \,|\, 0 \,\rangle \;=\; \Delta(x-y),$$

where

$$\Delta(x-y) \;=\; \Delta^{(+)}(x-y) + \Delta^{(-)}(x-y).$$

b Prove that $\Delta(x)$ vanishes for equal times

$$\Delta(\mathbf{x}-\mathbf{y}, 0) \;=\; 0.$$

Using Lorentz invariance, argue that the relation above implies the vanishing of $\Delta(\mathbf{x}-\mathbf{y})$ for all space-like intervals. From a physical point of view, the commutativity of the field for space-like intervals is a consequence of the causality

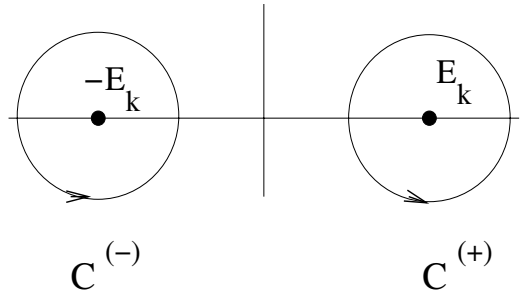


Fig. 7.21 *Contours of integration* $C^{(+)}$ *and* $C^{(-)}$ *for* $\Delta^{(+)}(x)$ *and* $\Delta^{(-)}(x)$.

principle: since space-like points cannot be related by light signals, the measures done at the two points cannot interfere and therefore the operators commute.

c Prove that $\Delta(x)$ and $\Delta^{(\pm)}(x)$ satisfy the homogenous equation

$$(\Box + m^2)\,\Delta(x) \;=\; (\Box + m^2)\Delta^{(\pm)}(x) \;=\; 0,$$

while the Feynman propagator, which corresponds to an infinite contour of integration, satisfies

$$(\Box + m^2)\,\Delta_F(x) \;=\; -i\,\delta^d(x).$$

7. Field theories with soliton solutions

Consider the lagrangian field theory in $1+1$ dimensions

$$\tilde{\mathcal{L}} \;=\; \frac{1}{2}(\partial_\mu\varphi)^2 + \frac{m^2}{\beta^2}\,\left[\cos(\beta\varphi) - 1\right].$$

a Expand in powers of β and show that this model corresponds to a Landau–Ginzburg theory with an infinite number of couplings.

b Write the equation of motion of the field $\varphi(x,t)$.

c Prove that the configurations

$$\varphi^{(\pm)}(x,0) \;=\; \pm 4\,\arctan\left[\exp(x - x_0)\right]$$

(where x_0 is an arbitrary point) are both classical solutions of the static version of the equation of motion.

d Show that these configurations interpolate between two next neighbor vacua. These configurations correspond to topological excitations of the field, called *solitons* and *antisolitons*.

e Compute the stress–energy tensor and use the formula $H = \int T^{00}(x)\,dx$ to determine the energy of the solitons. Since they are static, their energy corresponds to their M. Prove that

$$M \;=\; \frac{8m}{\beta}.$$

Note that the coupling constant is in the denominator, so that this is a non-perturbative expression.

8. Antiparticles

Consider the free theory of a complex field $\phi(x)$. In Minkowski space the action is

$$\mathcal{S} = \int d^d x \left(\partial_\mu \phi^* \, \partial^\mu \phi - m^2 \phi^* \, \phi\right).$$

a Show that the Hamiltonian is given by

$$H = \int d^{d-1} x \, (\pi^* \pi + \nabla \phi^* \cdot \nabla \phi + m^2 \phi^* \phi).$$

b Prove that the system is invariant under the continuous symmetry

$$\phi \to e^{i\alpha} \phi, \qquad \phi^* \to e^{-i\alpha} \phi.$$

Use Noether's theorem to derive the conserved charge

$$Q = -i \int d^{d-1} x (\pi^* \phi^* - \pi \phi).$$

c Diagonalize the hamiltonian by introducing the creation and annihilation operators. Show that the theory contains two sets of operators that can be distinguished by the different eigenvalues of the charge Q: the first set describes the creation and the annihilation of a particle A while the second one describes the same processes for an antiparticle $\bar{A}$.

d Show that the propagation of a particle in a space-like interval is the same as the propagation of an antiparticle back in time.

9. Conserved currents

Consider a multiplet of n scalar fields, $\Phi = (\phi_1, \ldots, \phi_n)$.

a Write the most general lagrangian that is invariant under a rotation of the vector Φ

$$\Phi_k \to (\mathcal{R})_{kl} \Phi_l.$$

b Use the Noether theorem to derive the conserved currents associated to this symmetry.

8
Renormalization Group

Everything must change so that nothing changes.
Giuseppe Tomasi di Lampedusa, *Il Gattopardo*

8.1 Introduction

At a critical point, the correlation length ξ diverges: the statistical fluctuations extend on all scales of the system and any attempt to solve the dynamics by taking into account only a finite number of degrees of freedom fails. In the absence of an exact solution of the model under consideration, the computation of the critical exponents is often obtained only by numerical methods and Monte Carlo simulations.

Leaving apart the problem of computing the critical exponents, there is however a general approach to phase transitions that has the advantage of conceptually simplifying many of their aspects. This approach goes under the name of the *renormalization group* (in short, RG). Beside its practical use, the fundamental ideas of the RG provide a theoretical scheme and a proper language to face critical phenomena and, in particular, to understand their universal properties and scaling laws.

It is worth stressing that the terminology is inappropriate for two reasons: (i) the transformations of the RG are irreversible and therefore they do not form a group, as usually meant in mathematics; (ii) moreover, they do not necessarily concern the renormalization of a theory, i.e. the cure of the divergencies of the perturbative series. As a matter of fact, the main concepts of the renormalization group have a wider spectrum of validity.

There are many specialized books on the renormalization group and its technical aspects. The interested reader can find a small list of them to the end of the chapter. The aim of this chapter is to present in the simplest possible way the physical scenario provided by the RG, introducing the appropriate terminology and emphasizing the main concepts with the help of some significant examples. Other important aspects of the RG will be discussed in more detail in Chapter 15, in relation with two-dimensional quantum field theories near to their critical points.

What is the key idea behind the renormalization group? The answer to this question is: a continuous family of transformations of the coupling constants in correspondence to a change of the length-scale of a physical system. In any physical system there are various length-scales and the main assumption of the renormalization group is that they are couple together in a local way. If one is interested in studying, for instance, the fluctuations of a magnetic system on a scale of the order of 1000 Å, it is reasonable to assume that it would be sufficient to consider only the degrees of freedom with

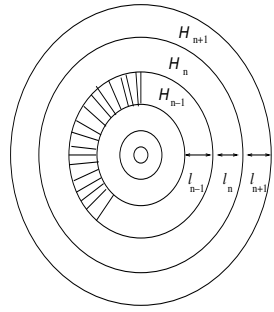


Fig. 8.1 *Length-scales and sequence of the effective hamiltonians for the degrees of freedom of each length shell.*

comparable wavelengths L, say those in the range 800 Å $< L <$ 1200 Å. The degrees of freedom with very short wavelength, of the order of a few atomic spacings, should not matter. If this is indeed the case, one is led to the conclusion that the interactions have a shell structure: the fluctuations of the system on scales of 1–2 Å only influence those on scales 2–4 Å, the last ones influence those on scales 4–8 Å, and so on. This sequence is ruled by a family of effective hamiltonians associated to the degrees of freedom that are relevant to each shell of the length-scale, as shown in Fig. 8.1.

There are two important aspects that emerge in this cascade scenario. The first aspect concerns the scaling invariant properties. With the absence of a characteristic length to compare with, the fluctuations of the intermediate lengths tend generally to be the same, besides a simple rescaling. There is, however, no scale invariance for those fluctuations with wavelengths comparable to a length parameter, such as the one provided for instance by the lattice spacing. The second aspect is the amplification or de-amplification phenomena that take place in the course of the cascade process. A small change of temperature may have a negligible effect at the atomic scale but, if this effect gets amplified to the large scales of the system, it may produce significant macroscopic changes. This is precisely what happens at the critical value T_c of the temperature, when the correlation length diverges, inducing all other thermodynamical singularities. Concerning the de-amplification effects, they are at the root of the universality properties of the critical phenomena: it is thanks to them that two magnetic materials, with quite different atomic compositions, may nevertheless share the same critical behavior.

To implement the ideas of the RG, the first steps consist of isolating a particular shell of length-scale and defining a procedure that permits us to pass to the next one. In the case of critical phenomena, this procedure involves a statistical average of all fluctuations within a certain range of lengths. This is equivalent to studying the behavior of the system under a length-scale $x \to x' = x/b$ or, tantamountly, under a rescaling of the lattice spacing $a \to a' = ba$. In doing so, one is simply looking at the system under a different magnifying glass. As a result, an effective hamiltonian is defined for the degrees of freedom that were not averaged. Implementing iteratively this average procedure, one is able to determine the amplification and de-amplification factors λ_i. These are the eigenvalues of the linearized version of the

iterative procedure: under a small change of the initial interactions, the λ_i are the quantities responsible for their amplification/deamplification to the next iteration. If a coupling constant gets amplified, it is called a *relevant coupling*. If, on the contrary, it gets deamplified, is called an *irrelevant coupling*. The instability nature of a critical point is determined by the number of its relevant couplings.

Roughly speaking, there are two different ways to implement the RG ideas. The first method is usually employed in contexts of quantum field theory to deal with the divergence of the Feynman diagrams discussed in the previous chapter. Since these computations are usually carried out in momentum space, this implementation of the RG goes under the heading of the *renormalization group in k space*. The second way is known as the *real space renormalization group*. This approach is more relevant in a statistical mechanics context, in particular in the discussion of systems defined on a lattice. Moreover, it is more intuitive. For this reason, in the following we will mainly follow this approach.

8.2 Reducing the Degrees of Freedom

Let us consider a statistical system defined on a d-dimensional regular lattice of lattice spacing a, with degrees of freedom s_i placed on its sites. Let $H(\{s_i\}, g_k)$ be the hamiltonian of the system, where g_k are the coupling constants of the various interactions among the spins s_i. For reasons that will become clear later, it is convenient to include in the hamiltonian all the possible coupling constants that are compatible with the nature of the degrees of freedom. For instance, if the s_i are Ising variables, the hamiltonian H can be written as $H = H^{(+)} + H^{(-)}$, where the $\pm$ signs refer to the even and odd sectors of the Z_2 symmetry of the model. In the even sector, the most general hamiltonian is given by

$$H^{(+)}(\{s_i\}, g_k) \;=\; \sum_{i,j} g^{(2)}_{ij}\, s_i\, s_j + \sum_{i,j,k,l} g^{(4)}_{ijkl} s_i\, s_j\, s_k\, s_l + \cdots \tag{8.2.1}$$

whereas in the odd sector, the most general hamiltonian is expressed by

$$H^{(-)}(\{s_i\}, g_k) \;=\; \sum_{i} g^{(1)}_{i}\, s_i + \sum_{i,j,k} g^{(3)}_{ijk} s_i\, s_j\, s_k + \dots \tag{8.2.2}$$

In the formulas above, the indices are not necessarily restricted to next neighbor sites. The partition function is given by

$$Z(\{g_k\}) \;=\; \sum_{\{s_i\}} \exp\left[-\,H(\{s_i\}, g_k)\right], \tag{8.2.3}$$

where we have included the factor $\beta = 1/KT$ in the definition of the coupling constants of the hamiltonian. At given values of the g_k, the system has a correlation length $\xi(g_k)$ that is a function of the couplings. This quantity measures the number of degrees of freedom effectively coupled together and one expects that, the smaller is $\xi(g_k)$, the more effective and accurate is a perturbative study of the model. This observation suggests we look for a scale transformation $a \to b\, a$ that establishes a correspondence

between the system with correlation length ξ and the one with correlation length $\xi' = \xi/b < \xi$. The idea is that, if such a transformation exists, its implementation may lead to a solvable or, at least, to a simpler model. Note, however, that, if the initial system is exactly at the critical point, this transformation will leave it invariant: in this case $\xi = \infty$ and therefore it remains a divergent quantity under any rescaling of the lattice spacing.

Spins within a sphere of radius ξ are correlated with each other. Therefore, those within a length shell ba ($b > 1$) satisfying

$$a \ll b\,a \ll \xi$$

act somehow as a single unit. We can imagine zooming in on the system, organizing the variables in spin blocks. Namely, let's divide the original lattice into blocks, denoted by $\mathcal{B}_k$, each of them made of b^d spins. If N is the total number of sites, there are $N\,b^{-d}$ blocks. Once this partition has been done, let's assign to each block a new variable $\sigma_i^{(1)}$ according to a certain law that involves the spins s_i present in each block

$$\sigma_i^{(1)} \;=\; f(\{s_i\}), \quad \text{with } i \in \mathcal{B}_k. \tag{8.2.4}$$

Postponing until later the discussion of the nature of this law, for the time being let's note that the effect of this transformation is to change the model into a new one, defined on a lattice with a new lattice spacing $a' = b\,a$. After all the dynamical variables have been changed according to the transformation (8.2.4), it is convenient to scale the new lattice by a factor b^{-1} (without altering, though, the spins), so that we come back to a lattice equal to the original one. What we have described above is the implementation of the real space renormalization group, which therefore consists of the iteration of the series of transformations

$$\sigma_k^{(n+1)} \;=\; f(\{\sigma_i^{(n)}\}), \quad \text{with } i \in \mathcal{B}_k \tag{8.2.5}$$

where $\sigma_i^{(n)}$ denote the spin variables of the n-th step of this procedure (see Fig. 8.2).

An important aspect of this transformation is its local nature: the definition of the variables $\sigma_k^{(n+1)}$ only involves the variables $\sigma_i^{(n)}$ and not the original spins s_i. Note that at each step of the procedure we lose information on the fluctuations of the spins that occur on the factor scale b.

8.3 Transformation Laws and Effective Hamiltonians

There are several reasonable choices of the transformation laws $f(\{\sigma_i\})$ for updating the spin variables and each of them gives rise to different RG coarse grainings of the system. However, one should realize that what really matters is the asymptotic behavior of the adopted iterative procedure. In the limit $n \to \infty$ the difference between the transformation laws may be washed out, leading to the same physical scenario. In the real space version of the renormalization group, the two versions mostly used are

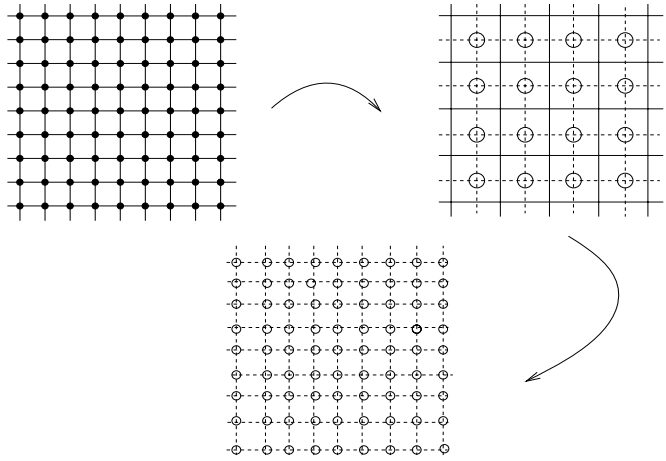


Fig. 8.2 *Sequence of a renormalization group transformation in real space: from the original lattice, with lattice spacing* $\mathbf{a}$ *and variables* s_i*, to a new lattice with* $\mathbf{a}' = b\mathbf{a}$ *and block spins* σ_i*. Finally, a scale transformation restores the original lattice spacing* $\mathbf{a}$*.*

the following:

- **Decimation**. This law assigns to the spin $\sigma_k^{(n+1)}$ the value of one of the spins $\sigma_i^{(n)}$ of the block $\mathcal{B}_k$, say the central one

$$\sigma_k^{(n+1)} \, = \, \sigma_j^{(n)}, j \, \in \, \mathcal{B}_k.$$

- **Majority rule**. This law assigns to the spin $\sigma_k^{(n+1)}$ the value of the majority of the spins $\sigma_i^{(n)}$ of the block $\mathcal{B}_k$, namely

$$\sigma_k^{(n+1)} \, = \, A^{(n)} \sum_{i \in \mathcal{B}_k} \sigma_i^{(n)},$$

  where $A^{(n)}$ is a normalization constant. One has to be careful to implement the latter procedure for it depends on the nature of the spins σ_i. For instance, if they are Ising variables with values ± 1, it is convenient to choose blocks with an odd number of spins, in order to avoid the possibility of generating a null value for the next block spins. The normalization constant $A^{(n)}$ is useful to re-establish the correct range of values ± 1 for the new variable.

Note the irreversible nature of both transformations above: knowing the value $\sigma_k^{(n+1)}$ it is indeed impossible to trace back the spins $\sigma_i^{(n)}$ that have generated it.

Given a transformation law, it is convenient to introduce the operator

$$T(\sigma_k^{(n+1)}, \sigma_i^{(n)}) \, = \, \begin{cases} 1 \text{ , if } \sigma_k^{(n+1)} \, = \, f(\{\sigma_i^{(n)}\}) \\ 0 \text{ , otherwise.} \end{cases} \tag{8.3.1}$$

It satisfies

$$\sum_{\{\sigma_k^{(n+1)}\}} T(\sigma_k^{(n+1)}, \sigma_i^{(n)}) \,=\, 1. \tag{8.3.2}$$

Fixed the transformation law of the spins, we pass to determine the *effective hamiltonian* $H^{(n+1)}(\{\sigma_i^{(n+1)}\}, g_k^{(n+1)})$ for the new block spins. Since the transformation (8.2.5) depends only on the configurations of the spins $\sigma_i^{(n)}$, the new hamiltonian will be determined by the n-th step hamiltonian $H^{(n)}(\{\sigma_i^{(n)}\}, g_k^{(n)})$ as follows. Let's denote by

$$P(\{\sigma_i^{(n)}\}) \,=\, \exp\left[-H^{(n)}\left(\{\sigma_i^{(n)}\}, g_k^{(n)}\right)\right],$$

the probability of realizing a configuration $\sigma_i^{(n)}$ and define the new hamiltonian by means of the conditional probability

$$\begin{aligned} &\exp\left[-H^{(n+1)}\left(\{\sigma_k^{(n+1)}\}, g_k^{(n+1)}\right)\right] \\ &\quad = \sum_{\{\sigma_i^{(n)}\}} \prod_{blocks} T(\sigma_k^{(n+1)}, \sigma_i^{(n)}) \exp\left[-H^{(n)}\left(\{\sigma_i^{(n)}\}, g_i^{(n)}\right)\right]. \end{aligned} \tag{8.3.3}$$

In other words, assigning the new block spins $\sigma_k^{(n+1)}$ according to the transformation law, the spins $\sigma_i^{(n)}$ of the previous step are averaged using as weight their Boltzmann factor. The result is the Boltzmann factor of the new block spins.

To avoid the introduction of some additive constants as we go on in the iteration of the procedure, it may be useful to fix a normalization condition for the sequence of hamiltonians, as for instance

$$\sum_{\{\sigma_i^{(n)}\}} H^{(n)}\left(\{\sigma_i^{(n)}\}, g_i\right) \,=\, 0.$$

Using this normalization, one has

$$\sum_{\{\sigma_k^{(n+1)}\}} \exp\left[-H^{(n+1)}\left(\{\sigma_k^{(n+1)}\}, g_k^{(n+1)}\right)\right] \,=\, \sum_{\{\sigma_i^{(n)}\}} \exp\left[-H^{(n)}\left(\{\sigma_i^{(n)}\}, g_i^{(n)}\right)\right] \tag{8.3.4}$$

and the same value of the partition functions

$$Z^{(n+1)}(g_k^{(n+1)}) \,=\, Z^{(n)}(g_i^{(n)}). \tag{8.3.5}$$

This equality also holds for the expectation value of any function X of the variables $\sigma_k^{(n+1)}$: this is independent whether we compute it by using $H^{(n+1)}$ or $H^{(n)}$ in view of the identity

$$\begin{aligned} \langle X \rangle \,&=\, \frac{1}{Z^{(n+1)}} \sum_{\{\sigma_k^{(n+1)}\}} X(\{\sigma_k^{(n+1)}\}) \exp\left[-H^{(n+1)}\left(\{\sigma_k^{(n+1)}\}, g_k^{(n+1)}\right)\right] \\ &=\, \frac{1}{Z^{(n)}} \sum_{\{\sigma_i^{(n)}\}} X(\{\sigma_k^{(n+1)}\}) \exp\left[-H^{(n)}\left(\{\sigma_i^{(n)}\}, g_i^{(n)}\right)\right]. \end{aligned} \tag{8.3.6}$$

This allows us to refer to the expectation values without specifying which effective hamitonian has been used.

Manifold of the coupling constants. To implement successfully the procedure of the RG it is obviously important that the new effective hamiltonian $H^{(n+1)}$ has the same functional form as $H^{(n)}$, so that the model remains the same at each step of the sequence, beside a change in the value of its coupling constants. As a matter of fact this is impossible if we restrict attention to the hamiltonians with a finite number of couplings, since at each step new couplings are generated: for instance, starting from a hamiltonian with interaction among the next neighbor spins, the new hamiltonian has a new interaction among the spins separated by more than a lattice spacing and, furthermore, interactions that involve more than two spins. For this reason, it is convenient to start from the very beginning with the ensemble of all possible coupling constants that are compatible with the symmetry of the model and the nature of the statistical variables. Let's introduce then the *manifold of the coupling constants* and denote by $\{g^{(n)}\} \equiv (g_1^{(n)}, g_2^{(n)}, \ldots)$ the set of all the couplings of the effective hamiltonian $H^{(n)}$. In such a manifold, the application of eqn (8.3.3) can be interpreted as a motion of the point $\{g\}$ that identifies the system. This motion is made in discrete time steps and ruled by

$$\{g^{(n+1)}\} = \mathcal{R}(\{g^{(n)}\}), \tag{8.3.7}$$

where $\mathcal{R}$ is, in general, a complicated nonlinear transformation. Starting from a point $\{g^{(0)}\}$ and applying (8.3.7), the point of the system evolves in the sequence $\{g^{(1)}\}$, $\{g^{(2)}\}, \ldots$, giving rise in this way to a *renormalization group trajectory*, as shown in Fig. 8.3. It is important to stress that all points of the trajectory describe the *same physical situation*: they simply correspond to an observation of the system with a different magnifying glass. Note that under the transformation (8.3.7), the correlation length has to be measured with respect to the new lattice spacing and therefore it changes as

$$\xi(g^{(n+1)}) = b^{-1}\, \xi(g^{(n)}). \tag{8.3.8}$$

Hence, it shrinks by a factor b at each step of the procedure.

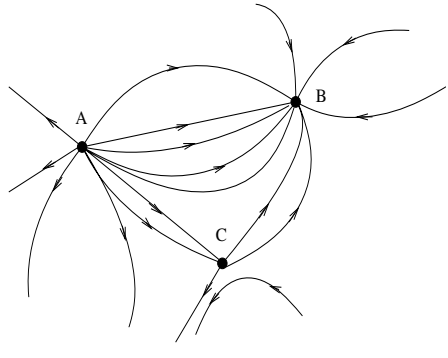


Fig. 8.3 *Trajectories of the renormalization group and fixed points: A is a repulsive fixed point, B is an attractive fixed point, whereas C is a mixed fixed point.*

8.4 Fixed Points

The mathematical nature of the renormalization group transformations is the same as *dynamical systems*, an important subject of physics and mathematics. An example of a dynamical system is provided by the logistic map discussed in Problem 1 at the end of the chapter. A priori, one could expect an arbitrary behavior for the trajectories that starts from a point P in the space of the coupling constants, with oscillations, discontinuities, or a zig-zag behavior. However, in all cases of physical relevance, one observes a smooth convergence toward some fixed points. A *fixed point* is a point in the manifold of the coupling constants that remains invariant under the mapping (8.3.7):

$$g^* \,=\, \mathcal{R}(g^*). \tag{8.4.1}$$

At a fixed point, the correlation length either diverges or vanishes, as can be easily seen from eqn (8.3.8) evaluated at $g = g^*$

$$\xi(g^*) \,=\, b^{-1}\,\xi(g^*). \tag{8.4.2}$$

Nature of fixed points. The fixed points where $\xi = \infty$ are called *critical points*, whereas those $\xi = 0$ are called *trivial fixed points*. The fixed points can be further classified by their stability nature: they can be *attractive*, *repulsive*, or *mixed*. One has an attractive fixed point if, in a neighborhood of g^*, the iteration of the transformations $g^{(n)}$ converges to g^*. One has instead a repulsive fixed point if the iteration of the RG transformations that start near g^* moves the point away from g^*. A mixed fixed point has both kinds of trajectories in its vicinity.

Linearization. The nature of the fixed points can be determined by studying the linear version of the transformation (8.3.7): putting $g = g^* + \delta g$, one has

$$g^* + \delta g' \,=\, \mathcal{R}(g^* + \delta g) \simeq \mathcal{R}(g^*) + \mathcal{K}\,\delta g \,=\, g^* + \mathcal{K}\,\delta g,$$

namely

$$\delta g'_a \,=\, \mathcal{K}_{ab}\,\delta g_b, \tag{8.4.3}$$

where the matrix $\mathcal{K}_{ab}$ is defined as

$$\mathcal{K}_{ab} \,=\, \frac{\partial \mathcal{R}_a}{\partial g_b}. \tag{8.4.4}$$

This matrix is not necessarily symmetric and for this reason it is necessary to distinguish between the right and the left eigenvectors. Denoting by λ^i its eigenvalues and by Δ^i its left eigenvectors $\mathcal{K}$, we have

$$\sum_a \Delta^i_a\,\mathcal{K}_{ab} \,=\, \lambda^i\,\Delta^i_a. \tag{8.4.5}$$

In terms of Δ^i_a let's now define a linear combination of the displacements δg_a

$$u_i \equiv \sum_a \Delta^i_a\,\delta g_a. \tag{8.4.6}$$

These linear combinations are called *scaling variables*. They have the important quality of transforming in a multiplicative way under the RG transformations

$$u_i^{'} = \sum_a \Delta_a^i \, \delta g_a^{'} \; = \; \sum_{a,b} \Delta_a^i \, \mathcal{K}_{ab} \, \delta g_b \qquad (8.4.7)$$
$$= \sum_b \lambda^i \, \Delta_b^i \, \delta g_b \; = \; \lambda^i \, u_i .$$

If b is the rescaling parameter of the block spins, it is common to parameterize λ_i as $\lambda^i = b^{y_i}$ where the quantities y_i are improperly called the eigenvalues of the renormalization group: in Section 8.9 we will show that they determine the critical exponents of the statistical model. Disregarding the case in which y_i is a complex number,[1] we can have the following cases:

1. $y_i > 0$. In this case the corresponding u_i is a *relevant variable*. A repeated application of the transformations moves its value away from the critical point.
2. $y_i < 0$. In this case u_i is an *irrelevant variable*. Starting sufficiently close to the fixed point, the iteration of the transformation shrinks the initial value to zero.
3. $y_i = 0$. In this case u_i is a *marginal variable*. Iterating the transformation, the value of this variable does not change.

Critical surface. To continue the analysis, let's assume that the dimension of the space of the coupling constants is m and let's consider a fixed point g^* with n relevant variables and $(m - n)$ irrelevant variables. This means that there exists a $(m - n)$-dimensional surface $\mathcal{C}$, called the *critical surface*, that is the attractive basin for the fixed point g^*. As shown below, on this surface the correlation length is infinite. The coupling constants g_k of the system depend generally on the external parameters of the system, such as temperature, pressure, or magnetic field. Varying these external parameters, the point $\{g\}$ of the coupling constants varies correspondingly. When there are n relevant variables, in order to intercept the critical surface it is necessary to choose appropriately n external control parameters. In all cases of physical interest, the temperature is one of these parameters and its value has to be tuned to its critical value $T = T_c$ to hit the critical surface. This may not be enough: if there are magnetic fields, they must be switched off and it may also be necessary to tune appropriately the chemical potential. Once such a fine tuning of the n experimental parameters has been done, the point $\{g\}$ in on the critical surface. If we now apply the RG transformations, their iterations of the RG move the point toward the critical point g^*, *independently of its initial position on* $\mathcal{C}$, as shown in Fig. 8.4. This is, in a nutshell, the origin of the universal behavior of the critical phenomena: hamiltonians that differ only for their irrelevant operators give rise to the same critical behavior.

Let's now prove that the correlation length diverges on the critical surface. Suppose that the physical system is represented by the point $\{g\}$ in the space of the coupling constants and, after n iterations, by $\{g^{(n)}\}$. Using eqn (8.3.8), we have the sequence

[1] In this case the trajectories are spirals that converge to the fixed point g^* if $\mathrm{Re}\, y_i < 0$ or diverge from it if $\mathrm{Re}\, y_i > 0$.

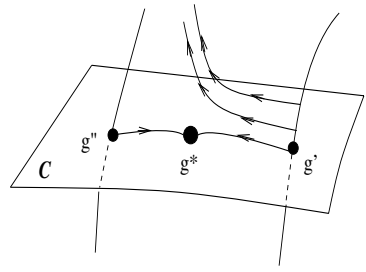


Fig. 8.4 *Once the trajectory reaches the critical surface, the evolution of the coupling constants under the renormalization group converges to the fixed point g^*, independently of the initial position $g^{'}$ or $g^{''}$. Points outside the critical surface move away from it.*

of identities

$$\xi(g) \,=\, b\,\xi(g^{(1)}) \,=\, b^2\,\xi(g^{(2)}) \,=\, \cdots \,=\, b^n\,\xi(g^{(n)}).$$

If the initial point $\{g\}$ was on the critical surface, in the limit $n \to \infty$ the sequence of $\{g^{(n)}\}$ converges to $\{g^*\}$, i.e. $\lim_{n\to}\{g^{(n)}\} \,=\, \{g^*\}$: since $\xi(g^*) = \infty$ and $b > 1$, we have that $\xi(g) \,=\, \infty$ for all points of the critical surface.

Properties of the RG flows. The physical nature of the problem is quite helpful in clarifying both the geometrical nature of the trajectories and some of their properties. For instance:

- The RG trajectories can only intersect at the fixed points.
- Switching on a relevant variable in a hamiltonian that is at a fixed point g_i^*, the corresponding flow moves the system away from it. At the end of this motion, the point $\{g\}$ reaches either a trivial fixed point (with a zero correlation length) or another critical point g_f^*. The approach to both final points is obviously along one of their irrelevant directions.
- During the motion along the RG flows, the point may pass close to other fixed points g_a^* $(a = 1, 2, \ldots)$, as shown in Fig. 8.5. If the trajectory is sufficiently close to them, there could be a series of interesting cross-over phenomena. According to the scale by which one monitors the system, one can observe the following behaviors: (i) over a short distance, the critical behavior ruled by the original fixed point g_i^*; (ii) on intermediate scales, the scaling behavior associate to the nearest fixed points met along the flow; (iii) at large distance, the scaling behavior ruled by the final fixed point g_f^*.

In order to clarify the concepts introduced so far, it is useful to discuss some simple examples.

8.5 The Ising Model

The first example is the one-dimensional Ising model. As the initial hamiltonian we take the one with the nearest neighbor interaction

$$H(s_i; J) \,=\, -J \sum_i s_i\, s_{i+1}. \tag{8.5.1}$$

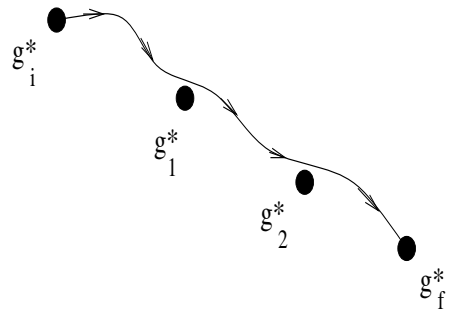


Fig. 8.5 *Renormalization group trajectory obtained by perturbing the hamiltonian of the fixed point g_i^* with a relevant variable. The final point corresponds to the hamiltonian of the new fixed point g_f^*, while the cross-over phenomena are ruled by the intermediate fixed points met along the trajectory.*

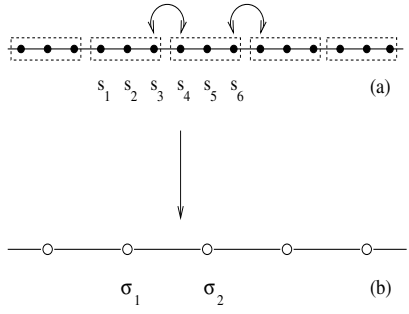


Fig. 8.6 *Spin blocks in the one-dimensional Ising model and the decimation transformation.*

Each pair of spins has the Boltzmann weight

$$W(s_i, s_{i+1}; v) \;=\; e^{J\, s_i\, s_{i+1}} \;=\; \cosh J\,(1 + v s_i\, s_{i+1}), \tag{8.5.2}$$

with $v = \tanh J$. To apply the RG transformations, we divide the system into blocks, each made of three spins, and then we apply the decimation rule: for each block we choose as a spin of the new system the one that is at the center, as shown in Fig. 8.6.

Consider two neighbor blocks. To implement the RG procedure, it is necessary to sum over the spins s_3 and s_4, keeping fixed, though, the values of the spins at the center of the two blocks, here denoted as $\sigma_1 \equiv s_2$ and $\sigma_2 = s_5$. In the partition function the terms that involve the degrees of freedom of two neighbor blocks are

$$e^{J\sigma_1\, s_3}\, e^{J s_3\, s_4}\, e^{J s_4\, \sigma_2}.$$

Using the identity $e^{J x_a x_b} = \cosh J\,(1 + v\, x_a\, x_b)$ for all the three terms of the previous equation, one has

$$(\cosh J)^3\,(1 + v\,\sigma_1\, s_3)\,(1 + v\, s_3\, s_4)\,(1 + v\, s_4\,\sigma_2).$$

Expanding this product and summing over s_3 and s_4, one gets

$$2^2(\cosh J)^3\,(1+v^3\sigma_1\sigma_2).$$

Beside a multiplicative normalization constant (independent of the spins), this expression is of the same form as (8.5.2) and therefore it defines the new Boltzmann weight $W(\sigma_1,\sigma_2;v')$ of the block spins σ_1 and σ_2 with

$$v' \;=\; v^3. \tag{8.5.3}$$

The new hamiltonian of the system is thus given by

$$H(\sigma_i;J') \;=\; N'\,p(J) - J'\sum_i \sigma_i\,\sigma_{i+1}, \tag{8.5.4}$$

where $N'=N/3$ is the number of sites of the new lattice, while the value of the new coupling constant is

$$J' \;=\; \tanh^{-1}\left[(\tanh J)^3\right], \tag{8.5.5}$$

$p(J)$ is the contribution to the free energy coming from the degrees of freedom on which we have summed, and it ensures the correct normalization of the partition functions of the two systems

$$p(J) \;=\; -\frac{1}{3}\log\left[\frac{(\cosh J)^3}{\cosh J'}\right] - \frac{2}{3}\,\log 2.$$

Let's now use the transformation law of the coupling constants, eqn (8.5.3), to study the physical content of the model. It is useful to make a plot of this mapping, as done in Fig. 8.7. It is easy to see that the mapping has two fixed points: $v_1^*=0$ and $v_2^*=1$. The first is an attractive fixed point, while the second is repulsive: unless v is exactly $v=1$, each iteration moves the values of v to the origin. Recall that we have absorbed in J a factor $\beta=1/kT$. This means that the high-temperature phase around $T\to\infty$ corresponds to the values close to $v\to 0$, while the low-temperature phase around $T\to 0$ corresponds to values $v\to 1$, with $v=1$ when $T=0$.

Since the effective coupling constant moves toward smaller values at each iteration, the large-scale degrees of freedom are described by an effective hamiltonian whose temperature increases: this is the region where the system is in its paramagnetic phase and has a finite correlation length. This happens for all values of v (except $v=1$) and therefore we are led to the conclusion that the one-dimensional Ising model is always in its disordered phase. As we have seen earlier, this conclusion is indeed confirmed by the exact solution of this model, discussed in Chapter 2. It is also easy to derive how the correlation length depends on the coupling constant: one simply needs to employ the transformation law

$$\xi(v') \;=\; \frac{1}{3}\,\xi(v), \tag{8.5.6}$$

and substitute v' with eqn (8.5.3). Hence, the correlation length satisfies the functional equation

$$\xi(v^3) \;=\; \frac{1}{3}\,\xi(v), \tag{8.5.7}$$

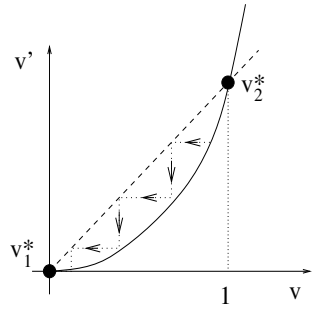


Fig. 8.7 *Renormalization group equation for the one-dimensional Ising model.* $v_1^* = 0$ *and* $v_2^* = 1$ *are the two fixed points, the former an attractive one, the latter a repulsive one. Starting from any value* $v \neq 1$*, the next iterations move the value of* v *toward the origin.*

whose solution is given by

$$\xi(v) \;=\; -\frac{\xi_0}{\log v} \;=\; -\frac{\xi_0}{\log \tanh J}. \tag{8.5.8}$$

This expression is in agreement with the behavior of $\xi(v)$ discussed in Chapter 2. Note that ξ is always finite, except when $J \to \infty$ ($T \to 0$), where it diverges as $\xi \simeq e^{1/T}$. This gives further evidence that the one-dimensional Ising model is always in a paramagnetic phase, expect when $T = 0$.

Even in the absence of simple analytic expressions, the arguments presented above help us to understand the phase diagram of the Ising model on higher dimensional lattices. Firstly, let's consider closely the one-dimensional case: if we refer to the spin variables of two neighbor blocks, in the limit $J \to \infty$ the equation that fixes the new coupling constant can be written as

$$J' \;\simeq\; J \,\langle s_3 \rangle_{\sigma_1=1} \,\langle s_4 \rangle_{\sigma_2=1}, \tag{8.5.9}$$

where $\langle s_3 \rangle_{\sigma_1=1}$ is the mean value of the spin at the edge of the block, with the condition that the spin in the middle of the block assumes value 1. Since these mean values are always less than 1 (except at $J = \infty$), one has $J' < J$ and therefore the low-temperature fixed point is always unstable.

However, for d-dimensional lattices (with $d > 1$), the situation is different. Consider once again the transformation law of the couplings in the limit $J \to \infty$. The value of the new coupling constants is essentially determined by the expectation values of the spins along the boundary of the blocks. Since there are b^{d-1} of them, we have

$$J' \;\simeq\; b^{d-1}\, J, \quad J \to \infty. \tag{8.5.10}$$

For $d > 1$, we have then $J' > J$, i.e. the low-temperature fixed point is now attractive! On the other hand, it is easy to convince oneself that the high-temperature fixed point is also attractive. The attractive nature of both fixed points implies that the Ising model in $d > 1$ should have a critical value at a finite value of the coupling constant, i.e. there should exist a critical temperature T_c at which the model undergoes a phase transition (see Fig. 8.8).

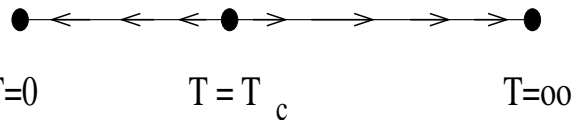


Fig. 8.8 *Phase diagram and renormalization group flows of the d-dimensional Ising model, with $d > 1$. In this case, both low and high temperature fixed points are attractive, while the fixed point between them is unstable with respect to the scaling variable associated to the temperature.*

8.6 The Gaussian Model

Another simple example of RG transformations is given by the gaussian model, whose variables $s_i = \varphi_i$ take values on all of the real axis. The hamiltonian of this model, expressed in the k-space, is expressed by

$$H \,=\, \frac{1}{2} \int_{|k|<1/a} (g_2 + k^2)\, |\varphi(k)|^2 \, d^d k. \tag{8.6.1}$$

The microscopic origin of the model is encoded in the cut-off $1/a$ present in the integration over the momenta. The partition function is

$$Z \,=\, \int \prod_{|k|<1/a} d\varphi(k)\, e^{-H}. \tag{8.6.2}$$

In order to implement the renormalization group procedure, let's integrate over the degrees of freedom of the field in the shell of the momenta $1/ba < |k| < 1/a$. This is equivalent to defining the new block spins, with scale parameter equal to b. The new hamiltonian is determined by the equation

$$e^{-H'} \,=\, \int \prod_{1/ba<|k|<1/a} d\varphi(k)\, e^{-H}. \tag{8.6.3}$$

Since each variable is decoupled from the others and each integral is gaussian, disregarding an inessential additive constant, the new hamiltonian is easily computed:

$$H' \,=\, \frac{1}{2} \int_{|k|<1/ba} (g_2 + k^2)\, |\varphi(k)|^2 \, d^d k. \tag{8.6.4}$$

To restore the initial lattice spacing, let's make the change of variable $k \to k/b$, so that

$$H' \,=\, \frac{1}{2} \int_{|k|<1/a} (g_2 + k^2\, b^{-2})\, b^{-d}\, |\varphi(k)|^2 \, d^d k. \tag{8.6.5}$$

We need to renormalize the new block variables. This can be done by requiring that the kinetic term $k^2\, |\varphi(k)|^2$ keeps the same form also in the new hamiltonian. Making the scale transformation on the field

$$\varphi' \,=\, b^{(d+2)/2}\, \varphi,$$

we arrive at

$$H' \,=\, \frac{1}{2}\int_{|k|<1/a} (g_2' + k^2)\,|\varphi'(k)|^2\, d^d k. \tag{8.6.6}$$

We have thus obtained the transformation law of the coupling constant

$$g_2' \,=\, b^2\, g_2 \tag{8.6.7}$$

which shows the relevant nature of this variable. The fixed points of the transformation law are: $g_2 = 0$, which is a critical point of the model, and $g_2 \to \infty$, which is a trivial fixed point. In fact, the correlation length is given by $\xi(r_0) = 1/\sqrt{g_2}$, and in the former case it diverges, while in the latter it goes to zero.

8.7 Operators and Quantum Field Theory

As discussed in the previous chapter, close to a critical point, where the correlation length ξ is much larger than the lattice spacing, it is natural to adopt the quantum field theory formalism to describe the dynamics of the statistical systems. For the sake of simplicity, let's focus our attention on a theory with a Z_2 internal symmetry. The order parameter is the scalar field $\varphi(x)$ that transforms as $\varphi \to -\varphi$ under the Z_2 symmetry. The more general action of such a model is given by

$$\mathcal{S} \,=\, \int d^d x \left[\frac{1}{2}(\partial_\mu \varphi)^2 + g_1 \varphi + \frac{g_2}{2}\varphi^2 + \cdots + \frac{g_n}{n!}\varphi^n + \cdots\right]. \tag{8.7.1}$$

The manifold of the coupling constants is described by the set $\{g\} = (g_1, g_2, \ldots, g_n, \ldots)$. The partition function of the system is expressed by the functional integral

$$Z[\{g\}, a] \,=\, \int \mathcal{D}\varphi\, \exp[-\mathcal{S}[\varphi, \{g\}]]. \tag{8.7.2}$$

In the previous chapter, we stressed the fundamental role played by the lattice spacing a even in the continuum theory. In terms of such a parameter, the "engineering" dimensions of the coupling constants

$$[g_n] \,=\, a^{nd/2-n-d} \,\equiv\, a^{\delta_n}, \tag{8.7.3}$$

can be regarded as their scaling dimensions with respect to the gaussian fixed point, identified by the condition $g_1 = g_2 = g_3 = \cdots = 0$. At the gaussian point, the relevant coupling constants are those with $\delta_n < 0$, the irrelevant ones those with $\delta_n > 0$, while the marginal ones are associated to the condition $\delta_n = 0$.

In addition to the gaussian fixed point, there may be other fixed points. They can be reached, for instance, by perturbing the gaussian action by some relevant operators, as shown in an explicit example discussed in Section 8.10. Let's then suppose that we are at a new fixed point, characterized by the action $\mathcal{S}^*$ and by a new set of m relevant

variables, denoted by λ_i $(i = 1, 2, \ldots, m)$. If $\phi_i(x)$ are the fields conjugated to these variables, in the vicinity of the new fixed point the action can be written as

$$\mathcal{S} \,=\, \mathcal{S}^* + \lambda_1 \int d^d x\, \phi_1(x) + \cdots + \lambda_m \int d^d x\, \phi_m(x). \tag{8.7.4}$$

Under a scaling transformation $x \to x/b$, λ_i scale as $b^{y_i}\, \lambda_i$ whereas the volume element $d^d x$ scales as $b^{-d}\, d^d x$. If we require that the action does not depend on the arbitrary parameter b, the fields $\phi_i(x)$ must then scale as[2]

$$\phi_i \to b^{x_i}\, \phi_i, \tag{8.7.5}$$

with

$$x_i \,=\, d - y_i. \tag{8.7.6}$$

It is easy to prove that, at the fixed point, the two-point correlation functions of the fields $\phi_i(x)$ have the scaling form

$$G_i(r_1 - r_2) \,=\, \langle \phi_i(r_1) \phi_i(r_2) \rangle \,=\, \frac{1}{|r_1 - r_2|^{2x_i}}. \tag{8.7.7}$$

Let's denote by $G_i(r_1 - r_2, \mathcal{S})$ the correlation function computed by the formula

$$\langle \phi_i(r_1) \phi_i(r_2) \rangle \,=\, \frac{1}{Z} \int \mathcal{D}\phi\, \phi_i(r_1)\, \phi_i(r_2)\, e^{-S}, \tag{8.7.8}$$

using the action $\mathcal{S}$. Making a RG transformation that changes the physical lengths by b, the action $\mathcal{S}$ becomes $\mathcal{S}'$, while the fields ϕ_i scale according to eqn (8.7.5). We arrive then at the functional equation

$$G_i((r_1 - r_2)/b, \mathcal{S}') \,=\, b^{2x_i}\, G_i(r_1 - r_2, \mathcal{S}). \tag{8.7.9}$$

If the initial action $\mathcal{S}$ is the one relative to the fixed point $\mathcal{S}^*$, we have $\mathcal{S}' = \mathcal{S}^*$ and therefore the equation above becomes

$$G_i((r_1 - r_2)/b, \mathcal{S}^*) \,=\, b^{2x_i}\, G_i(r_1 - r_2, \mathcal{S}^*). \tag{8.7.10}$$

The solution, except for a multiplicative factor, is provided by eqn (8.7.7).

Role of the microscopic scale. The expression (8.7.7) for the scaling form of the correlation function gives us the opportunity to make an important comment on the role played by the microscopic scale a. Suppose that the field $\phi_i(x)$ in (8.7.7) coincides with the field $\varphi(x)$ that appears in the action (7.2.3). Since the engineering dimension of this field is $a^{1-d/2}$, one could expect that its correlation function should take the form

$$G_i(r) \,=\, \frac{1}{r^{d-2}}. \tag{8.7.11}$$

If $2x_i \neq (d-2)$ obviously this expression does not coincide with the one given in eqn (8.7.7). This seems to preclude the possibility that the field $\varphi(x)$ could ever have

[2]The same argument also holds for the conjugate fields of the irrelevant variables, with their scaling dimension also given by (8.7.6). For simplicity we focus attention only on the relevant terms.

an anomalous dimension. There is, however, a way out: the scaling law (8.7.7) can be conciliated with the engineering dimension of the field if there exists a length-scale a able to absorb the canonical dimension of the field. In other words, the exact expression of the propagator of the scaling field is not that in eqn (8.7.7) but

$$G_i(r) \,=\, \frac{1}{a^{d-2}} \left(\frac{a}{r}\right)^{2x_i} . \tag{8.7.12}$$

If we make a scaling transformation of *all* the dimensional quantities of the problem, the correlator scales according to the canonical dimension of the field

$$G_i \,\to\, a^{-(d-2)}\, G_i .$$

Vice versa, if we would like to observe the system under a different magnifying glass, we shall rescale the length-scales but keep fixed the lattice spacing of the system: in this case, we obtain instead the anomalous behavior expressed by eqn (8.7.7). In conclusion, the expression (8.7.7) that is usually assumed as the propagator of the scaling fields has to be rather regarded as a shortening of the general formula (8.7.12), where the lattice spacing a was taken to be 1.

Note that the functional equation (8.7.9) can be empoyed to find the general form of the correlation functions of the scaling fields in the vicinity of the fixed point. Suppose, for simplicity, that $\mathcal{S}^*$ is perturbed by only one relevant field, with coupling constant λ_k. In this case, eqn (8.7.9) becomes

$$G_i(r/b, b^{y_k}\lambda_k) \,=\, b^{2x_i}\, G_i(r, \lambda_k), \tag{8.7.13}$$

whose general solution can be written as

$$G_i(r, \lambda_k) \,=\, \frac{1}{r^{2x_i}}\, f_i \left(\lambda^{1/y_k}\, r\right). \tag{8.7.14}$$

In this formula f_i is a homogeneous function of the distance r and of the coupling constant λ_k, whose explicit form can be obtained only by studying the details of the model.

8.8 Functional Form of the Free Energy

The linearized form of the renormalization group equations permits us to easily derive the scaling form of the free energy in the vicinity of the fixed point and the relationships between the critical exponents. Consider a statistical system with n relevant coupling constants λ_i. In the field theory formulation, in the vicinity of the fixed point the action is given by

$$\mathcal{S} \,=\, \mathcal{S}^* + \sum_i^n \lambda_i \int d^d x\, \phi_i(x). \tag{8.8.1}$$

In the Ising model, for instance, there are two relevant variables, given by the magnetic field $h \equiv \lambda_1$ and by the displacement of the temperature from the critical value $T - T_c \equiv \lambda_2$: the conjugate fields are $\phi_1(x)$, which correspond to the continuum limit of the spin

variable s_i, and $\phi_2(x)$, associated to the continuum limit of the energy density, given on the lattice by $\sum_j s_i\, s_{i+\hat{e}_j}$.

Since the variables λ_j in the action (8.8.1) have dimensions $[\lambda_j] = a^{y_j}$, the theory has a finite correlation length. Selecting one of the couplings, say λ_i, in the thermodynamic limit the correlation length can be expressed as

$$\xi(\{\lambda_j\}) \,=\, a\,(K_i\lambda_i)^{-\frac{1}{y_i}}\, L_i\left(\frac{K_1\lambda_j}{(K_i\lambda_i)^{\phi_{1i}}},\cdots,\frac{K_j\lambda_j}{(K_i\lambda_i)^{\phi_{ji}}},\cdots\right), \tag{8.8.2}$$

where $K_i \simeq 1/\lambda_i^{(0)}$ are some non-universal metric terms that depend on the units by which we measure the coupling constants, L_i are universal homogeneous functions of the $(n-1)$ ratios $K_j\lambda_j/(K_i\lambda_i)^{\phi_{ji}}$, with $j \neq i$, and finally

$$\phi_{ji} \,=\, \frac{y_j}{y_i}, \tag{8.8.3}$$

are the so-called *cross-over exponents*. There are many (but equivalent) ways of expressing this scaling law of the correlation length, according to which coupling constant we choose as prefactor. Each way selects its own scaling function L of the above ratio of the couplings. When $\lambda_k \to 0$ $(k \neq i)$ with $\lambda_i \neq 0$, eqn (8.8.2) can be written as

$$\xi_i \,=\, a\,\xi_i^0\,\lambda_i^{-\frac{1}{y_i}}, \quad \xi_i^0 \sim K_i^{-\frac{1}{y_i}}. \tag{8.8.4}$$

Consider now the free energy of the system, $f[\lambda_i]$, defined by

$$Z[\{\lambda_i\}] \,=\, \int \mathcal{D}\phi_i e^{-\left[\mathcal{S}^* + \sum_{i=1}^n \lambda_i \int \phi_i(x)\, d^d x\right]} \,\equiv e^{-N\, f(\lambda_i)}. \tag{8.8.5}$$

By virtue of the identity (8.3.4) of the partition functions, making a RG transformation we have

$$e^{-N\, f(\{\lambda\})} \,=\, e^{-Np(\{\lambda\}) - N'\, f(\{\lambda'\})},$$

where $p(\{\lambda\})$ is an additive constant related to the degrees of freedom on which we have integrated. Since the new number of sites is $N' = b^{-d}N$, we have the functional equation

$$f(\{\lambda\}) \,=\, p(\{\lambda\}) + b^{-d}\, f(\{\lambda'\}). \tag{8.8.6}$$

The function $p(\{\lambda\})$ is an analytic function of the coupling constant, since it involves a sum over a finite number of spins. If we are interested in studying the singular behavior of the free energy, we can safely discard this term and arrive at a functional equation that involves only the singular part of f

$$f_s(\{\lambda\}) \,=\, b^{-d}\, f_s(\{\lambda'\}). \tag{8.8.7}$$

Substituting in it the expression for the new coupling constants given by the RG transformations, we have

$$f_s(\{\lambda_k\}) \,=\, b^{-d}\, f_s(\{b^{y_k}\,\lambda_i\}). \tag{8.8.8}$$

Iterating this equation, the irrelevant variables go to zero (this is a manifestation of the universality of the critical behavior) and the free energy, as a function of the relevant variables alone, satisfies

$$f_s(\{\lambda_j\}) \;=\; b^{-nd}\, f_s(\{b^{ny_j}\,\lambda_i)\}. \tag{8.8.9}$$

As for the correlation length, there are many ways to express the general solution of this equation. Selecting once more one of the couplings, say λ_i, we have

$$f(\{\lambda_i\}) = f_i[\{\lambda_j\}] \equiv (K_i\lambda_i)^{-\frac{d}{y_i}}\; F_i\left(\frac{K_1\lambda_j}{(K_i\lambda_i)^{\phi_{1i}}},\cdots,\frac{K_j\lambda_j}{(K_i\lambda_i)^{\phi_{ji}}},\cdots\right)^{-\frac{d}{y_i}}. \tag{8.8.10}$$

The functions F_i are universal homogeneous functions of the $(n-1)$ ratios $K_j g_j/(K_i g_i)^{\phi_{ji}}$. As we will see below, there are some obvious advantages in considering different expressions for these scaling functions, obtained by changing the selecting variable λ_i. In fact, in several physical applications, there is only one coupling constant kept different from zero till the end, and the best choice of expressing the free energy depends on this situation. As we are going to show, even in the absence of an explicit expression of the F_i's (which can be explicitly found only by solving exactly the model by other methods), the functional dependence of the free energy is sufficient to obtain useful information on the critical behavior of the model.

8.9 Critical Exponents and Universal Ratios

Let's discuss the definition of several thermodynamical quantities associated to the derivates of the free energy. In the following, we adopt the notation $\langle ...\rangle_i$ to denote the expectation values computed with an action that has, at the end, only λ_i as coupling constant different from zero. The first quantities of interest are the expectation values of the fields ϕ_j that can be parameterized as

$$\langle\phi_j\rangle_i \;=\; -\frac{\partial f_i}{\partial\lambda_j}\bigg|_{\lambda_k=0} \equiv B_{ji}\,\lambda_i^{\frac{d-y_j}{y_i}}\;, \tag{8.9.1}$$

with

$$B_{ji} \sim\; K_j K_i^{\frac{d-y_j}{y_i}}. \tag{8.9.2}$$

Equivalently

$$\lambda_i \;=\; D_{ij}\,(\langle\phi_j\rangle_i)^{\frac{y_i}{d-y_j}}, \tag{8.9.3}$$

with

$$D_{ij} \;\sim\; \frac{1}{K_i K_j^{\frac{y_i}{d-y_j}}}. \tag{8.9.4}$$

The generalized susceptibilities are defined by

$$\hat\Gamma^i_{jk} = \frac{\partial}{\partial\lambda_k}\langle\phi_j\rangle_i \;=\; -\frac{\partial^2 f_i}{\partial\lambda_k\partial\lambda_j}. \tag{8.9.5}$$

These quantities are obviously symmetric with respect to the lower indices. By the fluctuation-dissipation theorem, they are related to the off-critical correlation functions as

$$\hat{\Gamma}^i_{jk} \;=\; \int dx\, \langle \phi_k(x)\phi_j(0)\rangle_i. \tag{8.9.6}$$

Taking out the dependence on the coupling constant λ_i, we have

$$\hat{\Gamma}^i_{jk} = \Gamma^i_{jk}\, \lambda_i^{\frac{d-y_j-y_k}{y_i}}, \tag{8.9.7}$$

with

$$\Gamma^i_{jk} \;\sim\; K_j K_k K_i^{\frac{d-y_j-y_k}{y_i}}. \tag{8.9.8}$$

As shown by the formulas above, the various quantities contain the metric factors K_i and their expressions are therefore not universal. However, we can consider some special combinations of these quantities in which the metric factors are cancelled out. Here we give some examples of the so-called *universal ratios*

$$(R_c)^i_{jk} \;=\; \frac{\Gamma^i_{ii}\Gamma^i_{jk}}{B_{ji}B_{ki}}; \tag{8.9.9}$$

$$(R_\chi)^i_j \;=\; \Gamma^i_{jj} D_{jj} B_{ji}^{\frac{D-4\Delta_j}{2\Delta_j}}; \tag{8.9.10}$$

$$R^i_\xi \;=\; \left(\Gamma^i_{ii}\right)^{1/D} \xi^0_i; \tag{8.9.11}$$

$$(R_A)^i_j \;=\; \Gamma^i_{jj}\, D_{ii}^{\frac{4\Delta_j+2\Delta_i-2D}{D-2\Delta_i}}\, B_{ij}^{\frac{2\Delta_j-D}{\Delta_i}}; \tag{8.9.12}$$

$$(Q_2)^i_{jk} \;=\; \frac{\Gamma^i_{jj}}{\Gamma^k_{jj}} \left(\frac{\xi^0_k}{\xi^0_j}\right)^{D-4\Delta_j}. \tag{8.9.13}$$

As the critical exponents, these pure numbers characterize the universality class of a given model. It is worth emphasizing that, from an experimental point of view, it should be simpler to measure universal amplitude ratios rather than critical exponents: in fact to determine the former quantities one needs to perform several measurements at a single, fixed value of the coupling which drives the system away from criticality whereas to determine the latter, one needs to make measurements over several decades along the axes of the off-critical couplings. Moreover, although not all of them are independent, the universal ratios are a set of numbers larger than the critical exponents and therefore permit a more precise determination of the class of universality. Finally, being universal quantities, they can be theoretically computed by analyzing the simplest representative of the class of universality under scrutiny.

Some of the quantities above have a familiar meaning in the context of the Ising model. In particular, they permit us to express all critical exponents in terms of rational functions of the eigenvalues y_i of the renormalization group. To compare with the

formulas of Chapter 1, it is convenient to use the notation $h = \lambda_1$ and $t = \lambda_2$. Consider, for instance, the specific heat

$$C(T) = \frac{\partial^2 f}{\partial t^2}\Big|_{h=0} = \begin{cases} C_+ \, t^{d/y_t - 2} & , \quad T > T_c \\ C_- (-t)^{d/y_t - 2} & , \quad T < T_c. \end{cases} \tag{8.9.14}$$

From the definition of the critical exponent α, $C(T) \simeq |t|^{-\alpha}$, we get

$$\alpha = 2 - d/y_t. \tag{8.9.15}$$

The spontaneous magnetization of the system is obtained by

$$M(T) = \frac{\partial f}{\partial h}\Big|_{h=0} = M_0 \, (-t)^{(d-y_h)/y_t}. \tag{8.9.16}$$

Comparing with the definition of the critical exponent β, $M(T) \simeq (-t)^{\beta}$, one has

$$\beta = \frac{d - y_h}{y_t}. \tag{8.9.17}$$

The susceptibility at zero magnetic field is given by

$$\chi(T) = \frac{\partial^2 f}{\partial h^2}\Big|_{h=0} = \begin{cases} \chi_+ t^{(d-2y_h)/y_t} & , \quad T > T_c \\ \chi_- (-t)^{(d-2y_h)/y_t} & , \quad T < T_c. \end{cases} \tag{8.9.18}$$

Comparing with the definition of the critical exponent γ, $\chi(T) \simeq |t|^{-\gamma}$, one obtains

$$\gamma = \frac{d - 2y_h}{y_t}. \tag{8.9.19}$$

In order to derive finally the exponent δ, one needs to consider the general expression of the magnetization

$$M(T,h) = \frac{\partial f}{\partial h} = \frac{1}{h_0} \left| \frac{t}{t_0} \right|^{(d-y_h)/y_t} \mathcal{F}' \left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}} \right), \tag{8.9.20}$$

with the choice of the scaling form of the free energy given by (t_0 and h_0 are the metric factors)

$$f(t,h) = \left| \frac{t}{t_0} \right|^{d/y_t} \mathcal{F} \left(\frac{h/h_0}{|t/t_0|^{y_h/y_t}} \right). \tag{8.9.21}$$

To have a finite limit of this expression when $|t| \to 0$, it is necessary that $\mathcal{F}'(x)$ behaves as $\mathcal{F}'_0 \, x^{d/y_h - 1}$ when $x \to \infty$, so that

$$M(T=0,h) \, \mathcal{M}_0 \; h^{(d-y_B)/y_t}. \tag{8.9.22}$$

Comparing with the definition of the exponent δ, $M(h) \simeq h^{1/\delta}$, one has

$$\delta = \frac{y_h}{d - y_h}. \tag{8.9.23}$$

Hence, all critical exponents are expressed in terms of the eigenvalues y_h and y_t of the renormalization group equations of this model. Moreover, there is a natural explanation of the existence of the scaling laws

$$\begin{aligned} \alpha + 2\beta + \gamma &= 2, \\ \alpha + \beta\,\delta + \beta &= 2. \end{aligned} \tag{8.9.24}$$

Using the scaling form (8.7.14) of the correlation function, with $\lambda_k = t$, we have

$$\xi = \begin{cases} \xi_+\, t^{-1/y_t}, & T > T_c \\ \xi_-\, (-t)^{-1/y_t}, & T < T_c \end{cases} \tag{8.9.25}$$

and comparing with the definition of $\xi \simeq |t|^{-\nu}$, one arrives at

$$\nu = 1/y_t. \tag{8.9.26}$$

Comparing, instead, with the critical expression of the correlator $G_i(r) = 1/r^{d-2+\eta}$, we can extract the last critical exponent

$$\eta = d + 2 - 2\,y_h. \tag{8.9.27}$$

In terms of these expressions one can easily check the validity of the other scaling laws

$$\begin{aligned} \alpha + d\nu &= 2 \\ \gamma &= \nu(2-\eta). \end{aligned} \tag{8.9.28}$$

Concerning the universal ratios of the Ising model, presently their analytic expressions are only available for the two-dimensional case.[3] They have been computed by using quantum field theory methods away from the critical point – a subject that will be discussed in the following chapters. Here we simply report the exact values of some of these quantities[4]

$$\begin{aligned} C_+/C_- &= 1, \\ \chi_+/\chi_- &= 37.6936..., \\ \xi_+/\xi_- &= 2, \\ C_+\,\chi_+/M_0^2 &= 0.318569..., \\ C_+^{1/2}\,\xi_+ &= 1/\sqrt{2\pi}, \\ \chi_+\,\mathcal{M}_0^{-\delta}\,M_0^{\delta-1} &= 6.7782\ldots. \end{aligned} \tag{8.9.29}$$

8.10 β-functions

In the previous sections we have seen that to a scaling transformation of the lengths there corresponds a change of the coupling constants that leaves the physical content

[3]The one-dimensional model has only the paramagnetic disorder phase and therefore there is only a subset of the previous defined universal ratios.

[4]See G. Delfino, *Integrable field theory and critical phenomena: the Ising model in a magnetic field*, J. Phys. A 37 (2004), R 45.

of a theory unchanged. In a continuum formulation of critical phenomena, we can make a scaling transformation $x \to x' = x/b$, where b is an arbitrary real parameter. If b is infinitesimally close to 1, $b \simeq 1 + \delta l$, the coupling constants change infinitesimally as well

$$g_a \to g'_a = g_a + \frac{dg_a}{dl}\,\delta l + \mathcal{O}((\delta l)^2), \tag{8.10.1}$$

and the RG transformation become the differential equation

$$\frac{dg_a}{dl} = \beta_a(\{g\}). \tag{8.10.2}$$

We have thus introduced the $\beta_a(\{g\})$-functions. They are the vector fields that fix the RG flow. These functions are fundamental quantities of the theory for the following reasons. First of all, their zeros identify the fixed point of the theory, since these are the points where the coupling constants do not vary. Secondly, their derivatives, computed at the fixed points, are directly related to the eigenvalues y_i of the renormalization group: in fact, at the fixed points, the linearized matrix of the renormalization group is given by $\mathcal{K}_{ab} = \delta_{ab} + \frac{\partial \beta_a}{\partial g_b}\,\delta l$, with eigenvalues $(1 + \delta l)^{y_i} \simeq 1 + y_i\,\delta l$. Hence, the y_i's are nothing else but the derivatives of the β_a-functions at the fixed points g^*.

In most cases, the $\beta_a(\{g\})$-functions are only known perturbatively as series expansions in the coupling constants. If g_a corresponds to a scaling variable with eigenvalue y_a with respect to the fixed point $g^* = (0, 0, 0, \ldots)$ (here chosen as the origin in the manifold of the couplings), the first term of their expansion is simply

$$\beta_a(\{g\}) = y_a\, g_a + \cdots \tag{8.10.3}$$

Solving at this order the differential equation (8.10.2), one easily finds the result previously discussed: the coupling constants with $y_a > 0$ grow and get away from the origin, those with $y_a < 0$ shrink to zero.

It is obviously interesting to evaluate the higher orders of the β-functions. It should be said that this computation becomes more and more difficult by increasing the orders of the perturbation theory and it is far beyond the scope of this book. However, as we will show in detail in Chapter 15, there is a particularly elegant result for the quadratic terms. At this order, in fact, the β_a-functions are expressed as

$$\beta_a(\{g\}) = y_a\, g_a - \sum_{b,c} C_{abc}\, g_b\, g_c + \cdots \tag{8.10.4}$$

where C_{abc} are the same coefficients that enter the three-point correlation function of the scaling fields $\phi_a i$ (conjugate to the variables g_i) at the critical point. Thanks to the scaling properties of the fields ϕ_i, the general form of these correlators is

$$\langle \phi_a(r_1)\phi_b(r_2)\phi_c(r_3)\rangle = C_{abc} r_{12}^{x_c - x_a - x_b}\, r_{13}^{x_b - x_a - x_c}\, r_{23}^{x_a - x_b - x_c}, \tag{8.10.5}$$

where $r_{ij} = |r_i - r_j|$ and $x_a = d - y_a$.

ϵ-expansion. Particularly significant is the case in which the new zero of the β_a-functions (8.10.4) is close to the origin. In this circumstance, one can deduce the

universal properties of the new fixed point in terms of the fixed point at the origin. This happens, for instance, in the lagrangian theory

$$\mathcal{S} = \int d^d x \left[\frac{1}{2} (\partial_\mu \Phi)^2 + g_2 \Phi^2 + g_4 \Phi^4 \right],$$

where $\Phi(x) = [\phi_1(x), \phi_2(x), \ldots, \phi_n(x)]$ is a field with n components. For this model, the lowest orders of the two β-functions are given by

$$\begin{aligned} \beta_2(\{g\}) &= \frac{dg_2}{dl} = 2g_2 - 8(n+2) g_2 \, g_4 + \cdots \\ \beta_4(\{g\}) &= \frac{dg_4}{dl} = (4-d) g_4 - 8(n+8) \, g_2^2 + \cdots \end{aligned} \tag{8.10.6}$$

It is easy to see that, in addition to the origin, there is a new fixed point at $g^* = (g_2, g_4) = (0, (4-d)/8(n+8)$. This has a perturbative nature (i.e. compatible with the perturbative approach used to computed the β-functions) if $\epsilon \equiv (4-d) \ll 1$. Computing the derivatives of the β_a at this new fixed point, we can extract the relative critical exponents

$$\begin{aligned} \alpha &= \frac{4-n}{2(n+8)} \, \epsilon + \cdots \\ \beta &= \frac{1}{2} - \frac{3}{2} \, \frac{\epsilon}{n+8} + \cdots \\ \gamma &= 1 + \frac{n+2}{2(n+8)} \, \epsilon \\ \delta &= 3 + \epsilon \\ \nu &= \frac{1}{2} + \frac{n+2}{4(n+8)} \, \epsilon. \end{aligned} \tag{8.10.7}$$

The first significant feature of these expressions is their difference with respect to the classical values obtained by the mean field theory discussed in Chapter 3. The second relevant feature is the analytic nature of the parameter n, which allows us to compute the critical exponents for any value of this parameter, not necessarily an integer. Note that for $n = 1$ we obtain a Z_2 invariant theory, while for $n \to \infty$ we get the spherical model. The third aspect that is worth a comment is the nature of these series expansions: it is generally believed that they are only asymptotic series, therefore with zero radius of convergence. There are, however, efficient techniques to cure the problem and sum this kind of series (such as the Borel resummation method). Note that, substituting in the formulas above $n = 1$ and $\epsilon = 1$, one gets an estimate of the critical exponents of the three-dimensional Ising model. We leave as an exercise the comparison of these values with those obtained by the numerical simulations that were reported in Chapter 3.

References and Further Reading

The version of the renormalization group discussed in this chapter was proposed by Kenneth Wilson. For further details, we refer to the review articles:

K. Wilson, *The renormalization group: Critical phenomena and the Kondo problem*, Rev. Mod. Phys. 47 (1975), 773.

K. Wilson, J. Kogut, *The renormalization group and the epsilon expansion*, Phys. Rept. 12 (1974), 75.

For applications in field theory and statistical mechanics, it is useful to read:

D. Amit, *Field Theory, the Renormalization Group and Critical Phenomena*, Mc-Graw Hill, New York, 1978.

N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, Frontiers in Physics, Vol 85, 1992.

J.L. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge University Press, Cambridge, 1996.

P. Pfeuty, G. Toulouse, *Introduction to the Renormalization Group and to Critical Phenomena*, John Wiley, New York, 1977.

Problems

1. Logistic map

Consider the map

$$x_{n+1} = f_r(x_n) = r\, x_n(1 - x_n).$$

This was introduced by P.F. Verhulst in 1845 to model the growth of a population in a region of a finite area: the population at time $n+1$, expressed by x_{n+1}, is proportional to the population x_n at time n, but at the same time is also proportional to the remaining area. This quantity is decreased proportional to x_n, namely $(1-x_n)$. Despite the innocent aspect, this map has a remarkable mathematical structure.

a Prove that for $r < 1$ the map has a unique (stable) fixed point at the origin.

b Prove that for $1 < r < 3$, there exist two fixed points, at $x = 0$ and $x = 1 - 1/r$, the first of them unstable, while the second is a stable one.

c Prove that also the second fixed point becomes unstable when $r > 3$. Show that for $3 < r < r_1$ there are two stable fixed points x_1 and x_2 of the map $f_r[f_r[x]] \equiv f_r^{(2)}(x)$, i.e. they are solutions of the equations $x_2 = f_r(x_1)$ and $x_1 = f_r(x_2)$.

d Show that there is a value $r_2 > r_1$ when the two previous fixed points x_1 and x_2 become unstable. More generally, prove that there exists a sequence of values r_n

such that for $r_{n-1} < r < r_n$ there is a set of 2^{n-1} points characterized by the conditions

$$f_r(x_i^*) = x_{i+1}^*, \qquad f_r^{(2^{n-1})}(x_i^*) = x_i^*.$$

e Define the family of functions $g_i(x)$ by the limit

$$g_i(x) \;=\; \lim_{n\to\infty} (-\alpha)^n \, f_{r_{n+i}}^{(2^n)} \left[\frac{x}{(-\alpha)^n} \right].$$

Show that they satisfy the functional equation

$$g_{i-1}(x) \;=\; (-\alpha)\, g_i \left[g_i \left(-\frac{x}{\alpha} \right) \right] \;\equiv\; T\, g_i(x).$$

Study the features of the function $g(x)$, defined as the "fixed point" of the transformation law T

$$g(x) \;=\; T\, g(x) \;=\; -\alpha\, g \left[g \left(-\frac{x}{\alpha} \right) \right].$$

Prove, in particular, that α is a universal parameter.

2. Universal ratios

Consider the mean field solution of the Ising model and, with the notation used in Chapter 1, compute the universal ratios

$$C_+ \, \chi_+ / M_0^2, \qquad \chi_+ \, M_0^{\delta-1} \, \mathcal{M}_0^{-\delta}.$$

3. Approximated values of the critical exponents

Use the formulas (8.10.7) to obtain an approximate values of the three-dimensional Ising model and the spherical model. Compare with the numerical values obtained for the three-dimensional Ising model and with the exact expressions of the spherical model.

4. β-functions

Consider a statistical system with the space of the coupling constants described by the variables (x, y^2). The fixed point is identified by the origin $(0, 0)$. Suppose that the β-functions of these coupling constants are given by

$$\begin{aligned} \frac{dx}{db} \;&=\; -y^2, \\ \frac{dy^2}{db} \;&=\; -2\, x\, y^2. \end{aligned}$$

Study the renormalization group flows, with initial conditions (x_0, y_0^2), and show that they are hyperbolas in the plane (x, y). Identify the nature of the coupling constants, i.e. if they are relevant, irrelevant, or marginal.

9
Fermionic Formulation of the Ising Model

There are different kinds of scientists, such as second or third rank physicists who try their best but do not get too far. There are also first class scientists, who make discoveries of great importance. But then there is the genius. Majorana was one of them. He had what nobody else in the world has, unfortunately he was lacking the most natural quality, simple common sense.

Enrico Fermi

9.1 Introduction

In this chapter we will study the continuum limit formulation of the two-dimensional Ising model, starting from the hamiltonian limit of its transfer matrix. We will first derive the quantum hamiltonian of the model and then we will study its most important properties, such as the duality transformation. This symmetry involves the order and disorder operators and we will clarify their physical interpretation. Afterwards, we will see how to diagonalize the quantum hamiltonian by means of particular fermionic fields. The operator mapping between the order/disorder operators and the fermionic fields is realized by the so-called *Wigner–Jordan* transformation: this brings the original hamiltonian to a quadratic form in the creation and annihilation operators of the fermions. The determination of the spectrum is then obtained by a Bogoliubov transformation, a technique extremely useful also in other contexts, such as superconductivity phenomena. In the limit in which the lattice spacing goes to zero, the Ising model becomes a theory of free Majorana fermions. They satisfy a relativistic dispersion relation and their mass is a direct measurement of the displacement of the temperature from the critical value T_c.

It is important to stress that the fermionic formulation of the two-dimensional Ising model is crucial for the understanding of many of its physical properties and for the computation of its correlation functions. This formulation will be used in other parts of the book to illustrate several other aspects of this model. Given the importance of this subject, in the final section of this chapter we present another approach to show the fermionic content of this model and to derive the Dirac equation satisfied by the Majorana fermion.

9.2 Transfer Matrix and Hamiltonian Limit

In this section we consider the transfer matrix on a square lattice with the standard orientation of the lattice. Consider a square lattice with $N = n^2$ spins, made of n rows and n columns. The lattice spacing along the vertical and horizontal directions are τ and α, respectively. The spins, here denoted by $\sigma_{i,j}$, satisfy periodic boundary conditions

$$\sigma_{i+n,j} = \sigma_{i,j}, \quad \sigma_{i,j+n} = \sigma_{ij}.$$

Below we will also use the notation σ_i and σ'_i to denote spins of next neighbor rows, where the index i labels in this case the position of the spins along these rows. Denoting with μ_a $(a = 1, 2, \dots n)$ the set of all spins that belong to the row a

$$\mu_a = \{\sigma_1, \sigma_2, \dots \sigma_n\}_{a\text{-}row},$$

a configuration of the system is specified by the ensemble $\{\mu_1, \dots \mu_n\}$. The a-th row interacts only with the next neighbor rows, namely μ_{a-1} and μ_{a+1}. Let $E(\mu_a, \mu_{a+1})$ be the interaction energy between two next neighbor rows and $E(\mu_a)$ the energy coming from the interactions of the spins placed on the a-th row, eventually also subjected to an external magnetic field B. Assuming the usual hamiltonian of the model, we have

$$\begin{aligned} E(\mu, \mu') &= -J' \sum_{k=1}^{n} \sigma_k \sigma'_k, \\ E(\mu) &= -J \sum_{k=1}^{n} \sigma_k \sigma_{k+1} - B \sum_{k=1}^{n} \sigma_k, \end{aligned}$$

where J' and J are the couplings along the vertical and horizontal directions, respectively (see Fig. 9.1). The total energy of a configuration of the system is then

$$E(\mu_1, \dots \mu_n) = \sum_{a=1}^{n} \left[E(\mu_a, \mu_{a+1}) + E(\mu_a) \right],$$

and its partition function is given by

$$Z = \sum_{\mu_1} \sum_{\mu_2} \dots \sum_{\mu_n} \exp\left[-\beta E(\mu_1, \dots \mu_n) \right]. \tag{9.2.1}$$

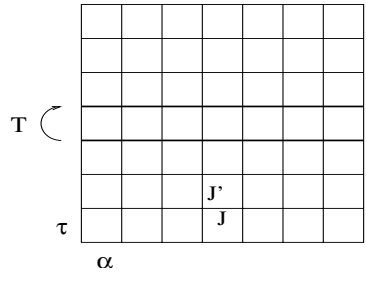


Fig. 9.1 *Lattice parameters and transfer matrix.*

Let's introduce the transfer matrix T. It is a $2^n \times 2^n$ matrix, with elements given by

$$\langle \mu \mid T \mid \mu' \rangle = \exp\left[-\beta\left(E(\mu,\mu') + E(\mu)\right)\right]. \tag{9.2.2}$$

In terms of T, the partition function is expressed as

$$\begin{aligned} Z &= \sum_{\mu_1}\sum_{\mu_2}\dots\sum_{\mu_n}\langle \mu_1 \mid T \mid \mu_2\rangle\langle \mu_2 \mid T \mid \mu_3\rangle\dots\langle \mu_n \mid T \mid \mu_1\rangle \\ &= \sum_{\mu_1}\langle \mu_1 \mid T^n \mid \mu_1\rangle = \mathrm{Tr}\, T^n. \end{aligned} \tag{9.2.3}$$

The operator T can be further decomposed in terms of three operators

$$T = V_3\, V_2\, V_1,$$

where the V_i are $2^n \times 2^n$ matrices whose elements are given by

$$\langle \sigma_1\dots\sigma_n \mid V_1 \mid \sigma'_1\dots\sigma'_n\rangle = \prod_{k=1}^{n} e^{L\,\sigma_k\,\sigma'_k}, \tag{9.2.4}$$

$$\langle \sigma_1\dots\sigma_n \mid V_2 \mid \sigma'_1\dots\sigma'_n\rangle = \delta_{\sigma_1\sigma'_1}\dots\delta_{\sigma_n\sigma'_n}\prod_{k=1}^{n} e^{K\,\sigma_k\,\sigma_{k+1}}, \tag{9.2.5}$$

$$\langle \sigma_1\dots\sigma_n \mid V_1 \mid \sigma'_1\dots\sigma'_n\rangle = \delta_{\sigma_1\sigma'_1}\dots\delta_{\sigma_n\sigma'_n}\prod_{k=1}^{n} e^{\beta\,B\,\sigma_k}. \tag{9.2.6}$$

In these formulas we have introduced the notation $K = \beta J$ and $L = \beta J'$. To have a more convenient expressions, let's introduce the operators

$$\tilde{\sigma}_1(a) = \mathbf{1}\times\mathbf{1}\times\dots\times\overbrace{\sigma_1}^{a}\times\mathbf{1}\dots\times\mathbf{1} \tag{9.2.7}$$

$$\tilde{\sigma}_2(a) = \mathbf{1}\times\mathbf{1}\times\dots\times\overbrace{\sigma_2}^{a}\times\mathbf{1}\dots\times\mathbf{1} \tag{9.2.8}$$

$$\tilde{\sigma}_3(a) = \mathbf{1}\times\mathbf{1}\times\dots\times\overbrace{\sigma_3}^{a}\times\mathbf{1}\dots\times\mathbf{1} \tag{9.2.9}$$

They are defined by the direct product of 2×2 matrices, where the σ_i are the usual Pauli matrices, whereas $\mathbf{1}$ is the unit matrix. For $a \neq b$ it is easy to see that these operators commute with each other:

$$[\tilde{\sigma}_i(a), \tilde{\sigma}_j(b)] = 0. \tag{9.2.10}$$

When $a = b$, they satisfy instead the commutation and anticommutation relations of the Pauli matrices

$$[\tilde{\sigma}_i(a), \tilde{\sigma}_j(a)] = 2i\,\epsilon_{ijk}\tilde{\sigma}_k(a), \tag{9.2.11}$$

$$\{\tilde{\sigma}_i(a), \tilde{\sigma}_j(a)\} = 2\delta_{ij}. \tag{9.2.12}$$

In terms of the $\tilde{\sigma}_i(a)$'s, the transfer matrix P can be put in an operatorial form as follows

$$T = \prod_{a=1}^{n} \left[e^{\beta B \tilde{\sigma}_3(a)} \; e^{K \tilde{\sigma}_3(a) \tilde{\sigma}_3(a+1)} \; e^{L \tilde{\sigma}_1(a)} \right]. \tag{9.2.13}$$

As discussed in Chapters 2 and 7, we can associate to a transfer matrix T of a classical statistical system in d dimensions a quantum hamiltonian H in $(d-1)$ dimensions. In our case we have

$$T \equiv e^{-\tau H}, \tag{9.2.14}$$

where τ is the lattice spacing along the vertical direction of the lattice and H is the one-dimensional quantum hamiltonian. To obtain an explicit expression for H it is necessary to use the Baker–Campbell–Hausdorf formula for the exponential of two non-commuting operators

$$e^A \, e^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}([A,B],B]+[A,[A,B]])+\cdots}.$$

Its expression, for a finite value of τ, is neither convenient nor particularly illuminating. To gain a better insight, it is useful to consider the so-called hamiltonian limit, i.e. the situation that arises when $\tau \to 0$. For simplicity, we will deal below only with the case when the magnetic field is absent, $B = 0$.

Matrix elements. In taking the hamiltonian limit, we shall be careful that the physical content of the system does not change: from the renormalization group analysis we know that this can be achieved by rescaling appropriately the coupling constants (see Fig. 9.2). To determine their dependence on the lattice spacing and, correspondingly, the expression of H that emerges in this limit, we can proceed as follows. For $\tau \to 0$, expanding the expression (9.2.14), we have

$$T \simeq 1 - \tau H. \tag{9.2.15}$$

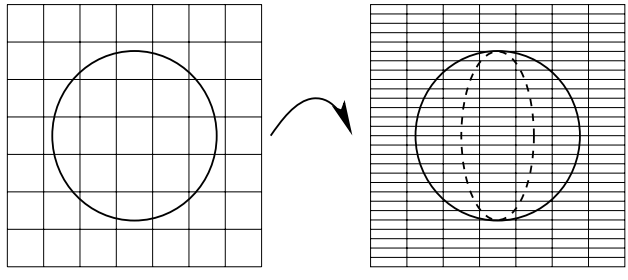


Fig. 9.2 *Hamiltonian limit. The circle in the lattice on the right is the set of points in which the correlation function* $\langle \sigma(r)\sigma(0) \rangle$ *is constant. If in the new lattice the coupling constants were not rescaled, the circle becomes an ellipse. Only an appropriate rescaling of the couplings leaves invariant the physical content of the original model.*

Let's now consider explicitly some matrix elements of the operator T. If there are non-spin flips going from the a-row to the $(a+1)$-row, we have

$$T(0\,\text{spin}-\text{flips}) \;=\; \exp\left[K\sum_i \sigma_i\,\sigma_{i+1}\right] \;=\; 1 + K\sum_i \sigma_i\,\sigma_{i+1} + \cdots \quad (9.2.16)$$
$$\simeq\; 1 - \tau\, H_{0\,\text{spin}-\text{flips}}.$$

When there is only one spin flip in going from a row to the next one, we have

$$T(1\,\text{spin}-\text{flip}) \;=\; \exp(-2\,L)\,\exp\left\{\frac{1}{2}\sum_i \left[\sigma_i\sigma_{i+1} + \sigma'_i\sigma'_{i+1}\right]\right\} \quad (9.2.17)$$
$$\simeq\; -\tau\, H_{1\,\text{spin}-\text{flip}},$$

and, finally, when there are k spin flips the matrix element is

$$T(k\,\text{spin}-\text{flips}) \;=\; \exp(-2k\,L)\,\exp\left\{\frac{1}{2}\sum_i \left[\sigma_i\sigma_{i+1} + \sigma'_i\sigma'_{i+1}\right]\right\} \quad (9.2.18)$$
$$\simeq\; -\tau\, H_{k\,\text{spin}-\text{flips}}.$$

From eqn (9.2.16), we infer that

$$K \;\sim\; \tau, \quad (9.2.19)$$

while from eqns (9.2.17) and (9.2.18)

$$\exp(-2\,L) \;\sim\; \tau. \quad (9.2.20)$$

From these two equations, we see that K and $\exp(-2L)$ have to be proportional to each other and we denote by λ the proportionality factor

$$K \;=\; \lambda\,\exp(-2\,L). \quad (9.2.21)$$

We can identify the vertical lattice spacing with

$$\tau \;=\; \exp(-2L) \quad (9.2.22)$$

and put the horizontal coupling constant of the spins equal to

$$K \;=\; \lambda\,\tau. \quad (9.2.23)$$

In summary, the physical content of the model does not change in the limit $\tau \to 0$ if we rescale the vertical coupling constant as in eqn (9.2.22) and the horizontal one as in eqn (9.2.23). These formulas show that, in the hamiltonian limit, L grows very large while K becomes extremely small.

Critical value. The value $\lambda = 1$ identifies the critical point of the model. In fact, the scenario that emerges from the rescaling of the couplings can be easily derived from the discussion of Chapter 4, where we have seen that the critical line is given by

$$\sinh 2K\,\sinh 2L \;=\; 1. \quad (9.2.24)$$

Any pair of the coupling constants K and L that satisfies this equation corresponds to a critical situation of the Ising model, with an infinite value of the correlation length.

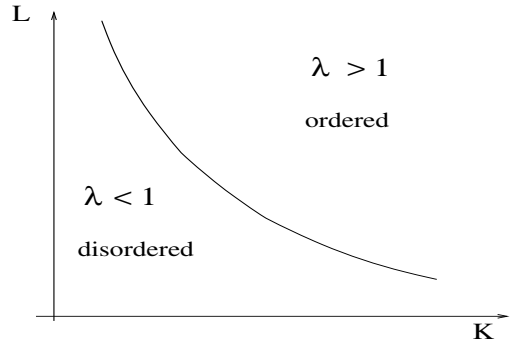


Fig. 9.3 *Phase diagram of the Ising model.*

If we consider the $L \to \infty$ asymptotic expression of this equation, when $\sinh 2L \to \frac{1}{2}\exp(2L)$ and $\sinh 2K \to 2K$, we get

$$K = \exp(-2L).$$

Comparing with eqn (9.2.21), we see that $\lambda = 1$ identifies the critical point of the model. Concerning the other values of λ, as already shown in Chapter 4, $\lambda > 1$ identifies the ordered phase of the model while $\lambda < 1$ corresponds to the disordered phase. The phase diagram is shown in Fig. 9.3. The parameter $1/\lambda$ provides a measurement of the displacement of the temperature from its critical value, as we will see in detail in the next section.

Once the lattice spacing τ has been identified with (9.2.22), we can proceed to derive the expression of the quantum hamiltonian that emerges in the limit $\tau \to 0$

$$H = -\lim_{\tau \to 0} \frac{1}{\tau} \log T.$$

The only processes that survive in this limit are those without a spin flip or those which induce only one spin flip, and correspondingly H is given by

$$H = -\sum_{a=1}^{n} \left[\tilde{\sigma}_1(a) + \lambda\, \tilde{\sigma}_3(a)\, \tilde{\sigma}_3(a+1)\right]. \tag{9.2.25}$$

In the hamiltonian limit, the two-dimensional classical Ising model is thus described by a simple one-dimensional quantum hamiltonian. In the basis in which the operators $\tilde{\sigma}_3(a)$ are diagonal, the term responsible for their spin flips is the operator $\tilde{\sigma}_1(a)$.

9.3 Order and Disorder Operators

In the thermodynamic limit, the sum is extended over all sites between $-\infty$ and $+\infty$ and the hamiltonian becomes

$$H = -\sum_{a=-\infty}^{\infty} \left[\tilde{\sigma}_1(a) + \lambda\, \tilde{\sigma}_3(a)\, \tilde{\sigma}_3(a+1)\right]. \tag{9.3.1}$$

To find its spectrum, let's first introduce the so-called *disorder operators*

$$\tilde{\mu}_3\left(r+\frac{1}{2}\right) = \prod_{\rho=-\infty}^{r} \tilde{\sigma}_1(\rho), \tag{9.3.2}$$

$$\tilde{\mu}_1\left(r+\frac{1}{2}\right) = \tilde{\sigma}_3(r)\tilde{\sigma}_3(r+1). \tag{9.3.3}$$

These operators are defined on the sites of the dual lattice, placed between two next neighbor sites of the original lattice. From their definition, $\tilde{\mu}_1(r+1/2)$ is sensitive to the alignment of two next neighbor spins. The other operator $\tilde{\mu}_3(r+1/2)$, acting on the original spins of the lattice, makes a spin flip of all those placed on the left-hand side of the point r, as shown in Fig. 9.4. Hence, starting from an ordered configuration of the spins, $\tilde{\mu}_3$ creates a *kink* excitation. This is a topological configuration that interpolates between the two states in which all spins are aligned either up or down, i.e. the ground states of the system. Since a kink changes the boundary conditions of the system, inspecting the values of the spins at the edge of the chain, one can easily infer whether there is an even or odd number of kinks in the system. It is also evident that the kink configurations tend to disorder the system and this justifies the terminology adopted for such an operator.

It is easy to check that the disorder operators $\tilde{\mu}_i$ satisfy the same algebra as the operators $\tilde{\sigma}_i$. Moreover, we have the algebraic relations

$$\begin{aligned}
&\tilde{\mu}_3^2 = \tilde{\mu}_1^2 = 1,\\
&\tilde{\mu}_3\left(r-\frac{1}{2}\right)\tilde{\mu}_3\left(r+\frac{1}{2}\right) = \tilde{\sigma}_1(r),\\
&\prod_{m<n}\tilde{\mu}_1\left(m+\frac{1}{2}\right) = \tilde{\sigma}_3(n+1),\\
&\left[\tilde{\mu}_1\left(r+\frac{1}{2}\right), \tilde{\mu}_3\left(r'+\frac{1}{2}\right)\right] = 2\delta_{r,r'},\\
&\left[\tilde{\mu}_3\left(r+\frac{1}{2}\right), \tilde{\mu}_3\left(r'+\frac{1}{2}\right)\right] = 0,\\
&\left[\tilde{\mu}_3\left(r+\frac{1}{2}\right), \tilde{\sigma}_1(r')\right] = 0.
\end{aligned} \tag{9.3.4}$$

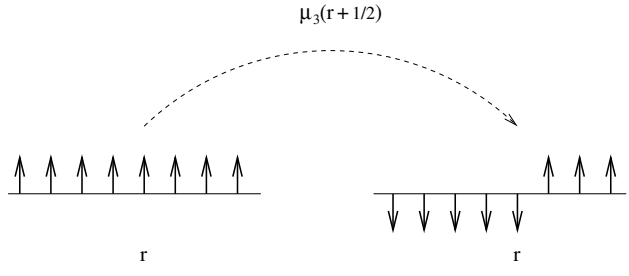


Fig. 9.4 *Action of $\tilde{\mu}_3$ on an ordered configuration of spins and creation of a kink state.*

We can now use the disorder operators to express the hamiltonian (9.3.1) as

$$\begin{aligned} H &= -\sum_r \left[\tilde{\mu}_3\left(r-\frac{1}{2}\right)\,\tilde{\mu}_3\left(r+\frac{1}{2}\right)+\lambda\,\mu_1\left(r+\frac{1}{2}\right)\right] \\ &= -\lambda\sum_r \left[\lambda^{-1}\,\tilde{\mu}_3\left(r-\frac{1}{2}\right)\,\tilde{\mu}_3\left(r+\frac{1}{2}\right)+\,\mu_1\left(r+\frac{1}{2}\right)\right]. \end{aligned} \tag{9.3.5}$$

This can be written as

$$H(\tilde{\sigma}\,;\lambda) \;=\; \lambda\, H(\tilde{\mu}\,;\lambda^{-1}). \tag{9.3.6}$$

Since the variables $\tilde{\sigma}_i$ and $\tilde{\mu}_i$ satisfy the same algebra, this expression simply expresses a symmetry of the system. This symmetry is nothing else but the Kramers–Wannier duality of the Ising model that expresses the invariance of the model under the substitutions

$$\begin{aligned} &\tilde{\mu}_1 \leftrightarrow \tilde{\sigma}_1 \\ &\tilde{\mu}_3 \leftrightarrow \tilde{\sigma}_3 \\ &\lambda \leftrightarrow \lambda^{-1}. \end{aligned} \tag{9.3.7}$$

Equation (9.3.6) implies that each eigenvalue H satisfies the functional equation

$$E(\lambda) \;=\; \lambda\, E(\lambda^{-1}). \tag{9.3.8}$$

Hence, there is a correspondence of the spectra for $\lambda > 1$ and $\lambda < 1$. Equation (9.3.8) leads to some important consequences, such as the exact value of λ for which the quantum hamiltonian is critical. To find this value, it is necessary to look for the vanishing of the mass gap, i.e. the difference between the two lowest eigenvalues. Denoting by $m(\lambda)$ the mass gap of the model, eqn (9.3.8) implies that, if $m(\lambda_*) = 0$ at a given critical value λ_*, then $m(\lambda)$ must also vanish at λ_*^{-1}. Assuming that there is only one critical point, the two values above must coincide and therefore

$$\lambda_* \;=\; \lambda_*^{-1} \qquad\rightarrow\qquad \lambda_* \;=\; 1. \tag{9.3.9}$$

As we previously mentioned, the critical value is indeed $\lambda_* = 1$.

9.4 Perturbation Theory

The function $m(\lambda)$ can be explicitly found by using perturbation theory. By adding a constant, let's first write the hamiltonian as

$$H \;=\; \sum_a \left[(1-\tilde{\sigma}_1(a)) - \lambda\tilde{\sigma}_3(a)\,\tilde{\sigma}_3(a+1)\right]. \tag{9.4.1}$$

In the high-temperature phase, λ is a small parameter and the hamiltonian can be split as

$$H \;=\; H_0 + \lambda V,$$

where

$$H_0 \;=\; \sum_a [1-\tilde{\sigma}_1(a)],$$
$$V \;=\; -\sum_a \tilde{\sigma}_3(a)\,\tilde{\sigma}_3(a+1).$$

To determine the first energy level in perturbation theory in λ, initially we have to identify the ground state of H_0 and its energy. It is easy to see that such a state has zero energy and it is characterized by the condition

$$\tilde{\sigma}_1(a)\,|\,0\,\rangle \;=\; |\,0\,\rangle, \quad \forall a. \tag{9.4.2}$$

In the basis in which $\tilde{\sigma}_3(a)$ is diagonal, the ground state is expressed by the tensor product of the vectors

$$|v_1(a)\rangle \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix},$$

each of them defined at the corresponding site of the lattice. The other eigenstate of $\tilde{\sigma}_1(a)$ (with eigenvalue -1) is expressed by the vector

$$|v_2(a)\rangle \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}.$$

Note that the operator $\tilde{\sigma}_3(a)$, which enters the perturbation V, maps one state to the other, $v_1(a) \leftrightarrow v_2(a)$.

With the ground state given by the tensor product $|\,0\,\rangle \;=\; \otimes_a |\,v_1(a)\,\rangle$, one can obtain an excited state by substituting, at an arbitrary point a of the system, the vector $v_2(a)$ for $v_1(a)$. Since the localization of this vector is arbitrary, this energy level has a degeneracy equal to the number n of the lattice sites.[1] One can take care of this degeneracy by introducing states with a well-defined quantum number of the lattice momentum. The state at zero momentum is obviously the only one invariant under translation and it is given by the linear combination

$$|\,-1\,\rangle \;=\; \frac{1}{\sqrt{n}}\sum_a \tilde{\sigma}_3(a)\,|\,0\,\rangle, \tag{9.4.3}$$

with $\langle\,-1\,|\,-1\,\rangle \;=\; 1$. The energy of this excited state can be computed perturbatively

$$E \;=\; \mathcal{E}_0 + \lambda\,\mathcal{E}_1 + \lambda^2\mathcal{E}_2 + \cdots, \tag{9.4.4}$$

where

$$\begin{aligned}
\mathcal{E}_1 \;&=\; \langle\,-1\,|V|\,-1\,\rangle,\\
\mathcal{E}_2 \;&=\; \langle\,-1\,|VgV|\,-1\,\rangle, \qquad (9.4.5)\\
\mathcal{E}_3 \;&=\; \langle\,-1\,|VgVgV|\,-1\,\rangle - \langle\,-1\,|V|\,-1\,\rangle\,\langle\,-1\,|Vg^2V|\,-1\,\rangle\,,\\
\cdots \;&=\; \cdots \qquad (9.4.6)
\end{aligned}$$

[1] n here provides an infrared cut-off, to be sent to infinity in the thermodynamic limit.

with the operator g defined by

$$g \;=\; \frac{1}{\mathcal{E}_0 - H}\,[1 - |\,-1\,\rangle\,\langle\,-1\,|]\,.$$

It is easy to see that $\mathcal{E}_0 = 2$. For the next term we have

$$\begin{aligned}\mathcal{E}_1 \;&=\; \langle\,-1\,|V|\,-1\,\rangle \\ &=\; -\frac{1}{n}\sum_{a,a'}\langle\,0\,|\tilde{\sigma}_3(a)\sum_b\,[\tilde{\sigma}_3(b)\,\tilde{\sigma}_3(b+1)]\;\,\tilde{\sigma}_3(a')\,|\,0\,\rangle.\end{aligned} \tag{9.4.7}$$

There are only two terms that contribute to this expression: $a = b$ and $b + 1 = a'$, or $a' = b$ and $b + 1 = a$. Since $\tilde{\sigma}_3^2 = 1$, both terms give a factor n that cancels the normalization factor. Hence $\mathcal{E}_1 = -2$.

If one carries on the computation to higher order (a task that we do not report here), there is a remarkable result: all of them are zero! In other words, the perturbative series truncates and coincides with its first two terms. For $\lambda < 1$, the exact mass gap is thus given by

$$m(\lambda) \;=\; 2(1 - \lambda).$$

This expression explicitly confirms that it vanishes at $\lambda = 1$. We can then use the duality relation, $m(\lambda) \;=\; \lambda\, m(\lambda^{-1})$, to obtain the mass gap for $\lambda > 1$. So, for all values of λ, the mass gap is expressed by

$$m(\lambda) \;=\; 2\,|1 - \lambda|. \tag{9.4.8}$$

9.5 Expectation Values of Order and Disorder Operators

In the ordered phase of the Ising model, described by $\lambda > 1$, the hamiltonian $H(\tilde{\sigma}\,;\lambda)$ with periodic boundary conditions has two possible vacuum states. The simplest way to obtain this result is to consider the limit $\lambda \to \infty$, in which the states with the minimum energy of the hamiltonian (9.3.1) are those in which the expectation values of $\tilde{\sigma}_3(a)$ and $\tilde{\sigma}_3(a + 1)$ coincide. Denoting by

$$|\,w_1(a)\,\rangle \;=\; \begin{pmatrix}1\\0\end{pmatrix}, \qquad |\,w_2(a)\,\rangle \;=\; \begin{pmatrix}0\\1\end{pmatrix}$$

the two eigenvectors of $\tilde{\sigma}_3(a)$ at the site a (with eigenvalues ± 1), there are then *two* degenerate states of minimum energy, given by

$$|\,0^+\,\rangle_{\lambda=\infty} \;=\; \otimes_a\,|\,w_1(a)\,\rangle, \qquad |\,0^-\,\rangle_{\lambda=\infty} \;=\; \otimes_a\,|\,w_2(a)\,\rangle. \tag{9.5.1}$$

The system will choose one of them, say $|\,0^+\,\rangle$, by the mechanics of spontaneous symmetry breaking: this can be induced by switching on a positive magnetic field B on all sites of the system and, once the spins are polarized in the direction of the field, switching B off.

For $\lambda > 1$ but finite, the corresponding vacuum states cannot be expressed by a simple expression such as those given above. However, even in the absence of an explicit

formula, let's denote by $|\,0\,\rangle_\lambda$ the vacuum state of $H(\tilde\sigma\,;\lambda)$ after the spontaneous symmetry breaking (this corresponds to $|\,0^+\,\rangle$). On this state, the operator $\tilde\sigma_3(r)$ has a non-vanishing expectation value

$$ {}_\lambda\langle\,0|\sum_a \tilde\sigma_3(a)|0\,\rangle_\lambda \neq 0. \tag{9.5.2}$$

The self-duality of the model allows us to interpret this result in an interesting way. Consider, in fact, the following hamiltonian

$$\lambda^{-1}\,H(\tilde\sigma\,;\lambda) + B\sum_a \tilde\sigma_3(a). \tag{9.5.3}$$

Applying a duality transformation, we have

$$\lambda^{-1}\,H(\tilde\sigma\,;\lambda) + B\sum_a \tilde\sigma_3(a) \;=\; H(\tilde\mu\,;\lambda^{-1}) + B\sum_a \prod_{b<a} \tilde\mu_1(b+1/2).$$

Expanding to first order in B, we obtain

$$ {}_\lambda\langle\,0|\sum_a \tilde\sigma_3(a)\,|\,0\,\rangle_\lambda \;=\; {}_{\lambda^{-1}}\langle\,0|\sum_a \prod_{b<a} \tilde\mu_1(b+1/2)\,|\,0\,\rangle_{\lambda^{-1}}, \tag{9.5.4}$$

where $|\,0\,\rangle_{\lambda^{-1}}$ is the vacuum state of $H(\tilde\mu\,;\lambda^{-1})$. Since the operators $\tilde\sigma_1$ and $\tilde\sigma_3$ are equivalent to $\tilde\mu_1$ and $\tilde\mu_3$, eqn (9.5.4) establishes that the operator $\tilde\mu_3(a+1/2) = \prod_{b<a}\tilde\sigma_1(m)$ has a non-zero expectation value in the high-temperature ground state of the model. This operator creates a kink excitation in the system and therefore the high-temperature ground state can be regarded as a *kink condensate*. Notice that the energy of the kink state is mostly localized around the point b and consequently the kink behaves as a localized particle. The kink state with the lowest energy is that at zero momentum, which can be written as

$$|\,1\,\mathrm{kink}\,\rangle \;=\; \frac{1}{\sqrt{n}}\sum_a \prod_{b<a} \tilde\sigma_1(b)\,|\,0\,\rangle_{\lambda^{-1}=0}. \tag{9.5.5}$$

From the duality relation, its mass coincides with the expression previously computed, $m(\lambda) = 2|1-\lambda|$.

For $\lambda < 1$, i.e. in the high-temperature phase, there is a similar argument: the operator $\tilde\sigma_3$ has a non-zero expectation value in the ground state of the system (which, in this phase, is unique) and consequently we have a vanishing expectation value of the disordered operator $\tilde\mu_3$ in the low-temperature phase.

9.6 Diagonalization of the Hamiltonian

The quantum hamiltonian of the Ising model can be explicitly diagonalized by means of the Wigner–Jordan transformation. To simplify the following expression, it is convenient to start from the hamiltonian

$$H \;=\; -\sum_a \left[\tilde\sigma_3(a) + \lambda\,\tilde\sigma_1(a)\,\tilde\sigma_1(a+1)\right] \tag{9.6.1}$$

which is unitarily equivalent to the previous one (9.3.1). Denoting the site of the lattice by $a = -n, -n+1, \ldots, n$ (we will consider later the limit $n \to \infty$), let's introduce the operators

$$c(a) = \prod_{b=-n}^{n} e^{i\pi\, \tilde{\sigma}^+(b)\, \tilde{\sigma}^-(b)}\ \tilde{\sigma}^-(a), \tag{9.6.2}$$

$$c^\dagger(a) = \tilde{\sigma}^+(a)\ \prod_{b=-n}^{n} e^{-i\pi\, \tilde{\sigma}^+(b)\, \tilde{\sigma}^-(b)}, \tag{9.6.3}$$

where

$$\begin{aligned} \tilde{\sigma}^+(a) &= \frac{1}{2}\left[\tilde{\sigma}_1(a) + i\tilde{\sigma}_2(a)\right], \\ \tilde{\sigma}^-(a) &= \frac{1}{2}\left[\tilde{\sigma}_1(a) - i\tilde{\sigma}_2(a)\right]. \end{aligned} \tag{9.6.4}$$

It is easy to show that $c(a)$ and $c^\dagger(a)$ are fermionic operators that satisfy the anticommutation relations

$$\left\{c(a), c^\dagger(b)\right\} \,=\, \delta_{a,b}, \quad \{c(a), c(b)\} \,=\, 0. \tag{9.6.5}$$

We can now write the hamiltonian (9.6.1) in a convenient form in terms of these new operators. As shown in Problem 2 at the end of the chapter, we have in fact

$$\begin{aligned} \tilde{\sigma}_3(a) \,&=\, 2\, c^\dagger(a)\, c(a) - 1, \\ \tilde{\sigma}_1(a)\, \tilde{\sigma}_1(a+1) \,&=\, \left[c^\dagger(a) - c(a)\right]\ \left[c^\dagger(a+1) - c(a+1)\right]. \end{aligned} \tag{9.6.6}$$

so that

$$H \,=\, -2 \sum_a c^\dagger(a)\, c(a) - \lambda \sum_a \left[c^\dagger(a) - c(a)\right]\ \left[c^\dagger(a+1) - c(a+1)\right] \tag{9.6.7}$$

is a quadratic expression of the fermionic operators. To diagonalize it, the first step is to take the Fourier transform

$$c(a) \,=\, \frac{1}{\sqrt{2n+1}} \sum_k e^{-ika}\, c_k, \qquad c^\dagger(a) \,=\, \frac{1}{\sqrt{2n+1}} \sum_k e^{ika}\, c_k^\dagger, \tag{9.6.8}$$

where c_k and $c_k^\dagger$ are the fermionic annihilation and creation operators in momentum space, and the momenta k take the discrete values

$$k \,=\, 0,\ \pm\frac{2\pi}{2n+1}\ \pm\ \frac{4\pi}{2n+1}, \ldots, \pm\frac{2\pi n}{2n+1}.$$

Substituting eqn (9.6.8) into (9.6.1), one has

$$H \,=\, -2 \sum_{k>0} (1 + \lambda\, \cos k)\, (c_k^\dagger\, c_k + c_{-k}^\dagger\, c_{-k}) + 2i\lambda \sum_{k>0} \sin k\, (c_k^\dagger\, c_{-k}^\dagger + c_k\, c_{-k}). \tag{9.6.9}$$

This hamiltonian is quadratic in c_k e $c_k^\dagger$ but this is still not sufficient to determine its spectrum. In fact, the ground state of H is not the state annihilated by the operators

c_k, since in the hamiltonian above there is also present a term $(c_k^\dagger\, c_{-k}^\dagger + c_k\, c_{-k})$. We need then to rewrite the hamiltonian in a canonical form in terms of some opportune creation and annihilation fermionic operators $\eta_k^\dagger$ and η_k:

$$H = \sum_k [\cdots]\, \eta_k^\dagger\, \eta_k. \tag{9.6.10}$$

Bogoliubov transformation. To achieve such a goal, it is necessary to make a Bogoliubov transformation: we introduce the operators

$$\begin{array}{lll} \eta_k &=& U_k\, c_k + iV_k\, c_{-k}^\dagger\ , \quad \eta_{-k} \;=\; U_k\, c_{-k} - iV_k\, c_k^\dagger \\ \eta_k^\dagger &=& U_k\, c_k^\dagger - iV_k\, c_{-k}\ , \quad \eta_{-k}^\dagger \;=\; U_k\, c_{-k} + iV_k\, c_k, \end{array} \tag{9.6.11}$$

where, in all these formulas, $k > 0$. The real coefficients U_k and V_k of this transformation are determined by the following requests:

- the operators η_k and $\eta_k^\dagger$ must be fermionic operators;
- the hamiltonian must be in a diagonal form in these new variables.

Imposing the validity of the relations

$$\{\eta_k, \eta_p^\dagger\} \;=\; \delta_{k,p}, \qquad \{\eta_k, \eta_p\} \;=\; \{\eta_k^\dagger, \eta_p^\dagger\} \;=\; 0, \tag{9.6.12}$$

we arrive at the equation

$$U_k^2 + V_k^2 \;=\; 1. \tag{9.6.13}$$

This allows us to parameterize these functions as

$$U_k \;=\; \cos\theta_k, \qquad V_k \;=\; \sin\theta_k. \tag{9.6.14}$$

Inverting now eqn (9.6.11)

$$\begin{array}{lll} c_k &=& U_k\, \eta_k - iV_k\, \eta_{-k}^\dagger, \quad c_{-k} \;=\; U_k\, \eta_{-k} + iV_k\, \eta_k^\dagger \\ c_k^\dagger &=& U_k\, \eta_k^\dagger + iV_k\, \eta_{-k}, \quad c_{-k}^\dagger \;=\; U_k\, \eta_{-k} - iV_k\, \eta_k, \end{array} \tag{9.6.15}$$

and inserting these formulas into the hamiltonian (9.6.9), we have

$$\begin{aligned} H = &\sum_{k>0} \left[-2(1+\lambda\cos k)(U_k^2 - V_k^2) + 4\lambda \sin k\, U_k\, V_k\right] (\eta_k^\dagger \eta_k + \eta_{-k}^\dagger \eta_{-k}) \\ &+ \sum_{k>0} \left[4i(1+\lambda\cos k)\, U_k\, V_k + 2i\lambda \sin k\, (U_k^2 - V_k^2)\right] (\eta_k^\dagger \eta_{-k}^\dagger + \eta_k \eta_{-k}). \end{aligned} \tag{9.6.16}$$

In order to bring the hamiltonian into the form (9.6.10), we need to impose

$$4(1+\lambda\cos k)\, U_k\, V_k + 2\lambda \sin k\, (U_k^2 - V_k^2) \;=\; 0. \tag{9.6.17}$$

Using (9.6.14), one has

$$2U_k\, V_k \;=\; \sin 2\theta_k, \qquad U_k^2 - V_k^2 \;=\; \cos 2\theta_k,$$

and eqn (9.6.17) becomes

$$4(1+\lambda\cos k)\,\sin 2\theta_k + 2\lambda\,\sin k\,\cos 2\theta_k \;=\; 0. \tag{9.6.18}$$

We have then the following equation for the angle θ_k

$$\tan 2\theta_k \;=\; -\frac{\lambda\,\sin k}{1+\lambda\,\cos k}. \tag{9.6.19}$$

The geometric interpretation of this equation is given in Fig. 9.5; taking into account the $(-)$ sign, its solution is expressed by

$$\begin{aligned} \sin 2\theta_k \;&=\; \frac{\lambda\,\sin k}{\sqrt{1+2\lambda\,\cos k+\lambda^2}}, \\ \cos 2\theta_k \;&=\; -\frac{1+\lambda\,\cos k}{\sqrt{1+2\lambda\,\cos k+\lambda^2}}\;\;. \end{aligned} \tag{9.6.20}$$

Spectrum of the hamiltonian. With this determination of U_k and V_k, coming back to eqn (9.6.16), we have

$$H \;=\; 2\sum_k \Lambda_k\;\eta_k^\dagger\,\eta_k + \text{costant} \tag{9.6.21}$$

where

$$\Lambda_k \;=\; \sqrt{1+2\lambda\cos k+\lambda^2}. \tag{9.6.22}$$

The plot of this function is given in Fig. 9.6. Its minimum is at $k=\pm\pi$, with a value

$$\Lambda_{\pm\pi} \;=\; 2|1-\lambda|$$

that is in agreement with the perturbative analysis done in the previous section.

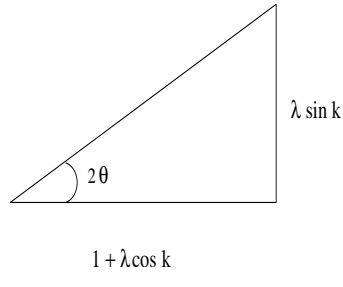


Fig. 9.5 *Relation between the parameters of the Bogoliubov transformation.*

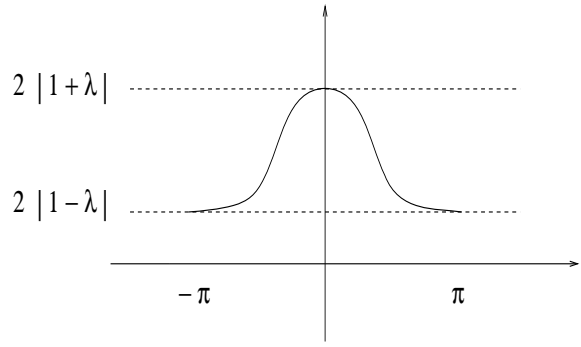


Fig. 9.6 *Dispersion relation of the fermionic particle.*

In taking the continuum limit, it is convenient to restore the lattice spacing α and measure the momentum with respect to its minimum value, i.e.

$$k \,=\, \pi + k'\alpha.$$

Let's also define the energy with the correct physical dimension

$$E(k') \,=\, \frac{\Lambda_k}{2\alpha}.$$

In the continuum limit $\alpha \to 0$, we shall take $k \to \pi$ in order to have a finite value of the momentum, and therefore

$$E(k') \,=\, \sqrt{\frac{(1-\lambda)^2}{\alpha^2} + \lambda k'^2}. \tag{9.6.23}$$

We arrive in this way at a relativistic dispersion relation. If λ is sufficiently close to the critical value, $\lambda \simeq 1$, we have the dispersion relation of a fermionic particle with mass

$$m \,=\, \frac{(1-\lambda)}{\alpha}. \tag{9.6.24}$$

At $\lambda = 1$, it becomes the dispersion relation of a massless particle

$$E(k') \simeq |k'|. \tag{9.6.25}$$

In terms of the fields

$$\eta(a) \,=\, \frac{1}{\sqrt{2n+1}} \sum_k e^{ika}\, \eta_k, \qquad \eta^\dagger(a) \,=\, \frac{1}{\sqrt{2n+1}} \sum_k e^{-ika}\, \eta_k^\dagger$$

we can define the new fermionic fields

$$\psi_1(a) \,=\, \frac{1}{2}(\eta(a) + \eta^\dagger(a)), \qquad \psi_2(a) \,=\, \frac{1}{2i}(\eta(a) - \eta^\dagger(a)). \tag{9.6.26}$$

They satisfy the relations

$$\begin{aligned} \psi_i^\dagger(a) &= \psi_i(a), \\ \{\psi_i(a), \psi_j(b)\} &= \delta_{ij}\,\delta_{ab}, \\ \psi_i^2(a) &= \frac{1}{2}. \end{aligned} \tag{9.6.27}$$

Hence they are neutral fermionic fields, also known in the literature as *Majorana fermions*.

9.7 Dirac Equation

In this section we present another way to show the fermionic content of the two-dimensional Ising model. Note that, using eqn (9.3.1), we can determine the equation of motion of the operator $\tilde{\sigma}_3(r)$

$$\frac{\partial}{\partial\tau}\tilde{\sigma}_3(r) = [H, \tilde{\sigma}_3(r)] = -i\tilde{\sigma}_2(r) = \tilde{\sigma}_1(r)\tilde{\sigma}_3(r). \tag{9.7.1}$$

Using the dual expression (9.3.5) of the hamiltonian, we can also derive the equation of motion of the operator $\tilde{\mu}_3(r+1/2)$

$$\begin{aligned} \frac{\partial}{\partial\tau}\tilde{\mu}_3\left(r+\frac{1}{2}\right) &= \left[H, \tilde{\mu}_3\left(r+\frac{1}{2}\right)\right] = -i\,\lambda\,\tilde{\mu}_2\left(r+\frac{1}{2}\right) \\ &= \lambda\,\tilde{\mu}_1\left(r+\frac{1}{2}\right)\,\tilde{\mu}_3\left(r+\frac{1}{2}\right) = \lambda\,\tilde{\sigma}_3(r)\,\tilde{\sigma}_3(r+1)\,\tilde{\mu}_3\left(r+\frac{1}{2}\right). \end{aligned} \tag{9.7.2}$$

These equations of motion are nonlinear and difficult to solve. However, they can be put in a linear form by defining

$$\psi_1(r) = \tilde{\sigma}_3(r)\,\tilde{\mu}_3\left(r+\frac{1}{2}\right), \tag{9.7.3}$$

$$\psi_2(r) = \tilde{\sigma}_3(r)\,\tilde{\mu}_3\left(r-\frac{1}{2}\right). \tag{9.7.4}$$

Using the algebraic properties of the variables $\tilde{\sigma}_i$ and $\tilde{\mu}_i$, the equation of motion for these new variables can be written as

$$\frac{\partial\psi_1(r)}{\partial\tau} = -\psi_2(r) + \lambda\,\psi_2(r+1), \tag{9.7.5}$$

$$\frac{\partial\psi_2(r)}{\partial\tau} = -\psi_1(r) + \lambda\,\psi_1(r-1). \tag{9.7.6}$$

Restoring the lattice spacing α with the substitution $(r\pm1)\to(r\pm\alpha)$ and going to the continuum limit $\alpha\to0$, we have

$$\frac{\partial\psi_1(r)}{\partial t} = -(1-\lambda)\,\psi_2(r) + \lambda\,\frac{\partial\psi_2(r)}{\partial r}\,\alpha, \tag{9.7.7}$$

$$\frac{\partial\psi_2(r)}{\partial t} = -(1-\lambda)\,\psi_1(r) - \lambda\,\frac{\partial\psi_1(r)}{\partial r}\,\alpha. \tag{9.7.8}$$

The two fields $\psi_1(r)$ and $\psi_2(r)$ can be organized in a spinorial field

$$\psi(r) \;=\; \begin{pmatrix} \psi_1(r) \\ \psi_2(r) \end{pmatrix}, \tag{9.7.9}$$

with anticommutation relations

$$\{\psi_1(r), \psi_2(r')\} = 2\delta_{r,r'} \tag{9.7.10}$$

$$\{\psi_2(r), \psi_2(r')\} = 2\delta_{r,r'}. \tag{9.7.11}$$

A compact expression for the equation of motion is then given by

$$\left(\gamma^0 \frac{\partial}{\partial t} + \gamma^3 \frac{\partial}{\partial r} + m\right) \psi \;=\; 0, \tag{9.7.12}$$

with $t = \alpha\tau$, $m \;=\; (1-\lambda)/\alpha$ and with the euclidean γ matrices given by

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We arrive in this way at a Dirac equation for a free fermionic neutral field, i.e. a Majorana fermion.

Note that the equation of motion can be derived from the euclidean action

$$\mathcal{S} \;=\; \int d^2x\, \bar{\psi}\; (\gamma^\mu \partial_\mu + m)\; \psi, \tag{9.7.13}$$

where $\bar{\psi} \equiv \psi\, \gamma^0$. For reasons that will become clearer later, it is convenient to introduce the complex coordinates $z = x + it$ e $\bar{z} = x - it$, with $\partial_z \;=\; \frac{1}{2}(\partial_x - i\partial_t)$ and $\partial_{\bar{z}} \;=\; \frac{1}{2}(\partial_x + i\partial_t)$, and define two new fermionic components by

$$\Psi(z,\bar{z}) \;=\; \frac{\psi_1 + i\psi_2}{\sqrt{2}}, \qquad \bar{\Psi}(z,\bar{z}) \;=\; \frac{\psi_1 - i\psi_2}{\sqrt{2}}. \tag{9.7.14}$$

In these new variables, the action becomes

$$\mathcal{S} \;=\; \int d^2z\, \left[\Psi\, \partial_{\bar{z}}\, \Psi + \bar{\Psi}\, \partial_z\, \bar{\Psi} + im\, \bar{\Psi}\, \Psi\right], \tag{9.7.15}$$

with the equation of motion given by

$$\partial_{\bar{z}}\, \Psi \;=\; \frac{im}{2}\, \bar{\Psi}, \qquad \partial_z\, \bar{\Psi} = -\frac{im}{2}\, \Psi. \tag{9.7.16}$$

When the mass of the fermion field vanishes, Ψ becomes a purely analytic field while $\bar{\Psi}$ a purely anti-analytic one. The duality of the Ising model is expressed by the invariance of this fermionic theory under the transformations

$$\begin{array}{c} m \to -m \\ \Psi \to \Psi \\ \bar{\Psi} \to -\bar{\Psi}. \end{array} \tag{9.7.17}$$

As we shall prove in Chapter 14, in the continuum limit of the model the order and disorder operators satisfy the operator relation

$$\sigma(z,\bar{z})\,\mu(z',\bar{z}') \;=\; \frac{1}{\sqrt{2}|z-z'|^{1/4}}\left[\omega\,(z-z')^{1/2}\,\Psi(z',\bar{z}') + \bar{\omega}\,(\bar{z}-\bar{z}')^{1/2}\,\bar{\Psi}(z',\bar{z}')\right], \tag{9.7.18}$$

when $|z-z'| \to 0$, with $\omega = \exp(i\pi/4)$, $\bar{\omega} = \exp(-i\pi/4)$.

The interpretation of the fermionic field in terms of particle excitations is obtained by making an analytic continuation to Minkowski space. In two dimensions it is convenient to parameterize the relativistic dispersion relations in terms of the rapidity θ as follows: $E = m\,\cosh\theta$ and $p = m\,\sinh\theta$. The mode expansion of the fermionic field becomes

$$\Psi(x,t) = \int \frac{d\theta}{2\pi}\left[\omega e^{\frac{\theta}{2}}\,A(\theta)\,e^{-im(t\cosh\theta - x\sinh\theta)} \;+\; \bar{\omega}e^{\frac{\theta}{2}}\,A^{\dagger}(\theta)\,e^{im(t\cosh\theta - x\sinh\theta)}\right] \tag{9.7.19}$$

$$\bar{\Psi}(x,t) = -\int \frac{d\theta}{2\pi}\left[\bar{\omega}e^{-\frac{\theta}{2}}\,A(\theta)\,e^{-im(t\cosh\theta - x\sinh\theta)} \;+\; \omega e^{-\frac{\theta}{2}}\,A^{\dagger}(\theta)\,e^{im(t\cosh\theta - x\sinh\theta)}\right],$$

where $A(\theta)$ and $A^{\dagger}(\theta)$ are, respectively, the annihilitation and creation operators of a neutral fermionic particle. They satisfy the anticommutation relations

$$\left\{A(\theta), A^{\dagger}(\theta)\right\} \;=\; 2\pi\,\delta(\theta-\theta). \tag{9.7.20}$$

Using this mode expansion and the anticommutation relations it is easy to compute the correlation functions of the fermionic field (see Problem 4).

References and Further Reading

The important role of the order and disorder operators is discussed in the papers:

E. Fradkin, L. Susskind, *Order and disorder in gauge systems and magnets*, Phys. Rev. D 17 (1978), 2637.

L.P. Kadanoff, H. Ceva, *Determination of an operator algebra for the two-dimensional Ising model*, Phys. Rev. B3 (1971), 3918.

For the fermionic formulation of the Ising model we refer the reader to the articles:

T.D. Schultz, D.C. Mattis, E.H. Lieb, *Two-dimensional Ising model as a soluble problem of many fermions*, Rev. Mod. Phys. 36 (1964), 856.

E.H. Lieb, T. D. Schultz, D.C. Mattis, *Two soluble models of anti-ferromagnetic chain*, Ann. Phys. 16 (1961), 407.

P. Pfeuty, *The one-dimensional Ising model in a transverse field*, Ann. Phys. 57 (1970), 79.

J.B. Zuber, C. Itzykson, *Quantum field theory and the two-dimensional Ising model*, Phys. Rev. D 15 (1977), 2875.

M. Bander, C. Itzykson, *Quantum field theory calculation of the two-dimensional Ising model correlation function*, Phys. Rev. D 15 (1977), 463.

For an attempt to extend the fermionic formulation to the three-dimensional Ising model see:

C. Itzykson, *Ising fermions (II). Three dimensions*, Nucl. Phys. B 210 (1982), 477.

Problems

1. Anticommutation relations

Prove that the operators $c(a)$ and $c^\dagger(b)$ satisfy the anticommutation relations

$$\{c(a), c^\dagger(b)\} \;=\; \delta_{a,b}.$$

2. Fermion identities

Prove that

$$\begin{aligned} c(a)\, c^\dagger(a+1) \;&=\; -\tilde\sigma^-(a)\, \tilde\sigma^+(a+1) \\ c^\dagger(a)\, c^\dagger(a+1) \;&=\; \tilde\sigma^+(a)\, \tilde\sigma^+(a+1) \\ c(a)\, c(a+1) \;&=\; -\tilde\sigma^-(a)\, \tilde\sigma^-(a+1). \end{aligned}$$

Moreover, show that the order operators of the Ising model can be expressed in terms of the fermion operators $c(a)$ and $c^\dagger(a)$ as

$$\begin{aligned} \tilde\sigma_3(a) \;&=\; 2\, c^\dagger(a)\, c(a) - 1, \\ \tilde\sigma_1(a)\, \tilde\sigma_1(a+1) \;&=\; \left[c^\dagger(a) - c(a)\right]\, \left[c^\dagger(a+1) - c(a+1)\right]. \end{aligned}$$

3. Duality

Show that the Dirac equation (9.7.12) with $m > 0$ is linked by a unitary transformation to the one with $m < 0$. The map between the two hamiltonians establishes the duality relation of the Ising model in its fermionic formulation.

4. Correlation functions

Use the mode expansion of the fermion field, given in eqn (9.7.19), and the anti-commutation relations of the operators $A(\theta)$ and $A^\dagger(\theta)$, to compute the correlation functions

$$\begin{aligned} G_1(x,t) &= \langle 0|\Psi(x,t)\Psi(0,0)|0\rangle \\ G_2(x,t) &= \langle 0|\bar\Psi(x,t)\Psi(0,0)|0\rangle. \end{aligned}$$

5. XXZ model

Consider the quantum spin chain of the XXZ model with Hamiltonian

$$H \;=\; J_1 \sum_{k=1}^{N} (S_k^x S_{k+1}^x + S_k^y S_{k+1}^y) + J_2 \sum_k S_k^z S_{k+1}^z.$$

Assume periodic boundary conditions. The spin operator $\vec{S}$ is in the 1/2 representation and given by $\vec{S} = 1/2\vec{\sigma}$, where $\vec{\sigma}$ is the set of Pauli matrices.

a Show that the sign of J_1 in H can be chosen freely without changing the physical properties of the model. Show that the same is not true for the sign of J_2.

b Discuss the symmetry of the system when $J_2/J_1 \to 0$ and $J_2/J_1 \to \infty$.

c Use the fermionic representation of the operators $S_k^{\pm} = S_k^x \pm i S_k^x$ and S_k^3 to write the hamiltonian as

$$\begin{aligned} H \;=\; & \frac{J_1}{2} \sum_{a=1}^{N} \left[c^\dagger(a) c(a+1) + c^\dagger(a+1) c(a) \right] \\ & + J_2 \sum_{a=1}^{N} \left(c^\dagger(a) c(a) - \frac{1}{2} \right) \left(c^\dagger(a+1) c(a+1) - \frac{1}{2} \right). \end{aligned}$$

d Consider the case $J_2 = 0$, the so-called XY model. Use the Bogoliubov transformation to find in this case the spectrum of the fermionic form of the hamiltonian.

10
Conformal Field Theory

A physical law must possess mathematical beauty.

P.A.M. Dirac

10.1 Introduction

In the previous chapters we have seen that, coming close to a critical point, the correlation length of a statistical system diverges and consequently there are fluctuations on all possible scales. In such a regime, the properties of the statistical systems can be efficiently described by a quantum field theory. Right at the critical point, the correlation length is infinite: the corresponding field theory is therefore massless and becomes invariant under a dilation of the length-scales

$$x_a \to \lambda\, x_a.$$

Under this transformation the fields Φ_i associated to the order parameters transform as

$$\Phi_i \to \lambda^{d_i}\, \Phi_i,$$

where d_i here denote their anomalous dimensions. Finding the spectrum of the anomalous dimensions is a central problem of the theory: in fact, from Chapter 8 we know that they determine the critical exponents of the various thermodynamic quantities. The singularities of these thermodynamic functions are associated to the fixed points of the renormalization group. In turn, in the vicinity of the fixed points there is a surface of instability that is spanned by the relevant operators present at the fixed points. In order to determine all fixed points and the spectrum of the operators nearby, A. Polyakov has put forward the hypothesis that the critical fluctuations are invariant under the set of conformal transformations. These transformations have the distinguished property of stretching locally the lengths of the vectors but leaving their relative angles invariant. It is important to stress that in systems with local interactions that are invariant under translations and rotations, the invariance under a global dilatation automatically implies an invariance under the conformal transformations. From this point of view, the classification of the fixed points of the renormalization group is equivalent to finding all possible quantum field theories with conformal symmetry.

It is worth emphasizing that the construction of such theories is based on an approach that is completely different from the lagrangian formalism that is usually used in quantum field theory and that was discussed in Chapter 7. The approach

that we present in this chapter is based on the algebra of local fields. We assume first of all the existence of a basis of local operators that include among others the order parameters. Secondly, we make the hypothesis that any other quantity, such as products of order parameters, can be expanded in terms of the local operators of the basis. In this way, one is naturally led to introduce the concept of the Operator Product Expansion (OPE) and its corresponding algebra.

In this chapter we study the general properties of these conformal invariant theories, pointing out the peculiar aspects that arise in two dimensions. The next chapter will be devoted to the analysis of an important class of two-dimensional conformal field theories, the so-called *minimal models*: their remarkable property is to have an operatorial algebra that closes within a finite number of conformal families. Other aspects of two-dimensional conformal theories will be the subject of subsequent chapters.

10.2 The Algebra of Local Fields

The main goal of a quantum field theory is the determination of the correlation functions

$$G(x_1, x_2, \dots, x_n) \,=\, \langle A_1(x_1) A_2(x_2) \dots A_n(x_n) \rangle .$$

where $A_i(x_i)$ are regular local functions that, for simplicity, we will assume are constructed in terms of only one fundamental field $\varphi(x)$. Below we consider the euclidean formulation of the theory and our definition of the vacuum expectation values is provided by the functional integral, with Boltzmann weight given by the action $S[\varphi]$ of the field φ

$$G(x_1, \dots, x_n) \,\equiv\, \frac{1}{Z} \int \mathcal{D}\varphi \, A_1(x_1) \dots A_n(x_n) \, e^{-S[\varphi]} . \tag{10.2.1}$$

The correct normalization is ensured by the partition function Z

$$Z \,=\, \int \mathcal{D}\varphi \, e^{-S[\varphi]} .$$

In order to clarify the important concept of the algebra of the local field, it is useful to initially consider the free bosonic field $\phi(x)$. In this case, using $\phi(x)$ and the normal ordered product of its powers,[1] we can define the local scalar densities

$$\phi(x), : \phi^2(x) :, \; : \phi^3(x) :, \; \dots : \phi^n(x) : . \tag{10.2.2}$$

In a euclidean D-dimensional space-time, the scalar field $\phi(x)$ has dimension equal to $d_\varphi = (D-2)/2$ (in mass units) and the dimensions of the composite fields (10.2.2) are given by[2]

$$d_\phi, \, 2d_\phi, \, 3d_\phi, \dots, \, nd_\phi \tag{10.2.3}$$

[1] By normal ordered product we mean here the subtraction of the divergent term coming from the propagator of the product of two operators, i.e. $: \phi^2(x) := \lim_{y \to x} \phi(x)\phi(y) - \langle 0 | \phi(x)\phi(y) | 0 \rangle$. All other normal ordered products can be iteratively constructed starting from this relation.

[2] In the following discussion we assume $D \neq 2$. The $D = 2$ case will be discussed in detail later.

as is easily seen by applying Wick's theorem. For instance

$$\langle : \phi^2(x) :: \phi^2(y) : \rangle \;=\; 2\,[\langle \phi(x)\phi(y)\rangle]^2$$

and the singularity $1/r^{2d_\phi}$ of the correlator $\langle \phi(x)\phi(y)\rangle$ for $r = |x_1 - x_2| \to 0$, gives rise to

$$\langle : \phi^2(x) :: \phi^2(y) : \rangle \;\sim\; \frac{1}{r^{4d_\phi}}.$$

An analogous result holds for the composite operators $: \phi^n(x) :$, so that we arrive at the sequence of dimensions (10.2.3).

The set of fields $\{: \phi^n(x) :\}$, to which we have to add the identity operator $\mathbf{I} \equiv \phi^0$, can be used to express any other regular density $A(x)$ of the free bosonic field. This is done by the series expansion of $A(x)$

$$A(x) \;=\; \sum_{n=0}^{\infty} a_n \;: \phi^n(x) : . \tag{10.2.4}$$

Basis of local operators. As in the free case, we make the hypothesis that there is a similar set of fields also in the interacting theories. Namely, we assume the existence of a numerable set of fields[3] $\varphi_i(x)$, that are eigenvectors of the dilation operator, whose dimensions are defined by the condition

$$\varphi_i(x) \;=\; \lambda^{d_i}\, \varphi_i(\lambda x). \tag{10.2.5}$$

Moreover, we assume that any other operator can be expressed as a linear combination of the fields above:

$$A(x) \;=\; \sum_{n=0}^{\infty} a_n \varphi_n(x). \tag{10.2.6}$$

At the critical point, the propagator of these operators is

$$\mathcal{D}_{ij}(x_1 - x_2) \;\equiv\; \langle \varphi_i(x_1)\varphi_j(x_2)\rangle \;=\; \frac{\mathcal{A}_{ij}}{|\, x_1 - x_2\,|^{d_i + d_j}}, \tag{10.2.7}$$

where the numerical constant $\mathcal{A}_{ij}$ expresses their normalization.

In an interactive theory, however, the dimensions d_i are not expressed in a simple way as in the free theory. As a matter of fact, one of the fundamental problems is the determination of their values.

It is worth saying that the validity of the operatorial identity (10.2.6), as well as other identities of similar nature that we will meet later on, has to be understood in a weak sense: this means that it can be used straighforwardly in expressions that involve correlation functions but it can give rise to inconsistencies if interpreted strictly as an operatorial identity (see Problem 1). So, for instance, to calculate the correlation

[3]For simplicity we consider below only scalar fields. Quantities with spin will be considered later.

function $\langle A(x)B(y)\dots C(z)\rangle$, we can use the expansion (10.2.6) to express any of these fields – say $A(x)$ – arriving in this way at the expression

$$\langle A(x)B(y)\dots C(z)\rangle \;=\; \sum_{n=0}^{\infty} a_n \,\langle \varphi_n(x)B(y)\dots C(z)\rangle.$$

In the vicinity of the critical point, the theory has a mass scale m that is a small quantity (it is related to the correlation length ξ by the relation $\xi = m^{-1}$), and the vacuum expectation values of the fields φ_n can be expressed as

$$\langle \varphi_n(x)\rangle \;=\; \mu_n\, m^{d_n}, \tag{10.2.8}$$

where μ_n are dimensionless quantities. At the critical point $m \to 0$ and all fields $\varphi_n(x)$ $(n = 1, 2, \dots)$, but the identity operator, have zero expectation value[4]

$$\langle \varphi_n(x)\rangle = 0, \quad n = 1, 2, \dots \tag{10.2.9}$$

Operator algebra. So far we have discussed the operatorial expressions that refer to a given point. Consider now the product of two fluctuating fields $A(x_1)B(x_2)$ placed at two distinct points. If their separation is much less than the correlation length ξ of the system, in particular if we consider the limit $x_1 \to x_2$, it is a natural hypothesis that the properties of this composite operator become those of a local operator, so that it can be expanded in the same basis φ_n given above:

$$A(x_1)B(x_2) \;=\; \sum_{k=0}^{\infty} \beta(x_1, x_2)\, \varphi_k(x_2), \tag{10.2.10}$$

where the coefficients $\beta(x_1, x_2)$ contain the dependence on the coordinates x_1 and x_2. If we specialize this relation to the case in which both $A(x_1)$ and $B(x_2)$ are themselves members of the basis, we arrive at the operator algebra

$$\varphi_p(x_1)\varphi_q(x_2) \;=\; \sum_{r=0}^{\infty} C^r_{pq}(x_1, x_2)\, \varphi_r(x_2). \tag{10.2.11}$$

The function $C^r_{pq}(x_1, x_2)$ can be further constrained. If the system is invariant under translations, it can only depend on the separation $\mid x_1 - x_2 \mid$ of the two points. At the critical point, the system is also invariant under the scale transformation $x \to \lambda x$ and, for the transformation law of the fields φ_n, it is easy to see that $C^r_{pq}(\mid x_1 - x_2 \mid)$ is a homogeneous function of degree $d_r - d_p - d_q$. Hence, it can be written as

$$C^r_{pq}(x_1, x_2) \;=\; c^r_{pq} \frac{1}{\mid x_1 - x_2 \mid^{d_p + d_q - d_r}}, \tag{10.2.12}$$

where c^r_{pq} are pure numbers, known as the *structure constants of the operator algebra*.

[4]This result is obvious if $d_n > 0$. We will see later that, for the conformal invariance of the theories of the fixed points, the same conclusion also holds when there are fields with negative dimensions.

To summarize, the hypothesis of an algebra of the local fields consists of: (i) the existence of a basis made of scaling operators $\varphi_i(x)$ with dimensions d_i; (ii) the validity of the operatorial algebra

$$\varphi_p(x_1)\varphi_q(x_2) \;=\; \sum_{r=0}^{\infty} c_{pq}^{r} \frac{1}{\mid x_1 - x_2 \mid^{d_p+d_q-d_r}} \varphi_r(x_2), \tag{10.2.13}$$

which holds at the critical point of the theory. The solution of the field theories that describe the fixed points of the renormalization group consists of finding the spectrum of the dimensions d_i of the scaling fields, together with the set of structure constants c_{pq}^{r}: once all these quantities are known, one can compute in principle any other correlation functions (see Problem 2).

Some consequences. Note that there are some immediate consequences of the operator algebra: first of all, to be consistent, the algebra (10.2.13) has to be associative. Consider the four-point correlation functions of the fields $\varphi_i(x)$ shown in Fig. 10.1:

$$G_{ijkl}(x_1, x_2, x_3, x_4) \;=\; \langle \varphi_i(x_1)\varphi_j(x_2)\varphi_k(x_3)\varphi_l(x_4) \rangle. \tag{10.2.14}$$

As shown in Fig. 10.2, using the operator algebra this correlator can be computed in two equivalent ways:

- expanding the product $\varphi_i(x_1)\varphi_j(x_2)$ and then contracting the resulting field $\varphi_m(x_2)$ with the field $\varphi_n(x_4)$ that comes from an analogous expansion of the product $\varphi_k(x_3)\varphi_l(x_4)$ (with a sum over the intermediate indices m and n);
- alternatively, expanding the two pairs $\varphi_i(x_1)\varphi_k(x_3)$ and $\varphi_j(x_2)\varphi_l(x_4)$, with a final sum over the indices of the propagators of the resulting fields.

The equivalence of these two procedures is known as the *duality symmetry* of the four-point correlation function. It corresponds to the associativity condition of the operatorial algebra and consists of the infinite number of constraints

$$\sum_{m,n} C_{ij}^{m}(x_1{-}x_2)\, \mathcal{D}_{m,n}(x_2{-}x_4)\, C_{kl}^{n}(x_3{-}x_4) \;=\; \sum_{m,n} C_{ik}^{m}(x_1{-}x_3)\, \mathcal{D}_{m,n}(x_3{-}x_4)\, C_{jl}^{n}(x_2{-}x_4). \tag{10.2.15}$$

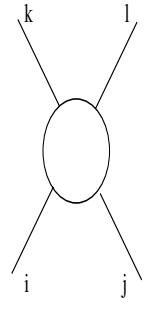


Fig. 10.1 *Four-point correlation functions of the scaling fields.*

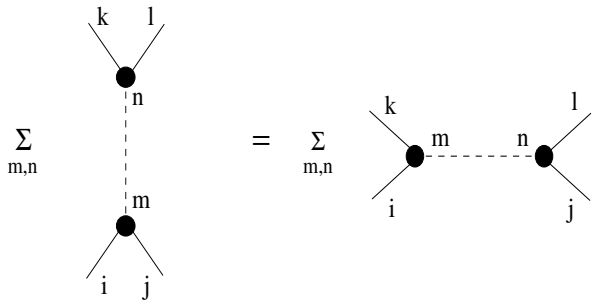


Fig. 10.2 *Duality of the four-point correlation function that corresponds to the associativity condition of the algebra (10.2.13).*

Using the expressions (10.2.7) and (10.2.12), this set of equations can be in principle enforced to determine all the scaling dimensions d_i and the structure constants c^r_{pq} of the algebra. For this reason they are known as *conformal bootstrap equations*:[5] if we were able to solve them, the dynamical data would be determined self-consistently from the theory itself. They were proposed originally by A. Polyakov as an alternative approach to solve the quantum field theories of the critical points. Unfortunately their direct solution proved to be extremely difficult and not much progress has been achieved by their analysis. An important step forward can instead be obtained by studying the important consequences of an additional symmetry of the fixed points, namely the conformal symmetry.

10.3 Conformal Invariance

At the critical point, a statistical system is invariant under a global dilatation of the length-scale, $x \to \lambda x$. Consider now two subsystems, separated by a distance L considerably larger than their linear dimension l: in this case, it is obvious that the fluctuations will be more uncorrelated when the ratio L/l becomes large as in Fig. 10.3. However, near the critical point, there does not exist any length-scale and the ratio above can be made arbitrarily large. In this way we arrive at the conclusion that the two subsystems are statistically independent, i.e. that the system is not only invariant under a global dilatation but also under the *local* scale transformations

$$x \to \lambda(x)\, x \tag{10.3.1}$$

also called *conformal transformations*. The previous considerations can be formulated as a theorem, developed originally by A. Polyakov:[6] a physical system with local interactions that is invariant under translations, rotations, and a global dilatation, is also automatically invariant under the larger class of conformal transformations. Before presenting the detailed analysis of this important result, it is useful to discuss the main properties of conformal transformations.

[5]The term "bootstrap" denotes the circumstance in which a physical theory owes its validity to its internal consistency. Later we will meet other theories based on bootstrap methods.

[6]A.M. Polyakov, *Conformal symmetry of critical fluctuations*, JETP Lett. 12 (1970), 381.

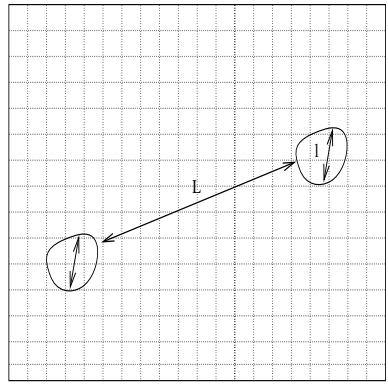


Fig. 10.3 *Subsystems of linear dimension l separated by a distance $L \gg l$.*

10.3.1 Conformal Transformations in D Dimensions

By definition, a conformal transformation of the coordinates $x \to x'$ is an invertible mapping that leaves invariant the metric tensor $g_{\mu,\nu}(x)$ up to a local scale factor

$$g'_{\mu\nu}(x') \;=\; \Lambda(x)\, g_{\mu\nu}(x). \tag{10.3.2}$$

It is useful to characterize the infinitesimal form of such transformations: using the tensor properties of the metric $g_{\mu,\nu}(x)$, under the transformation

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x) \tag{10.3.3}$$

we have

$$g_{\mu\nu} \to g_{\mu\nu} \to g_{\mu\nu} + (\partial_\mu \epsilon_\nu(x) + \partial_\nu \epsilon_\mu(x)).$$

If we now impose the conformal invariance of the metric, eqn (10.3.2), we have

$$\partial_\mu \epsilon_\nu(x) + \partial_\nu \epsilon_\mu(x) \;=\; \rho(x)\, g_{\mu\nu}, \tag{10.3.4}$$

where the local function $\rho(x)$ can be easily determined by taking the trace of both terms of this expression

$$\rho(x) \;=\; \frac{2}{D} \partial \cdot \epsilon$$

where D is the dimension of the space. If we now take the limit of a flat euclidean space, with the usual metric $\delta_{\mu\nu} = \mathrm{diag}(1, 1, \ldots, 1)$, we obtain the differential equations that characterize the infinitesimal conformal transformations

$$\partial_\mu \epsilon_\nu(x) + \partial_\nu \epsilon_\mu(x) \;=\; \frac{2}{D} \partial \cdot \epsilon\, \delta_{\mu\nu}. \tag{10.3.5}$$

Note that from these equations it follows that

$$[\delta_{\mu\nu} \Box + (D-2) \partial_\mu \partial_\nu]\, \partial \cdot \epsilon \;=\; 0. \tag{10.3.6}$$

The two previous equations imply that the third derivative of ϵ must vanish, so that ϵ can be at most a quadratic function of x. Hence we have the following cases:

- If ϵ is of zero order in x

$$\epsilon^\mu \;=\; a^\mu \tag{10.3.7}$$

  we recover the usual translation transformations.
- If ϵ is of first order in x, there are two different situations:

$$\epsilon^\mu \;=\; \omega^{\mu\nu}\, x^\nu \tag{10.3.8}$$

  with $\omega^{\mu\nu}$ an antisymmetric tensor $\omega^{\mu\nu} = -\omega^{\nu\mu}$, or

$$\epsilon^\mu \;=\; \lambda\, x^\nu. \tag{10.3.9}$$

  The first case corresponds to rotations while the second one is associated to the global dilatation.
- When ϵ is a quadratic function of x one has

$$\epsilon^\mu \;=\; b^\mu x^2 - 2x^\mu b\cdot x. \tag{10.3.10}$$

  Its finite form gives rise to the so-called *special conformal transformation*

$$\frac{x'^\mu}{x'^2} \;=\; \frac{x^\mu}{x^2} + b^\mu. \tag{10.3.11}$$

  This can be interpreted as the result of an inversion plus a translation.

Since in D-dimensional euclidean space there are D translation axes and $D(D-1)/2$ possible rotations, it is easy to see that the set of all conformal transformations forms a group, with the number of generators equal to

$$\frac{(D+1)(D+2)}{2}. \tag{10.3.12}$$

All these transformations can be characterized by a very simple geometrical property: they translate, rotate, or stretch the vectors placed at a given point but leave invariant their relative angle.

Constraints on correlation functions. It is instructive to study the constraints on the functional form of the scalar functions $G(x_1,\ldots,x_N)$ that are conformally invariant. To be invariant under translations and rotations, these functions can only depend on the relative distances $|x_i - x_j|$. Furthermore, to be invariant under a dilatation, the dependence on the distances can only be through their ratio, such as

$$\frac{|x_i - x_j|}{|x_k - x_l|}.$$

Since under the special conformal transformation (10.3.11) we have

$$|x'_i - x'_j| \;=\; \frac{|x_i - x_j|}{(1+2b\cdot x_i + b^2x_i^2)(1+2b\cdot x_j + b^2x_j^2)},$$

it is evident that, for $N \le 3$, it is impossible to define conformally invariant transformations of the distances and therefore, in this case, the only conformally invariant

functions with $N \leq 3$ variables are only the constants. On the contrary, for $N \geq 4$, by using four different points, we can define the so-called *harmonic ratios*

$$R_{ikmn} \equiv \frac{|x_i - x_k||x_m - x_n|}{|x_i - x_m||x_k - x_n|}. \tag{10.3.13}$$

These quantities are invariant under all conformal transformations. For N points the number of independent harmonic ratios is $N(N-3)/2$ and any arbitrary function of them is consequently conformally invariant.

10.3.2 Polyakov's Theorem

Let us now present the theorem due to Polyakov on the conformal symmetry of physical systems with local interactions that are invariant under translations, rotations, and a global dilatation. Its proof is simple.

Due to the locality of the theory, there exists a local field $T_{\mu\nu}(x)$, called the *stress–energy tensor*,[7] defined by the variation of the local action $S[\varphi]$ under the transformation (10.3.3)

$$\delta S = \frac{1}{(2\pi)^{D-1}} \int d^D x \, T_{\mu\nu}(x) \, \partial^\mu \epsilon^\nu(x). \tag{10.3.14}$$

Let's derive the equations fulfilled by the stress–energy tensor for the conformal invariance of the theory, expressed by the condition $\delta S = 0$. The translation invariance implies the conservation law

$$\partial_\mu T^{\mu\nu}(x) = 0, \tag{10.3.15}$$

obtained by integrating eqn (10.3.14) by parts and using the invariance of S under an arbitrary variation of the parameter a^μ in the expression (10.3.7) for ϵ. The rotational invariance implies that the stress–energy tensor is symmetric with respect to its indices

$$T^{\mu\nu}(x) = T^{\nu\mu}(x), \tag{10.3.16}$$

as can be seen by substituting in eqn (10.3.14) the transformation (10.3.8). The invariance under dilatation, given by the transformation (10.3.9), finally leads to a zero trace condition

$$T^\mu_\mu(x) = 0. \tag{10.3.17}$$

In view of eqns (10.3.15), (10.3.16), and (10.3.17), the action is automatically invariant under the conformal transformations that satisfy eqn (10.3.5).

10.4 Quasi-Primary Fields

The conformal invariance of a statistical system at a critical point permits us to prove a series of important results on the correlation functions of a special class of its operators.

[7] The factor $(2\pi)^{D-1}$, which is unusual with respect to the definition of this field as derived from Noether's theorem, is introduced here for later convenience.

These operators are called *quasi-primary* and in this section they will be denoted as $Q_n(x)$. They have the property of transforming under a generic conformal mapping as

$$Q_n(x) \;=\; \left|\frac{\partial x'}{\partial x}\right|^{d_n/D} Q_n(x'), \tag{10.4.1}$$

where $\left|\frac{\partial x'}{\partial x}\right|$ is the Jacobian of the mapping. For the transformations associated to the translations and rotations we have

$$\left|\frac{\partial x'}{\partial x}\right| \;=\; 1,$$

while for the dilatation and the special conformal transformations we have, respectively,

$$\left|\frac{\partial x'}{\partial x}\right| = \lambda^D, \quad \left|\frac{\partial x'}{\partial x}\right| = \frac{1}{(1+2b\cdot x+b^2x^2)^D}.$$

The transformation law (10.4.1) of the primary fields is obviously more specific than the simple scaling law (10.2.5) and therefore it imposes more restrictive conditions on the correlation functions of these fields. They satisfy the equation

$$\langle Q_1(x_1)Q_2(x_2)\dots Q_n(x_n)\rangle \;=\; \left|\frac{\partial x'}{\partial x}\right|^{d_1/D}_{x=x_1} \dots \left|\frac{\partial x'}{\partial x}\right|^{d_1/D}_{x=x_n} \langle Q_1(x_1')Q_2(x_2')\dots Q_n(x_n')\rangle. \tag{10.4.2}$$

Let us consider the two-point correlation functions

$$G^{(2)}_{ab}(x_1,x_2) \;=\; \langle Q_a(x_1)Q_b(x_2)\rangle. \tag{10.4.3}$$

Due to the translation and rotational invariance, it depends on the relative distance $\mid x_1 - x_2 \mid$. The dilatation invariance implies that $G^{(2)}_{ab}$ behaves as $\mid x_1 - x_2 \mid^{-d_a-d_b}$. Finally, using the invariance under the special conformal transformation (10.3.11), we arrive at the condition

$$\delta\, G^{(2)}_{ab} \;=\; -[d_a\,(b\cdot x_1) + d_b\,(b\cdot x_2)]\, G^{(2)}_{ab}. \tag{10.4.4}$$

This implies that $G^{(2)}_{ab}(\mid x_1 - x_2 \mid)$ is different from zero only when $d_a = d_b$. Hence, the two-point correlation functions of the quasi-primary fields satisfy an orthogonality condition. By an appropriate normalization of these fields, their general expression is then

$$\langle Q_a(x_1)Q_b(x_2)\rangle \;=\; \frac{\delta_{ab}}{\mid x_1 - x_2 \mid^{2d_a}}. \tag{10.4.5}$$

Consider now the three-point correlation functions of the quasi-primary fields

$$G^{(3)}_{abc}(x_1,x_2,x_3) \;=\; \langle Q_a(x_1)Q_b(x_2)Q_c(x_3)\rangle. \tag{10.4.6}$$

As usual, the invariance under translations and rotations implies that this correlator depends on the relative distances $x_{ij} \equiv \mid x_i - x_j \mid (i,j = 1,2,3)$. The invariance

under the infinitesimal transformations of eqn (10.3.11) gives rise to the homogeneous equations

$$\frac{1}{2}\sum_{i<j} x_{ij}\frac{\partial G^{(3)}}{\partial x_{ij}}\left[(b\cdot x_i)+(b\cdot x_j)\right] = -\sum_i d_i\,(b\cdot x_i)\,G^{(3)},$$

whose solution is given by

$$\langle Q_a(x_1)Q_b(x_2)Q_c(x_3)\rangle = \frac{C_{abc}}{x_{12}^{d_3-d_1-d_2}x_{13}^{d_2-d_1-d_3}x_{23}^{d_1-d_2-d_3}}. \tag{10.4.7}$$

The numerical constant C_{abc} cannot be fixed by the conformal invariance. However, using the orthogonality condition (10.4.5), it is easy to see that it coincides with the structure constant c_{abc} of the algebra (10.2.13), relative to these three fields.

In conclusion, thanks to the conformal invariance, one can determine the exact expression of the two- and three-point correlation functions of the quasi-primary operators. The same is not true for the higher point correlation functions. In this case, in fact, the N-point correlation functions can have an arbitrary dependence on the $N(N-3)/2$ harmonic ratios. For instance, the general form of the four-point function can be expressed as

$$G^{(4)}(x_1,x_2,x_3,x_4) = F\left(\frac{x_{12}x_{34}}{x_{13}x_{24}},\frac{x_{12}x_{34}}{x_{23}x_{41}}\right)\prod_{i<j}x_{ij}^{-(d_i+d_j)+d/2}, \tag{10.4.8}$$

with $d=\sum_{i=1}^4 d_i$ whereas F is an arbitrary function of the two independent harmonic ratios. For $D\neq 2$ and in the absence of additional information, the explicit expression of F cannot be determined. As we will see in the next sections, this goal can, however, be achieved when $D=2$.

10.5 Two-dimensional Conformal Transformations

Let's consider in more detail the conformal transformations in a euclidean space when $D=2$. In this case, something remarkable happens: in fact, eqns (10.3.5) become

$$\partial_1\epsilon_1 = \partial_2\epsilon_2, \quad \partial_1\epsilon_2 = -\partial_2\epsilon_1, \tag{10.5.1}$$

i.e. the Cauchy–Riemann conditions relative to the functions of a complex variable! In other words, if we introduce the notation

$$\begin{aligned}\epsilon(z) &= \epsilon_1 + i\,\epsilon_2\\ \bar\epsilon(\bar z) &= \epsilon_1 - i\,\epsilon_2\end{aligned} \tag{10.5.2}$$

and consider the new variables

$$\begin{aligned}z &= x_1 + i\,x_2\\ \bar z &= x_1 - i\,x_2\end{aligned}$$

in two dimensions the conformal transformations coincide with the analytic mappings of the complex plane

$$z\to f(z), \quad \bar z\to \bar f(\bar z). \tag{10.5.3}$$

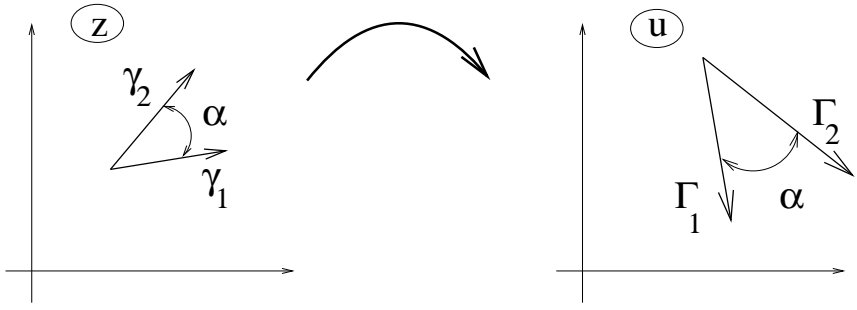


Fig. 10.4 *Conformal transformation of two vectors.*

It is well known that the analytic maps of the complex plane have the property of preserving the angle between two vectors placed at the same point. The proof is simple. Consider an analytic map $u = f(z)$ and suppose that $f'(z_0) \neq 0$, where z_0 is the intersection point of the two vectors γ_1 and γ_2, as shown in Fig. 10.4. We have

$$f'(z_0) \,=\, \lim_{\Delta z \to 0} \frac{\Delta u}{\Delta z} \,\equiv\, k\, e^{i\beta}.$$

Suppose now we choose the direction in which the variation Δz goes to zero along the curve γ_1, so that Δu moves along the transformed curve Γ_1. In this case we have

$$\beta \,=\, \arg f'(z_0) \,=\, \lim_{\Delta u \to 0} \arg \Delta u - \lim_{\Delta z \to 0} \arg \Delta z \,=\, \Phi_1 - \phi_1,$$

where ϕ_1 and Φ_1 are the angles of the curves γ_1 and Γ_1 with respect to the horizontal axis of the corresponding planes. However, the same quantity is obtained by taking the limit along the direction of the curve γ_2, since the derivative of a function of a complex variable is independent of the path along which it is computed. We have then

$$\beta \,=\, \Phi_2 - \phi_2,$$

where ϕ_2 and Φ_2 are respectively the angles of the curves γ_2 and Γ_2 with respect to the horizontal axis of the corresponding planes. Taking the difference between these two equations we have

$$\phi_2 - \phi_1 \,=\, \Phi_2 - \Phi_1 \,=\, \alpha.$$

Hence, the angle between two arbitrary curves that pass through the point z_0 at which $f'(z_0) \neq 0$ is equal to the angle between the two transformed curves. The modulus of the two vectors is instead dilated by the factor k.

As we will see later, in two-dimensional conformal field theories there is a natural splitting in the variables z and $\bar{z}$. To simplify the notation, in the following we focus

attention only on the transformations that involve the variable z, keeping in mind that the same analysis applies to the variable $\bar{z}$ as well. Moreover, thanks to the splitting of the two variables, it is useful to consider z and $\bar{z}$ as two independent quantities. This is equivalent to regarding the original coordinates (x_1, x_2) as variables of C^2 rather than of $\Re^2$ and to consider the new quantities $(z, \bar{z})$ simply as a change of variables. Obviously the original real variables are recovered imposing in C^2 the equation $z = \bar{z}$. As we will show below, this procedure permits a remarkable efficiency in the development of the formalism. From the definition of z and $\bar{z}$ we have

$$\begin{array}{ll} \partial_z = 1/2(\partial_1 - i\partial_2) & \partial_1 = (\partial_z + \partial_{\bar{z}}) \\ \partial_{\bar{z}} = 1/2(\partial_1 + i\partial_2) & \partial_2 = i\,(\partial_z - \partial_{\bar{z}}). \end{array} \tag{10.5.4}$$

In euclidean two-dimensional space, the line element is

$$ds^2 = g_{\mu\nu} dx^\mu\, dx^\nu = dx^2 + dy^2 = dz\, d\bar{z}.$$

The complex metric is then $g_{zz} = g_{\bar{z},\bar{z}} = 0$ and $g_{z\,\bar{z}} = g_{\bar{z}\,z} = \frac{1}{2}$ and therefore the components of the stress–energy tensor in these coordinates are given by

$$\begin{array}{l} T_{zz} \equiv T(z,\bar{z}) = \frac{1}{4}(T_{11} - T_{22} + 2iT_{12}), \\ T_{\bar{z}\bar{z}} \equiv \bar{T}(z,\bar{z}) = \frac{1}{4}(T_{11} - T_{22} - 2iT_{12}), \\ T_{z\bar{z}} = T_{\bar{z}z} \equiv \frac{1}{4}\,\Theta(z,\bar{z}) = \frac{1}{4}(T_{11} + T_{22}) = \frac{1}{4}T^\mu_\mu. \end{array} \tag{10.5.5}$$

In complex coordinates, the conservation law (10.3.15) becomes

$$\begin{array}{l} \partial_{\bar{z}}\, T(z,\bar{z}) + \frac{1}{4}\partial_z\, \Theta(z,\bar{z}) = 0; \\ \partial_z\, \bar{T}(z,\bar{z}) + \frac{1}{4}\partial_{\bar{z}}\, \Theta(z,\bar{z}) = 0. \end{array} \tag{10.5.6}$$

Furthermore, at the critical point the trace $\Theta(z,\bar{z})$ of the stress–energy tensor vanishes:

$$\partial_{\bar{z}} T(z,\bar{z}) = \partial_z \bar{T}(z,\bar{z}) = 0. \tag{10.5.7}$$

Hence, at criticality, T depends only on z while $\bar{T}$ depends on $\bar{z}$: the former is then an analytic operator while the latter is an anti-analytic one. These properties will be important in the sequel.

Let's now use the complex coordinates to investigate some aspects of two-dimensional conformal invariance. Imposing

$$z' = z + \epsilon(z), \quad \epsilon(z) = \sum_{n=-\infty}^{\infty} a_n z^{n+1},$$

every conformal transformation can be characterized by the coefficients a_n of the Taylor–Laurent expansion of $\epsilon(z)$ at the origin. An analogous result holds for the variable $\bar{z}$. As a result of these transformations, to first order in ϵ a scalar function $g(z,\bar{z})$ changes as

$$\begin{aligned} g(z',\bar{z}') - g(z,\bar{z}) \equiv \delta g &= -(\epsilon(z)\,\partial + \bar{\epsilon}(\bar{z})\,\bar{\partial})\, g(z,\bar{z}) \\ &= \sum_{n=-\infty}^{\infty} (a_n l_n + \bar{a}_n\, \bar{l}_n)\, g(z,\bar{z}), \end{aligned}$$

where we have introduced the operators

$$l_n \,=\, -z^{n+1}\,\partial, \quad \bar{l}_n \,=\, -\bar{z}^{n+1}\,\bar{\partial}, \tag{10.5.8}$$

with $\partial = \partial_z$ and $\bar{\partial} = \partial_{\bar{z}}$. They satisfy the commutation relations

$$\begin{array}{l} [l_n, l_m] \,= (n-m)\, l_{n+m}; \\ [\bar{l}_n, \bar{l}_m] = (n-m)\, \bar{l}_{n+m}; \\ [l_n, \bar{l}_m] = 0. \end{array} \tag{10.5.9}$$

The conformal symmetry is then expressed in terms of the direct sum of two isomorphic infinite dimensional algebras.

An apparent paradox. The above result may appear paradoxical: how can we have in two dimensions an infinite dimensional symmetry when, in the last section, we have established that the number of real parameters of the conformal symmetry is finite and equal, in D dimensions, to $(D+1)(D+2)/2$? In the light of this, for $D = 2$ we should have only six real parameter transformations and not an infinite number of them.

The explanation of this paradox is as follows: in two dimensions there is only *one* particular class of analytic functions that are well-defined and invertible on all of the Riemann sphere (i.e. the complex plane plus the point at infinity) (see Fig. 10.5). These are the Moebius transformations, given by

$$w(z) \,=\, \frac{az+b}{cz+d}, \tag{10.5.10}$$

with $a, b, c, d \in C$ and $ad - bc \neq 0$. Since a uniform rescaling of the parameters does not change the final expression of $w(z)$, we can always assume that they satisfy the condition $ad - bc = 1$. The set of all these transformations is isomorphic to the group

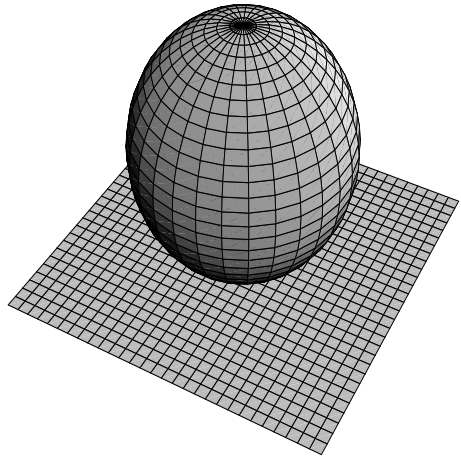


Fig. 10.5 *The Riemann sphere, alias the complex plane plus the point at infinity.*

$SL(2,C)/Z_2$, where the quotient with respect to Z_2 comes from the invariance of the mapping (10.5.10) under the multiplication of all the parameters for -1. Hence there are only three complex parameters, alias six real parameters, in agreement with the general result of the previous section.[8] The detailed study of these functions – an argument that is itself interesting – can be found in Appendix 10A. Here we simply note that a Moebius transformation is uniquely defined by the correspondence between three different points of the z-plane and three different points of the w-plane. If we denote the chosen three points of the z-plane z_1, z_2, and z_3 and the corresponding points of the w-plane w_1, w_2, and w_3, it is easy to prove the identity

$$\frac{(w_1 - w)(w_3 - w_2)}{(w_1 - w_2)(w_3 - w)} = \frac{(z_1 - z)(z_3 - z_2)}{(z_1 - z_2)(z_3 - z)}. \tag{10.5.11}$$

This relation is an implicit expression of the desired Moebius function. By solving eqn (10.5.11) for w, we get an explicit expression of the coefficients a, b, c, d of the Moebius transformation in terms of the given numbers z_1, z_2, z_3 and w_1, w_2, w_3. This result also shows that $w(z)$ depends at most on three complex parameters. Note that eqn (10.5.11) states that the harmonic ratio of four points of the complex plane is invariant under a Moebius transformation.

An equivalent way to achieve the same result consists of considering the vector field that generates the conformal transformations in the space of the functions $g(z)$:

$$v(z) = -\sum_n a_n \, l_n = \sum_n a_n \, z^{n+1} \, \partial_z.$$

Requiring that the vector field is regular near the origin, we have $a_n \neq 0$ only for $n \geq -1$. We can study its behavior in a neighborhood of the point at infinity by using the transformation $z = -1/y$

$$v(z) = \sum_n a_n \left(-\frac{1}{y}\right)^{n+1} \left(\frac{dz}{dy}\right)^{-1} \partial_y = \sum_n a_n \left(-\frac{1}{y}\right)^{n-1} \partial_y.$$

Requiring that the vector field is also regular at infinity, we have $a_n \neq 0$ only for $n \leq 1$. In conclusion, the only conformal transformations that are globally defined on the Riemann sphere are those expressed by a quadratic expression z. They correspond to $a_n l_n$ with $n = 0, \pm 1$, with three arbitrary complex parameters. The same result is obtained for the anti-analytic vector field. However the number of parameters does not double since the generators of the two algebras (10.5.9) that preserve the real section of C^2 are expressed by the linear combinations

$$l_n + \bar{l}_n, \quad i(l_n - \bar{l}_n). \tag{10.5.12}$$

Note that the generators $l_0, l_{\pm 1}$ form a finite subalgebra of (10.5.9), where $l_{-1} = -\partial_z$ is the generator of the translations, $l_0 = -z\partial_z$ the generator of the dilatations,

[8] The real projection of C^2 is obtained by taking (10.5.10) together with its complex conjugate $\bar{w}(\bar{z})$. Obviously this procedure does not enlarge the number of independent parameters.

and finally $l_1 = -z^2\partial_z$ the generator of the special conformal transformations. In particular, $l_0 + \bar{l}_0$ is the operator that generates the global dilatations on the real section of C^2, while $i(l_0 - \bar{l}_0)$ is the operator of the rotations.

To summarize, in two dimensions, it is necessary to distinguish between the *global* and the *local* conformal transformations: the former are given by the Moebius mappings whereas the latter are expressed by all other analytic functions. Note that, except for the Moebius transformations, any other analytic function is either not uniquely invertible or not well-defined on all of the complex plane.[9] For instance, the function $f(z) = z^3$ is defined on all the complex plane but is not uniquely invertible. Vice versa, $f(z) = \log z$ has a branch cut along the positive axis and jumps by $2\pi i$ each time we cross it. It is also worth recalling that, for the well-known Liouville theorem, an analytic function cannot be small everywhere unless it is a constant: one has to keep in mind this result for properly interpreting the Polyakov theorem that is based on the infinitesimal nature of the transformations.

The considerations above should clarify some basic aspects of conformal invariance in two dimensions but the question of how to interpret the apparent symmetry of the theory under the infinite dimensional class of local conformal transformations is still open. As we will see in the next sections, the answer to this question can be summarized as follows: the local conformal transformations $z \to f(z)$ present, at the quantum level, an anomaly. In particular, the two algebras (10.5.9) need a central extension for the fluctuations of a quantum field theory near the phase transition point. For this anomaly, the theory is invariant only under the Moebius transformations and not under all conformal transformations. But, even in the absence of an exact invariance of the physical systems under a generic analytic map, we will show that we can nevertheless control how the system behaves under a local conformal map and this proves to be the key to achieving the exact solution of many statistical models.

10.6 Ward Identity and Primary Fields

The correlation functions of the local fields, by the definition given by eqn (10.2.1), satisfy a Ward identity. To derive it, consider an arbitrary infinitesimal change of the coordinates

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x)$$

and denote by δA_i the corresponding variation of the fields under this transformation

$$A_i(x) \to A_i(x) + \delta A_i(x). \tag{10.6.1}$$

Under this change of variables, the action changes as

$$\delta S = \frac{1}{(2\pi)^{D-1}} \int d^D x \, T_{\mu\nu}(x) \, \partial^\mu \epsilon^\nu(x).$$

[9] Analytic functions are also called *holomorphic functions*. Analytic functions with a finite number of poles are called *meromorphic functions* whereas those having only zeros in any finite region of the complex plane are called *entire functions*.

Since a change of coordinates does not alter the description of the physical system, we have $\delta\langle A_1(x_1)\dots A_n(x_n)\rangle = 0$. This leads to the Ward identity

$$\sum_{i=1}^{n}\langle A_1(x_1)\dots\delta A_i(x_i)\dots A_n(x)\rangle - \frac{1}{(2\pi)^{D-1}}\int d^Dx\langle T_{\mu\nu}(x)A_1(x_1)\dots A_n(x_n)\rangle\,\partial^\mu\epsilon^\nu(x)$$
$$-\left[\frac{1}{(2\pi)^{D-1}}\int d^Dx\langle T_{\mu\nu}(x)\rangle\,\partial^\mu\epsilon^\nu(x)\right]\langle A_1(x_1)\dots A_n(x_n)\rangle \;=\; 0. \tag{10.6.2}$$

Let's now specialize this equation to the case of two-dimensional quantum field theory. First of all, let's define the *primary operators* $\phi_i(z,\bar z)$: they have the distinguishing property to change as follows

$$\phi_i(z,\bar z) \;=\; \left(\frac{df}{dz}\right)^{\Delta_i}\left(\frac{d\bar f}{d\bar z}\right)^{\bar\Delta_i}\phi_i'(z',\bar z'), \tag{10.6.3}$$

under an *arbitrary* conformal transformation $z\to z'=f(z)$, $\bar z\to\bar z'=\bar f(\bar z)$, where the two real quantities Δ_i and $\bar\Delta_i$ are the *conformal weights*. Note that the expression above can be written as

$$\phi_i'(z',\bar z')dz'^{\Delta_i}\,d\bar z'^{\bar\Delta_i} \;=\; \phi_i(z,\bar z)dz^{\Delta_i}\,d\bar z^{\bar\Delta_i}.$$

This enables us to consider the primary field $\phi_i(z,\bar z)$ as a "tensor" of degree $(\Delta_i,\bar\Delta_i)$ under the conformal transformations. Obviously a primary field is also a quasi-primary field but the reverse is not true, since the quasi-primary operators transform as "tensors" only under global conformal transformations (expressed, in two dimensions, by the Moebius mappings). To derive the infinitesimal transformation law of the primary fields, impose $z'=z+\epsilon(z)$ and $\bar z'=\bar z+\bar\epsilon(\bar z)$, and expand up to first order

$$\delta\phi_i \;\equiv\; \phi(z,\bar z)-\phi'(z',\bar z') \;=\; [(\Delta_i\,\partial\epsilon+\epsilon\partial)+(\bar\Delta_i\,\bar\partial\bar\epsilon+\bar\epsilon\bar\partial)]\,\phi_i(z,\bar z). \tag{10.6.4}$$

Consider the n-point correlation function $\langle\phi_1(x_1)\dots\phi_n(x_n)\rangle$ of the primary fields in the two-dimensional euclidean plane. Denote by S a circle of radius sufficiently large to include all the points $x_1,x_2,\dots,x_n$ as shown in Fig. 10.6. Suppose now we make an infinitesimal change of coordinates $x^\mu\to x^\mu+\epsilon^\mu(x)$, where $\epsilon^\mu(x)$ is a regular analytic function for $|z|\le R$ whereas, outside this region, only a differentiable function but rapidly decreasing to zero at infinity. $\epsilon^\mu(x)$ is then an arbitrarily small quantity in any region of the plane. Such a transformation is obviously controlled by the Ward identity (10.6.2). At the critical point of the theory $\langle T_{\mu\nu}(x)\rangle = 0$ and therefore we can neglect the last term in (10.6.2). For the second term, the only contribution to the integral comes from the region $|x|>R$ since, inside the circle, the theory is conformally invariant and we can apply the Polyakov theorem . If we denote generically by X the product of the n primary field, we have

$$\frac{1}{2\pi}\int d^2x\,\langle T_{\mu\nu}(x)\,X\rangle\partial^\mu\epsilon^\nu(x) \;=\; \frac{1}{2\pi}\int_{|x|>R}d^2x\,\langle T_{\mu\nu}(x)X\rangle\partial^\mu\epsilon^\nu(x). \tag{10.6.5}$$

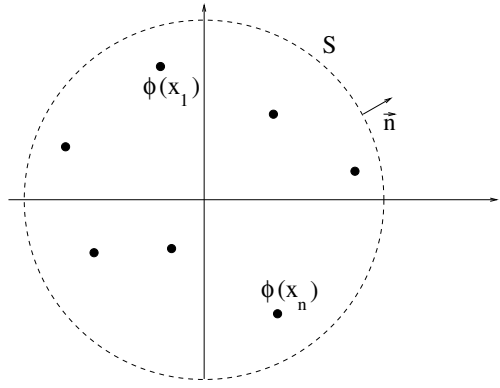


Fig. 10.6 *Correlation functions of the primary fields.*

Integrating by parts the right-hand side of this expression and neglecting the surface term at infinity (which vanishes by the rapidly decreasing behavior of ϵ^ν), we have

$$\frac{1}{2\pi}\int_{|x|>R} d^2x\,\langle T_{\mu\nu}(x)X\rangle\partial^\mu\epsilon^\nu(x) \;=\; -\frac{1}{2\pi}\int_{|x|>R} d^2x\,\epsilon^\nu\,\langle\partial^\mu T_{\mu\nu}(x)X\rangle + \frac{1}{2\pi}\int_C d\Sigma\, n^\mu\epsilon^\nu\,\langle T_{\mu\nu}X\rangle,$$

where n^μ is a unit vector orthogonal to the surface Σ of the circle C. The first term in the right-hand side vanishes by the conservation law of the stress–energy tensor. By using the complex coordinates $z, \bar{z}$ and the corresponding components of the stress–energy tensor, we can express the second term on the right-hand side as

$$\frac{1}{2\pi}\int_C d\Sigma\, n^\mu\epsilon^\nu\,\langle T_{\mu\nu}X\rangle \;=\; \frac{1}{2\pi i}\oint_C dz\epsilon(z)\,\langle T(z)X\rangle - \frac{1}{2\pi i}\oint_C d\bar{z}\;\bar{\epsilon}(\bar{z})\,\langle\bar{T}(\bar{z})X\rangle,$$

with $\epsilon(z) = \epsilon^1 + i\epsilon^2$ and $\bar{\epsilon}(\bar{z}) = \epsilon^1 - i\epsilon^2$.

Consider now the first term of the Ward identity (10.6.2) which, for the primary fields, is expressed by

$$\sum_{i=1}^{n}[(\Delta_i\partial_i\epsilon(z) + \epsilon(z)\partial_i) + (\bar{\Delta}_i\bar{\partial}_i\bar{\epsilon}(\bar{z}) + \bar{\epsilon}(\bar{z})\bar{\partial}_i)]\langle\phi_1(x_1)\dots\phi_n(x_n)\rangle.$$

We can use the complex coordinates $(z_i, \bar{z}_i)$ to identify the points in the plane and apply the Cauchy theorem to write the analytic and anti-analytic terms as

$$\sum_{i=1}^{n}[(\Delta_i\partial_i\epsilon(z) + \epsilon(z)\partial_i)]\langle\phi_1(x_1)\dots\phi_n(x_n)\rangle$$
$$= \frac{1}{2\pi i}\oint_C dz\epsilon(z)\sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2} + \frac{1}{z-z_i}\partial_i\right]\langle\phi_1(z_1,\bar{z}_1)\dots\phi_n(z_n,\bar{z}_n)\rangle,$$

$$\sum_{i=1}^{n}[(\bar{\Delta}_i\bar{\partial}_i\bar{\epsilon}(z)+\bar{\epsilon}(z)\bar{\partial}_i)]\langle\phi_1(x_1)\dots\phi_n(x_n)\rangle$$
$$= -\frac{1}{2\pi i}\oint_C d\bar{z}\,\bar{\epsilon}(z)\sum_{i=1}^{n}\left[\frac{\bar{\Delta}_i}{(\bar{z}-\bar{z}_i)^2}+\frac{1}{\bar{z}-\bar{z}_i}\bar{\partial}_i\right]\langle\phi_1(z_1,\bar{z}_1)\dots\phi_n(z_n,\bar{z}_n)\rangle.$$

Now putting together all the terms of the Ward identity, we arrive at

$$\frac{1}{2\pi i}\oint_C dz\,\epsilon(z)\left[\sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2}+\frac{1}{z-z_i}\partial_i\right]\langle\phi_1(z_1,\bar{z}_1)\dots\rangle-\langle T(z)\phi_1(z_1,\bar{z}_1)\dots\rangle\right]$$
$$-\frac{1}{2\pi i}\oint_C d\bar{z}\bar{\epsilon}(\bar{z})\left[\sum_{i=1}^{n}\left[\frac{\bar{\Delta}_i}{(\bar{z}-\bar{z}_i)^2}+\frac{1}{\bar{z}-\bar{z}_i}\bar{\partial}_i\right]\langle\phi_1(z_1,\bar{z}_1)\dots\rangle-\langle\bar{T}(\bar{z})\phi_1(z_1,\bar{z}_1)\dots\rangle\right]=0.$$

Since the two functions $\epsilon(z)$ and $\bar{\epsilon}(\bar{z})$ can be varied independently, the two terms of this equation must vanish separately. So, we have the conformal Ward identity for the analytic sector

$$\langle T(z)\phi_1(z_1,\bar{z}_1)\dots\rangle = \frac{1}{2\pi i}\oint_C dz\,\epsilon(z)\left[\sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2}+\frac{1}{z-z_i}\partial_i\right]\langle\phi_1(z_1,\bar{z}_1)\dots\rangle\right], \tag{10.6.6}$$

and an analogous one for the anti-analytic sector

$$\langle\bar{T}(\bar{z})\phi_1(z_1,\bar{z}_1)\dots\rangle = \frac{1}{2\pi i}\oint_C d\bar{z}\bar{\epsilon}(\bar{z})\left[\sum_{i=1}^{n}\left[\frac{\bar{\Delta}_i}{(\bar{z}-\bar{z}_i)^2}+\frac{1}{\bar{z}-\bar{z}_i}\bar{\partial}_i\right]\langle\phi_1(z_1,\bar{z}_1)\dots\rangle\right]. \tag{10.6.7}$$

The two sectors are decoupled. We can use these Ward identities and the Cauchy theorem to extract the singular terms of the OPE of the primary field $\phi_i(z,\bar{z})$ with the analytic component $T(z)$ of the stress–energy tensor:

$$T(z_1)\phi_i(z_2,\bar{z}_2) = \frac{\Delta_i}{(z_1-z_2)^2}\phi_i(z_2,\bar{z}_2)+\frac{1}{z_1-z_2}\partial\phi(z_2,\bar{z}_2)+\text{regular terms.} \tag{10.6.8}$$

An analogous result holds for the anti-analytic part:

$$\bar{T}(\bar{z}_1)\phi_i(z_2,\bar{z}_2) = \frac{\bar{\Delta}_i}{(\bar{z}_1-\bar{z}_2)^2}\phi_i(z_2,\bar{z}_2)+\frac{1}{\bar{z}_1-\bar{z}_2}\partial\phi(z_2,\bar{z}_2)+\text{regular terms.} \tag{10.6.9}$$

Notice that, for what concerns the Ward identity, the primary field $\phi_i(z,\bar{z})$ may be regarded as made up of a product of two *chiral primary fields* $\Phi(z)$ and $\bar{\Phi}(\bar{z})$, the first depending only on z while the second only on $\bar{z}$, $\phi_i(z,\bar{z}) = \Phi_i(z)\,\bar{\Phi}_i(\bar{z})$. This factorization is extremely useful to deal with the algebraic properties of the primary fields but is not a faithful representation of the actual nature of the primary fields. As we shall show later, the correlation functions of the primary field $\phi_i(z,\bar{z})$ are not simply given by the product of the correlation functions of the chiral primary fields $\Phi_i(z)$ and $\bar{\Phi}(\bar{z})$.

The primary fields play an important role in conformal field theory. As a matter of fact, their transformation law (10.6.3) is the simplest possible and leads to an operator product expansion with $T(z)$ and $\bar{T}(\bar{z})$ in which there are at most second-order poles. Any other field of the theory has an OPE with the stress–energy tensor of higher order poles: to see this, it is sufficient to consider the operator product expansion of $T(z)$ with a derivative of the primary field. In addition to the simplicity of their operator product expansion, the primary fields are also the building blocks of the representation theory of conformal symmetry. As we will show in the following sections, *all* conformal fields of the theory are organized in *conformal families* that are uniquely identified by the primary fields. These families form irreducible representations of the quantum version of the conformal algebra.

10.7 Central Charge and Virasoro Algebra

In this section we analyze the quantum version of the conformal algebra that, as we will see, is deeply related to the stress–energy tensor. First of all, it is necessary to note that the role played by the stress–energy tensor in the theory is twofold: on one hand, it is the generator of the conformal transformations; on the other hand it is a conformal field itself. Since it satisfies a conservation law, its scaling dimension coincides with its canonical dimension, equal to $d_T = d_{\bar{T}} = 2$. The two-point correlation function of its analytic part is generically different from zero and can be expressed as

$$\langle T(z_1)T(z_2)\rangle \;=\; \frac{c/2}{(z_1 - z_2)^4}, \tag{10.7.1}$$

where the real coefficient c is the *central charge* of the conformal algebra. The same holds for the anti-analyitic component

$$\langle \bar{T}(\bar{z}_1)\bar{T}(\bar{z}_2)\rangle \;=\; \frac{\bar{c}/2}{(\bar{z}_1 - \bar{z}_2)^4}. \tag{10.7.2}$$

For a relativistic and parity invariant theory, it is easy to show that $c = \bar{c}$. From now on we focus attention only on $T(z)$, keeping in mind that the same results will also hold for $\bar{T}(\bar{z})$. The quantity c is generally different from zero, as can be seen by the analysis of two simple but significant examples of conformal field theories.

10.7.1 Example 1. Free Neutral Fermion

Consider the lagrangian of a neutral bidimensional fermion (Majorana fermion)

$$\mathcal{L} \;=\; \frac{\lambda}{2\pi}\left[\psi\frac{\partial}{\partial\bar{z}}\psi + \bar{\psi}\frac{\partial}{\partial z}\bar{\psi}\right].$$

The equations of motion are

$$\frac{\partial}{\partial\bar{z}}\psi \;=\; \frac{\partial}{\partial z}\bar{\psi} \;=\; 0.$$

Hence $\psi(z)$ is a purely analytic field (with conformal weight $\Delta = \frac{1}{2}$, as can be easily seen directly from the lagrangian) while $\bar{\psi}(\bar{z})$ is a purely anti-analytic field

with conformal weight $\bar{\Delta} = \frac{1}{2}$. Their two-point correlation functions are

$$\begin{aligned} \langle \psi(z_1)\psi(z_2) \rangle &= \frac{1}{\lambda} \frac{1}{z_1 - z_2}; \\ \langle \bar{\psi}(\bar{z}_1)\bar{\psi}(\bar{z}_2) \rangle &= \frac{1}{\lambda} \frac{1}{\bar{z}_1 - \bar{z}_2}. \end{aligned} \tag{10.7.3}$$

The analytic part of the stress–energy tensor is obtained by Noether's theorem:

$$T(z) = -\frac{\lambda}{2} : \psi(z) \frac{\partial}{\partial z} \psi(z) : . \tag{10.7.4}$$

The two-point correlator of $T(z)$ can be obtained by the correlator (10.7.3) applying Wick's theorem

$$\begin{aligned} \langle T(z_1)T(z_2) \rangle &= \frac{\lambda^2}{4} \langle : \psi(z_1)\partial_1\psi(z_1) :: \psi(z_2)\partial_2\psi(z_2) : \rangle \\ &= \frac{\lambda^2}{4} \left[\langle \psi(z_1)\partial_2\psi(z_2) \rangle \langle \partial_1\psi(z_1)\psi(z_2) \rangle - \langle \psi(z_1)\psi(z_2) \rangle \langle \partial_1\psi(z_1)\partial_2\psi(z_2) \rangle \right] \\ &= \frac{1}{4} \left[-\frac{1}{(z_1 - z_2)^4} + \frac{2}{(z_1 - z_2)^4} \right] \\ &= \frac{1}{4(z_1 - z_2)^4}. \end{aligned} \tag{10.7.5}$$

For this system we then have $c = \frac{1}{2}$.

10.7.2 Example 2. Free Bosonic Field

Consider now the lagrangian of a neutral free boson

$$\mathcal{L} = \frac{g}{4\pi} (\partial_\mu \Phi)^2.$$

The correlation function of this field is given by

$$G(z, \bar{z}) = \langle \Phi(z_1, \bar{z}_1)\Phi(z_2, \bar{z}_2) \rangle = -\frac{1}{2g} \log z_{12} - \frac{1}{2g} \log \bar{z}_{12}. \tag{10.7.6}$$

Note that the free bosonic field $\Phi(z, \bar{z})$ is not a scaling operator itself. However we can construct scaling operators as the fields $\partial_z \Phi$ or $e^{i\alpha\Phi}$ using the field Φ. For the analytic part of the stress–energy tensor, derived from Noether's theorem, we have

$$T(z) = -g : (\partial_z \Phi)^2 : \tag{10.7.7}$$

Its two-point correlation function can be computed by (10.7.6) using Wick's theorem

$$\begin{aligned} \langle T(z_1)T(z_2) \rangle &= g^2 \langle : (\partial_1 \Phi(z_1))^2 :: (\partial_2 \Phi(z_2) :)^2 \rangle \\ &= g^2 \left[2 \left(\langle \partial_1 \Phi(z_1) \partial_2 \Phi(z_2) \rangle \right)^2 \right] \\ &= \frac{1}{2(z_1 - z_2)^4}. \end{aligned}$$

Therefore the central charge of this system is $c = 1$.

Conformal anomaly. Let's come back to the general discussion on the stress–energy tensor. In the presence of a central charge different from zero, the OPE of $T(z)$ with itself has the singular terms

$$T(z_1)T(z_2) \;=\; \frac{c/2}{(z_1-z_2)^4} + \frac{2}{(z_1-z_2)^2}T(z_2) + \frac{1}{z_1-z_2}\partial T(z_2) + \cdots \tag{10.7.8}$$

Therefore, the infinitesimal conformal transformation of $T(z)$ is given by

$$\delta T(z) \;=\; (2\partial\epsilon + \epsilon\partial)T(z) + \frac{c}{12}\partial^3\epsilon(z). \tag{10.7.9}$$

The term proportional to c may be interpreted as a quantum anomaly. Consider, in fact, the Ward identity for the one-point function of this operator

$$\delta\langle T(w)\rangle \;=\; \frac{1}{2\pi i}\oint dz\,\epsilon(z)\,\langle T(z)T(w)\rangle \;=\; \frac{c}{12}\partial^3\epsilon(z),$$

where, in the last line, we used the expression (10.7.1) of the correlator and then the Cauchy theorem. This term is obviously zero for the *global* conformal transformations[10] but is instead different from zero for all the *local* conformal mappings. This means that, passing from the euclidean plane in which $\langle T(z)\rangle_{piano} = 0$ to another geometry with the conformal transformation $z \to f(z)$, in the new geometrical domain the energy density is different from zero! It is for this reason that the central charge is also called a "conformal anomaly". It is also called a "trace anomaly", because in a conformal field theory defined on a curved manifold it is no longer true that the trace Θ of the stress–energy tensor vanishes as a consequence of the scaling invariance of the theory: in fact the curvature R of the manifold introduces a length-scale and in this case it is possible to prove that

$$\langle\Theta\rangle \;=\; -\frac{c}{12}R. \tag{10.7.10}$$

A non-zero value of c gives rise to a measurable physical effect, known as the *Casimir effect*. This will be analyzed in Section 9, using the main properties of $T(z)$ that we are going to discuss.

Properties of the stress–energy tensor. The first property is its transformation law under a local conformal mapping $z \to \eta$

$$T(z) \;=\; T(\eta)\left(\frac{d\eta}{dz}\right)^2 + \frac{c}{12}\{\eta, z\}, \tag{10.7.11}$$

where the last term is the *Schwartz derivative*

$$\{\eta, z\} \;\equiv\; \frac{d^3\eta/dz^3}{d\eta/dz} - \frac{3}{2}\left(\frac{d^2\eta/dz^2}{d\eta/dz}\right)^2. \tag{10.7.12}$$

[10] For these transformations $\epsilon(z)$ is at most quadratic in z.

The second important aspect of the stress–energy tensor is its Taylor–Laurent expansion (say, around the origin) in terms of the operators L_n

$$T(z) = \sum_{-\infty}^{\infty} \frac{L_n}{z^{n+2}}. \tag{10.7.13}$$

The Schwartz derivative. The Schwartz derivative of a function of a complex variable has the following properties:

1. $\{\eta, z\} = 0$ if and only if $\eta(z)$ is a Moebius transformation $\eta(z) = \frac{az+b}{cz+d}$;
2. it satisfies

$$\left\{\frac{a\eta+b}{c\eta+d}, z\right\} = \{\eta, z\}$$
$$\left\{\eta, \frac{az+b}{cz+d}\right\} = \{\eta, z\}\,(cz+d)^4;$$

3. under the sequence of transformations $z \to \eta \to \gamma$ one has

$$\{\gamma, z\} = \{\gamma, \eta\}\left(\frac{d\eta}{dz}\right)^2 + \{\eta, z\}.$$

The last equation ensures the correct transformation properties of the stress–energy tensor. In fact, for the two individual mappings we have

$$T(z) = T(\eta)\left(\frac{d\eta}{dz}\right)^2 + \frac{c}{12}\{\eta, z\}$$
$$T(\eta) = T(\gamma)\left(\frac{d\gamma}{d\eta}\right)^2 + \frac{c}{12}\{\gamma, \eta, \}$$

and therefore, substituting the second of these equations into the first

$$\begin{aligned} T(z) &= \left[T(\gamma)\left(\frac{d\gamma}{d\eta}\right)^2 + \frac{c}{12}\{\gamma, \eta, \}\right]\left(\frac{d\eta}{dz}\right)^2 + \frac{c}{12}\{\eta, z\} \\ &= T(\gamma)\left(\frac{d\gamma}{dz}\right)^2 + \frac{c}{12}\{\gamma, z\}. \end{aligned}$$

Their action on a generic conformal field $A(z, \bar{z})$ is defined as follows

$$L_n\, A(z_1, \bar{z}_1) = \frac{1}{2\pi i}\oint_{C_1} dz\,(z - z_1)^{n+1}\, T(z)\, A(z_1, \bar{z}_1), \tag{10.7.14}$$

where C_1 is a closed contour around the point z_1. In other words, the application of L_n to $A(z_1, \bar{z}_1)$ filters the conformal field that appears in front of the power $(z - z_1)^{-n}$

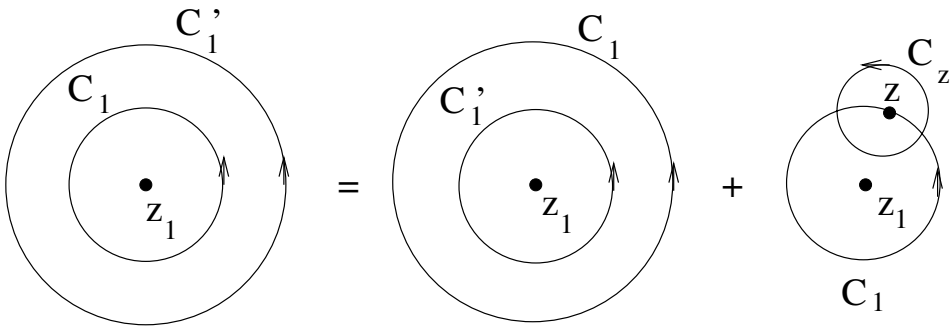


Fig. 10.7 *Exchange of the integration contours for the commutator* $[L_n, L_m]$.

in the operator product expansion of $T(z)$ with $A(z_1, \bar{z}_1)$. The application of two of these operators is given by

$$L_n\, L_m\, A(z_1, \bar{z}_1) \;=\; \left(\frac{1}{2\pi i}\right)^2 \oint_{C_1'} dz' \oint_{C_1} dz\, (z' - z_1)^{n+1}\, (z - z_1)^{m+1}\, T(z')T(z)\, A(z_1, \bar{z}_1), \tag{10.7.15}$$

where both the two contours C_1' and C_1 have the point z_1 inside, with C_1'' external to C_1.

It is interesting to note that the operator expansion (10.7.8) can be equivalently expressed in terms of the commutator $[L_n, L_m]$. To compute this quantity, we need to exchange the two integration contours, paying attention to the singular terms of the OPE encountered in this exchange. The situation is graphically given in Fig. 10.7: it involves[11] an integral over the contour C_z around the singularities and an integral over the contour C_1 of the point z_1

$$\begin{aligned}
[L_n, L_m] \;&=\; \left(\frac{1}{2\pi i}\right)^2 \oint_{C_z} dz' \oint_{C_1} dz\, (z' - z_1)^{n+1}\, (z - z_1)^{m+1}\, T(z')T(z) \\
&= \left(\frac{1}{2\pi i}\right)^2 \oint_{C_z} dz' \oint_{C_1} dz\, (z' - z_1)^{n+1}\, (z - z_1)^{m+1} \\
&\quad \times \left(\frac{c/2}{(z' - z)^4} + \frac{2}{(z' - z)^2} T(z) + \frac{1}{z' - z} \partial T(z)\right).
\end{aligned}$$

[11] In the integrals we have omitted the field $A(z_1, \bar{z}_1)$ since it appears on both members of the equation.

Let's consider separately the results of integration of each term. For the first term we have

$$\begin{aligned}
&\frac{c}{2}\left(\frac{1}{2\pi i}\right)^2 \oint_{C_z} dz' \oint_{C_1} dz\, \frac{(z'-z_1)^{n+1}\,(z-z_1)^{m+1}}{(z'-z)^4} \\
&= \frac{c}{2\cdot 3!}(n+1)n(n-1)\,\frac{1}{2\pi i}\oint_{C_1} dz(z-z_1)^{n+m-1} \\
&= \frac{c}{12}n(n^2-1)\,\delta_{n+m,0}.
\end{aligned}$$

For the second term

$$\begin{aligned}
&2\left(\frac{1}{2\pi i}\right)^2 \oint_{C_z} dz' \oint_{C_1} dz\, \frac{(z'-z_1)^{n+1}\,(z-z_1)^{m+1}}{(z'-z)^2}\,T(z) \\
&= 2\,(n+1)\,\frac{1}{2\pi i}\oint_{C_1} dz(z-z_1)^{n+m+1}\,T(z) \\
&= 2(n+1)\,L_{n+m}.
\end{aligned}$$

For the last term

$$\begin{aligned}
&\left(\frac{1}{2\pi i}\right)^2 \oint_{C_z} dz' \oint_{C_1} dz\, \frac{(z'-z_1)^{n+1}\,(z-z_1)^{m+1}}{z'-z}\,\partial T(z) \\
&= \frac{1}{2\pi i}\oint_{C_1} dz(z-z_1)^{n+m+2}\,\partial T(z) \\
&= -\frac{1}{2\pi i}\oint_{C_1} dz\partial(z-z_1)^{n+m+2}\,T(z) - (n+m+2)\,L_{n+m}.
\end{aligned}$$

Now putting together all the expressions above and keeping in mind that analogous results hold for the anti-analytic part of the stress–energy tensor, we arrive at the commutation relations

$$\begin{aligned}
[L_n, L_m] &= (n-m)\,L_{n+m} + \frac{c}{12}n(n^2-1)\,\delta_{n+m,0}, \\
\left[\bar{L}_n, \bar{L}_m\right] &= (n-m)\,\bar{L}_{n+m} + \frac{c}{12}n(n^2-1)\,\delta_{n+m,0}, \\
\left[L_n, \bar{L}_m\right] &= 0.
\end{aligned} \tag{10.7.16}$$

These relations define the so-called *Virasoro algebra*: it provides the quantum version of the classical conformal algebra (10.5.9) and the two coincide when $c = 0$.

An important remark. As a result of this analysis we have achieved a very important conceptual point that is worth emphasizing: in two dimensions the problem of classifying all possible universality classes of critical phenomena simply consists of identifying all irreducible representations of the Virasoro algebra. From this point of view, the numerous variety of critical phenomena is on the same footing as the different behavior of the irreducible representations of the rotation group where, according to the value of the angular momentum, the phenomenology is different but the underlying algebraic structure is the same.

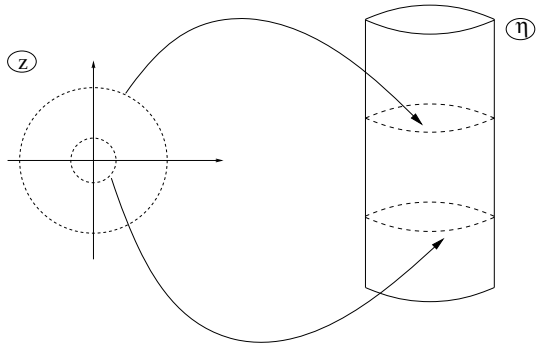


Fig. 10.8 *Conformal transformation of the complex plane onto the cylinder. Circles of different radius are mapped onto different sections of the cylinder.*

10.8 Representation Theory

Let us discuss the representations of the Virasoro algebra. They can be equivalently analyzed in terms of conformal fields or in terms of vectors in a Hilbert space. There is in fact an isomorphism between the two pictures that can be established as follows.

Radial quantization. Consider the conformal transformation

$$\eta \,=\, \frac{L}{2\pi}\,\log z \tag{10.8.1}$$

that maps the entire complex plane into the infinite cylinder strip of width L, as can be seen by writing $\eta = \tau + i\sigma$ and expressing z as $z = \rho e^{i\alpha}$ (see Fig. 10.8)

$$\tau \,=\, \frac{L}{2\pi}\log\rho, \sigma \,=\, \frac{L}{2\pi}\,\alpha.$$

Circles of the z-plane are mapped in orthogonal sections of the cylinder. In particular, the origin is transformed in the section of the cylinder placed at $-\infty$, whereas the point at infinity of the z-plane is mapped in the section of the cylinder at $+\infty$. For this reason, the map (10.8.1) gives rise to the so-called *radial quantization scheme* in which the longitudinal direction of the cylinder is regarded as *time*, while the transverse direction as (compactified) *space*. Circles in the z-plane correspond to surfaces of equal time on the cylinder. Note that the temporal inversion $\tau \to -\tau$ is implemented by the transformation $z \to 1/\bar{z}$.

In the radial quantization scheme we can introduce the R-ordered product of the fields, analogously to the usual T-ordered product

$$R[\phi_1(z)\phi_2(w)] \,=\, \begin{cases} \phi_1(z)\phi_2(w)\ , \text{ if } |z| < |w| \\ \phi_2(w)\phi_1(z)\ , \text{ if } |w| < |z|. \end{cases} \tag{10.8.2}$$

We can also relate their operator product expansion with the commutation relations, as we have already seen for the Virasoro generators. To this aim, let $\beta(z)$ and $\gamma(z)$ be

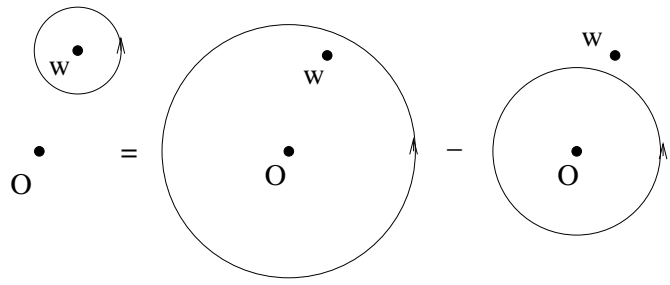


Fig. 10.9 *Commutator as the difference of circular contours.*

two analytic fields and consider the integral

$$\oint_w \beta(z)\,\gamma(w)\,dz, \tag{10.8.3}$$

around the point w, taken in an anti-clockwise direction. By using the radial quantization of these operators, (10.8.3) can be expressed as a difference of integrals computed along the circular contours of radius $|w| \pm \epsilon$, as shown in Fig. 10.9. These contours correspond to two slightly different time instants and therefore

$$\oint_w \beta(z)\,\gamma(w)\,dz \;=\; \oint_{C_1} \beta(z)\gamma(w)\,dz - \oint_{C_2} \gamma(w)\beta(z)dz \;=\; [\Gamma, \beta(w)]\,, \tag{10.8.4}$$

where the operator Γ is given by the integral

$$\Gamma \;=\; \oint \gamma(z)\,dz$$

taken along a circle around the origin. In the limit $\epsilon \to 0$ the commutator so obtained corresponds to the usual equal time commutator of quantum field theory.

Equation (10.8.4) allows us to compute the commutator of the generators of the Virasoro algebra with any primary field of conformal weight Δ. Using eqn (10.6.8), we have

$$\begin{aligned}
[L_n, \phi(w,\bar{w})] &= \frac{1}{2\pi i}\oint_w dz\, z^{n+1}\, T(z)\phi(w,\bar{w}) \\
&= \frac{1}{2\pi i}\oint_w dz\, z^{n+1}\left[\frac{\Delta\,\phi(w,\bar{w})}{(z-w)^2} + \frac{\partial\phi(w,\bar{w})}{z-w} + \dots\right] \\
&= \Delta(n+1)w^n\phi(w,\bar{w}) + w^{n+1}\partial\phi(w,\bar{w}).
\end{aligned} \tag{10.8.5}$$

A similar expression holds for the anti-analytic generators $\bar{L}_n$.

Hilbert space of conformal states. In the cylinder geometry it is possible to introduce a Hilbert space and a hamiltonian H that will rule the (euclidean) time evolution

of the states. The explicit form of H will be given later. H also determines the time evolution of the fields $A(\sigma, \tau)$

$$A(\sigma, \tau) \;=\; e^{H\tau}\, A(\sigma, 0)\, e^{-H\tau}. \tag{10.8.6}$$

To construct the Hilbert space we assume firstly the existence of a vacuum state $\mid 0\rangle$. Any other state of this space can be constructed by acting on the vacuum state with certain operators, as it happens for the creation operators of usual quantum field theories. The initial states are those at $t \to -\infty$ and, thanks to the conformal mapping on the cylinder, they can be defined as

$$\mid A_{in}\rangle \;=\; \lim_{z,\bar{z}\to 0} A(z, \bar{z}) \mid 0\rangle. \tag{10.8.7}$$

To introduce the final state, we need to define the adjoint operator of a conformal operator, here given by

$$[A(z, \bar{z})]^{\dagger} \;=\; A\left(\frac{1}{\bar{z}}, \frac{1}{z}\right) \frac{1}{z^{2\bar{\Delta}}} \frac{1}{\bar{z}^{2\Delta}}. \tag{10.8.8}$$

The reason for this definition lies in the relationship that exists between the usual definition of the adjoint operator in the present formulation and in the formulation done in Minkowski space: the factor i that is present in the Minkowski formulation and is instead absent in the time evolution (10.8.6) must be compensated by an explicit transformation of the time inversion $\tau \to -\tau$. The other extra factors in (10.8.8) are necessary to preserve the transformation properties of the adjoint operator under the conformal transformations. In fact, parameterize the point at infinity in terms of the map $w = 1/z$ and denote by $\hat{A}(w, \bar{w})$ the operator in these new coordinates. It is natural to impose

$$\langle A_{fin} \mid = \lim_{w,\bar{w}\to 0} \langle 0 \mid \hat{A}(w, \bar{w}). \tag{10.8.9}$$

For the primary and quasi-primary fields, it is now possible to link $\hat{A}(w, \bar{w})$ to $A(z, \bar{z})$ since

$$\hat{A}(w, \bar{w}) \;=\; A(z, \bar{z}) \left(\frac{\partial z}{\partial w}\right)^{\Delta} \left(\frac{\bar{\partial}\, \bar{z}}{\bar{\partial}\, \bar{w}}\right)^{\bar{\Delta}} \;=\; A\left(\frac{1}{w}, \frac{1}{\bar{w}}\right) (-w^{-2})^{\Delta}\, (-\bar{w}^{-2})^{\bar{\Delta}},$$

and therefore

$$\begin{aligned} \langle A_{fin} \mid \;&=\; \lim_{w,\bar{w}\to 0} \langle\, 0 \mid \hat{A}(w, \bar{w}) \;=\; \lim_{z,\bar{z}\to 0} \langle\, 0 \mid A\left(\frac{1}{z}, \frac{1}{\bar{z}}\right) \frac{1}{z^{2\Delta}} \frac{1}{\bar{z}^{2\bar{\Delta}}} \\ &=\; \lim_{z,\bar{z}\to 0} \langle\, 0 \mid [A(z, \bar{z})]^{\dagger} \;=\; \left[\lim_{z,\bar{z}\to 0} A(\bar{z}, z) \mid 0\rangle\right] = \mid A_{in}\rangle^{\dagger}. \end{aligned}$$

The final states can thus be defined as

$$\langle A_{fin} \mid \equiv \lim_{z,\bar{z}\to\infty} \langle 0 \mid A(z, \bar{z})\, z^{2\Delta}\, \bar{z}^{2\bar{\Delta}}. \tag{10.8.10}$$

For the stress–energy tensor, the definition (10.8.8) of the adjoint operator implies the equality of the expressions

$$T^\dagger(z) \,=\, \sum_{-\infty}^{\infty} \frac{L_n^\dagger}{\bar{z}^{n+2}} \quad \text{and} \quad T\left(\frac{1}{\bar{z}}\right)\,\frac{1}{\bar{z}^4} \,=\, \sum_{-\infty}^{\infty} \frac{L_n}{\bar{z}^{-n+2}},$$

namely

$$L_n^\dagger \,=\, L_{-n}. \tag{10.8.11}$$

An analogous formula holds for the generators $\bar{L}_n$ of $\bar{T}(\bar{z})$. Applying now $T(z)$ ($\bar{T}(\bar{z})$) to the vacuum state

$$T(z)\,|0\rangle \,=\, \sum_{-\infty}^{\infty} \frac{L_n}{z^{n+2}}\,|0\rangle,$$

and demanding their regularity at the origin, we arrive at the conditions that identify this state

$$\begin{array}{l} L_n\,|0\rangle \,=\, 0, \quad n \geq -1, \\ \bar{L}_n\,|0\rangle \,=\, 0, \quad n \geq -1. \end{array} \tag{10.8.12}$$

In particular, the conditions $L_{0,\pm 1}|0\rangle = 0$ and $\bar{L}_{0,\pm 1}|0\rangle = 0$ establish that the vacuum state is invariant under $SL(2,C)$ transformations: the vacuum state is the same for all the observers related by the global conformal transformations. Moreover, these relations imply that the vacuum expectation values of T and $\bar{T}$ vanish:

$$\langle 0 \mid T(z) \mid 0 \rangle \,=\, \langle 0 \mid \bar{T}(\bar{z}) \mid 0 \rangle \,=\, 0. \tag{10.8.13}$$

10.8.1 Representation Theory: The Space of the Conformal States

For the sake of simplicity, let us focus our attention only on the analytic sector of the theory (similar results hold for the anti-analytic one). Consider the state created by the analytic component of the primary field $\phi_\Delta(z)$ of conformal weight Δ

$$\mid \Delta \,\rangle \,\equiv\, \phi_\Delta(0) \,\mid\, 0\,\rangle. \tag{10.8.14}$$

Using the operator product expansion (10.6.8) and the definition (10.7.13) of $T(z)$ in terms of the L_n's, it is easy to show that

$$\begin{array}{l} L_0 \,\mid \Delta\rangle \,=\, \Delta \,\mid \Delta\rangle \\ L_n \,\mid \Delta\rangle \,=\, 0, \quad n > 0. \end{array} \tag{10.8.15}$$

Hence, $\mid \Delta \,\rangle$ is an eigenstate of L_0 (with eigenvalue Δ). It can be normalized as

$$\langle \Delta | \Delta \rangle \,=\, 1. \tag{10.8.16}$$

Consider now the *descendant states* of $\mid \Delta\rangle$, i.e. the states that are obtained by acting on $\mid \Delta\rangle$ by the operators L_n with a negative index. To avoid an over-counting of these

states,[12] it is convenient to introduce an ordering, for instance

$$\begin{gathered} \mid \Delta; n_1, n_2, \ldots n_k\rangle \equiv L_{-n_1} L_{-n_2} \ldots L_{-n_k} \mid \Delta\rangle \\ n_1 \leq n_2 \leq \ldots \leq n_k. \end{gathered} \tag{10.8.17}$$

Using the commutation relations of the Virasoro algebra, we have

$$L_0 \mid \Delta; n_1, n_2, \ldots n_k\rangle = (\Delta + N) \mid \Delta; n_1, n_2, \ldots n_k\rangle, \quad N = \sum_{i=1}^{k} n_i. \tag{10.8.18}$$

This shows that the descendant states are also eigenstates of L_0 with an eigenvalue related to their level N. The negative modes L_{-m} of the Virasoro algebra behave then as the creation operators of the familiar quantum harmonic oscillator, the only difference is that they move by m the eigenvalues of the state they act on. This situation is graphically represented in Fig. 10.10.

Structure of the Hilbert space. The Hilbert space of the conformal states has a nested structure. To reach the level N, for instance, we can act directly with the operator L_{-N} on the state $\mid \Delta\rangle$ or we can act on any descendant of a level M ($M < N$) with L_{M-N} or with any other ordered sequence of operators $L_{-n_j} \ldots L_{-n_k}$ that satisfy the condition $\sum_{i=j}^{k} n_i = M - N$. This nested structure gives rise to an exponential growth of the dimensions of the L_0-eigenspaces. These dimensions can be computed noting that the problem consists of determining in how many ways a positive integer number N can be expressed as sum of all possible integer numbers less that it. This combinatorial problem can be solved in terms of the generating function

$$f(q) = \prod_{n=1}^{\infty} \frac{1}{1 - q^n}. \tag{10.8.19}$$

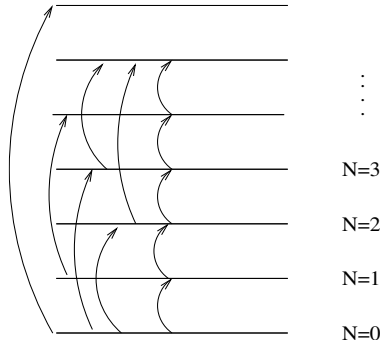


Fig. 10.10 *Levels of different N and action of the operators L_{-m}.*

[12]For the commutation relations of the Virasoro algebra, any other ordering can be expressed as a linear combination of the states given in the text.

Denoting by $P(N)$ the dimension of the space at level N, we have

$$\sum_{N=0}^{\infty} P(N)\, q^N \;=\; \prod_{n=1}^{\infty} \frac{1}{1-q^n}. \tag{10.8.20}$$

To check the validity of this expression is sufficient to expand each factor on the right-hand side in terms of the geometrical series and then gather together the various terms to form the powers of q^N. Expanding in series the function $f(q)$ we have

$$f(q) \;=\; 1 + q + 2q^2 + 3q^3 + 5q^4 + 7q^5 + 11q^6 + 15q^7 + 22q^8 + 30q^9 + 42q^{10} + \cdots$$

from which we can read the values of $P(N)$. As anticipated, they grow extremely fast and their asymptotic estimate is given by the Hardy–Ramanujan formula

$$P(N) \;\simeq\; \frac{\exp\left[\pi\sqrt{\frac{2N}{3}}\right]}{4\sqrt{3}N}. \tag{10.8.21}$$

Let's now investigate in more detail the descendent states. At the level $N = 1$ there is only the state $L_1 \mid \Delta\rangle$ and its norm is easily obtained using eqn (10.8.11), the commutation relations (10.8.16), and the properties (10.8.15) of the state $\mid \Delta\rangle$:

$$\begin{aligned}
\langle \Delta \mid L_{-1}^{\dagger}\, L_{-1} \mid \Delta\rangle \;&=\; \langle \Delta \mid L_1\, L_{-1} \mid \Delta\rangle \\
&=\; \langle \Delta \mid [L_1, L_{-1}] \mid \Delta\rangle \\
&=\; \langle \Delta \mid 2L_0 \mid \Delta\rangle \\
&=\; 2\Delta\langle\Delta|\Delta\rangle \;=\; 2\Delta.
\end{aligned}$$

We can also easily compute the norm of the descendant state $L_{-m} \mid \Delta\rangle$:

$$\begin{aligned}
\langle \Delta \mid L_m\, L_{-m} \mid \Delta\rangle \;&=\; \langle \Delta \mid [L_m, L_{-m}] \mid \Delta\rangle \\
&=\; 2m\,\langle \Delta \mid L_0 \mid \Delta\rangle + \frac{c}{12} m(m^2-1)\,\langle\Delta|\Delta\rangle \\
&=\; 2m\Delta + \frac{c}{12} m(m^2-1).
\end{aligned}$$

The computation soon becomes more involved for the other matrix elements of the $P(N) \times P(N)$ matrix, called the Gram matrix, given by the scalar product of the various descendants of the level N

$$M^{(N)} \;=\; \begin{pmatrix}
\langle\Delta|L_1^N L_{-1}^N|\Delta\rangle & . & . & \langle\Delta|L_1^N L_{-N}|\Delta\rangle \\
. & . & . & . \\
. & . & . & . \\
. & . & . & . \\
\langle\Delta|L_N L_{-1}^N|\Delta\rangle & . & . & \langle\Delta|L_N L_{-N}|\Delta\rangle
\end{pmatrix}. \tag{10.8.22}$$

As an explicit example, we present here the computation of the Gram matrix of level $N = 2$, given by

$$M^{(2)} \;=\; \begin{pmatrix} 4\Delta(2\Delta+1) & 6\Delta \\ 6\Delta & 4\Delta + c/2 \end{pmatrix}. \tag{10.8.23}$$

When the determinant of all the Gram matrices $M^{(N)}$ is different from zero, all the descendant states are linearly independent and their set provides, by construction, an irreducible representation of the Virasoro algebra. The space of states V_Δ so

constructed is called the *conformal family* (or the *Verma module*) of the primary field $\phi_\Delta(z)$ where the seed state $\mid \Delta\rangle$ behaves as the *highest state vector* of the Virasoro algebra.

10.8.2 Representation Theory: The Space of Conformal Fields

The representation theory of the Virasoro algebra can also be developed on the space of conformal fields, similarly to that in the Hilbert space of the states. This study provides, however, useful information on the structure of the conformal families.

Given a conformal field $A(z)$, by the definition of the operators L_n, we have

$$T(z)A(w) \;=\; \sum_{-\infty}^{\infty} \frac{1}{(z-w)^{n+2}} (L_n A)(w). \tag{10.8.24}$$

If we specialize this expression to the case where $A(z)$ is a primary field of conformal weight Δ, with an OPE given by (10.6.8), we can easily extract the action of L_n on this primary field

$$\begin{aligned} (L_0\phi)(z) \;&=\; \Delta\,\phi(z), \\ (L_{-1}\phi)(z) \;&=\; \partial\phi(z), \\ (L_n\phi)(z) \;&=\; 0 \quad n \geq 1. \end{aligned} \tag{10.8.25}$$

The other L_m with negative index create the descendant fields

$$\phi^{(m)}(z) \;\equiv\; (L_{-m}\,\phi)(z),$$

and we can recover all other fields by recurrence

$$\phi^{(n_1,n_2,\ldots,n_k)}(z) \;\equiv\; (L_{-n_1} L_{-n_2} \ldots L_{-n_k}\phi)(z), \tag{10.8.26}$$

adopting the usual ordering $n_1 \leq n_2 \leq \cdots \leq n_k$. These fields are also eigenvectors of L_0 with eigenvalues given by

$$L_0\,\phi^{(n_1,n_2,\ldots,n_k)}(z) \;=\; (\Delta + n_1 + n_2 + \ldots n_k)\,\phi^{(n_1,n_2,\ldots,n_k)}(z). \tag{10.8.27}$$

Note that a significant example of a descendant field is provided by the stress–energy tensor! In fact, taking the identity operator I as a primary field, we have

$$(L_{-2}I)(w) \;=\; \frac{1}{2\pi i}\oint dz \frac{1}{z-w} T(z)I \;=\; T(w). \tag{10.8.28}$$

This explains the more complicated transformation law of $T(z)$, given by eqn (10.7.11), with respect to that of a primary field: it is because it is a descendant field.

When the descendant fields (10.8.26) are all linearly independent, they form together with the primary field $\phi(z)$ an irreducible representation of the Virasoro algebra. Since $(L_{-1}\phi)(z) = \partial\phi(z)$, we also deduce that in the conformal family of the operator $\phi(z)$ there are automatically all the derivatives of the primary fields and its descendants.

Let's now prove a result that will be extremely important for the development of the formalism:

> All correlation functions of the descendant fields can be obtained by acting with *linear differential operators* $\mathcal{L}_a$ on the correlation functions of the primary fields. The operators $\mathcal{L}_a$ are uniquely fixed by the Virasoro algebra.

We present this result for the simplest case of a correlation function of the primary fields $\phi_i(z_i)$ $(i = 1, \ldots, n-1)$ and the descendant field $(L_{-k}\,\phi_n)(z)$ of the primary field ϕ_n, where we have

$$\langle \phi_1(z_1) \ldots \phi_{n-1}(z_{n-1})(L_{-k}\phi_n)(z)\rangle \;=\; \mathcal{L}_{-k}\,\langle \phi_1(z_1) \ldots \phi_{n-1}(z_{n-1})\phi(z)\rangle. \tag{10.8.29}$$

The linear differential operator $\mathcal{L}_{-k}$ is expressed as

$$\mathcal{L}_{-k} \;=\; -\sum_{i=1}^{n-1}\left[\frac{(1-k)\Delta_i}{(z_i-z)^k} + \frac{1}{(z_i-z)^{k-1}}\frac{\partial}{\partial z_i}\right]. \tag{10.8.30}$$

To prove eqn (10.8.29) is convenient to start from the Ward identity

$$\langle T(z)\phi_1(z_1) \ldots \phi_n(z_n)\rangle \;=\; \sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2} + \frac{1}{z-z_i}\frac{\partial}{\partial z_i}\right]\,\langle \phi_1(z_1) \ldots \phi_n(z_n)\rangle,$$

and consider the limit $z \to z_n$. Using eqn (10.8.24), the left-hand side of the Ward identity becomes

$$\sum_{k\geq 0}^{\infty}(z-z_n)^{k-2}\langle \phi_1(z_1) \ldots (L_{-k}\phi_n)(z_n),$$

and, for the Cauchy formula

$$\begin{aligned}&\langle \phi_1(z_1) \ldots \phi_{n-1}(z_{n-1})(L_{-k}\phi_n)(z)\rangle \\ &\quad= \frac{1}{2\pi i}\oint_{z_n} dz(z-z_n)^{1-k}\left[\sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2} + \frac{1}{z-z_i}\frac{\partial}{\partial z_i}\right]\,\langle \phi_1(z_1) \ldots \phi_n(z_n)\rangle\right].\end{aligned}$$

Since the residue at infinity of this expression vanishes, we can use the residue theorem of complex analysis to express the contour integral around the point z_n in terms of the contour integrals (taken clockwise) around the points z_i $(i = 1, \ldots, n-1)$. However, the last quantities are simply the opposite of the contour integrals taken in the usual anticlockwise direction, and we have then the situation shown in Fig. 10.11.

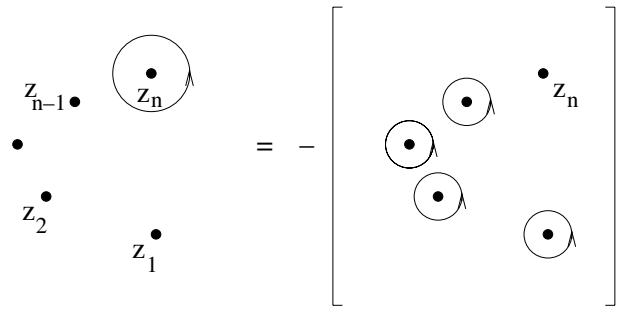


Fig. 10.11 *Theorem of the residues applied to eqn (10.8.31).*

Hence

$$\begin{aligned}
&\langle \phi_1(z_1) \dots \phi_{n-1}(z_{n-1)}(L_{-k}\phi_n)(z)\rangle \\
&\quad = -\frac{1}{2\pi i}\sum_{j=1}^{n-1}\oint_{z_j} dz(z-z_n)^{1-k}\left[\sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2}+\frac{1}{z-z_i}\frac{\partial}{\partial z_i}\right]\langle\phi_1(z_1)\dots\phi_n(z_n)\rangle\right] \\
&\quad = -\sum_{j=1}^{n-1}\left[\frac{(1-k)\Delta_j}{(z_j-z_n)^k}+\frac{1}{(z_j-z_n)^{k-1}}\frac{\partial}{\partial z_j}\right]\langle\phi_1(z_1)\dots\phi_n(z_n).
\end{aligned}$$

We obtain in this way eqn (10.8.29). Similar formulas can be easily derived for all other descendant fields.

There are several important consequences of eqn (10.8.29) and the like.

Orthogonality of conformal families. The first consequence is on the orthogonality condition of the two-point correlation functions of the descendant fields. In fact, the orthogonality condition (10.4.5) of the primary fields automatically implies that also the two-point correlation functions of the descendant fields of two different families vanish. Hence, there is complete orthogonality between two different conformal families.

Structure constants of descendant fields. The second important consequence concerns the structure constants of the descendant fields in the operator algebra (10.2.13). As an outcome of the existence of the linear differential operators $\mathcal{L}_a$, these quantities are proportional to the structure constants c_{ijp} of the primary fields, with a proportionality coefficient that is *uniquely* determined by the Virasoro algebra. In more detail, denoting by $C_{ijp}^{(k,\bar{k})}$ the structure constant of two primary fields ϕ_i, ϕ_j with a descendant $\phi_p^{(k,\bar{k})}$ at the levels k and $\bar{k}$ of the primary field ϕ_p, we have

$$C_{ijp}^{(k)} \;=\; C_{ijp}\,\beta_{ijp}^{(k)}\,\bar{\beta}_{ijp}^{(\bar{k})}, \tag{10.8.31}$$

where $\beta_{ijp}^{(k)}$ is a rational expression of the conformal weights Δ_i (alias) and the central charge c. The same for $\bar{\beta}_{ijp}^{(k)}$, a function of $\bar{\Delta}_i$ and c. Both quantities can be computed

in a purely algebraic way by applying the relative operators $\mathcal{L}_a$ of the descendant field to its three-point correlation functions with the two primary fields.

Notice that eqn (10.8.31) implies that, if $C_{ijp} = 0$, then all other structure constants of the descendant fields vanish as well. In another words, if two primary fields ϕ_i and ϕ_j do not couple to the primary field ϕ_p, they do not couple either to any of its descendants. Hence, in two-dimensional conformal field theories, the determination of the structure constants of the operatorial algebra (10.2.13) simply reduces to determining only the structure constants of the primary fields.

10.9 Hamiltonian on a Cylinder Geometry and the Casimir Effect

Consider a conformal theory defined on a cylinder of width L with periodic boundary conditions. The coordinates along the cylinder are given by $-\infty < \tau < +\infty$ and $0 \le \sigma \le L$. This theory can be analyzed in terms of the conformal transformation

$$w \equiv \tau + i\sigma = \frac{L}{2\pi} \log z, \tag{10.9.1}$$

that maps the plane into the cylinder. Using the transformation law (10.7.11) of the stress–energy tensor, we have

$$T_{cyl}(w) = \left(\frac{2\pi}{L}\right)^2 \left[T_{pl}(z)\, z^2 - \frac{c}{24}\right], \tag{10.9.2}$$

with an analogous expression for $\bar{T}$. We can now define the hamiltonian of this conformal theory in terms of the space integral of $\hat{T}_{\tau\tau}$

$$\begin{aligned} H &= \frac{1}{2\pi} \int_0^L \hat{T}_{\tau\tau}(\sigma)\, d\sigma = \frac{1}{2\pi} \int_0^L (T(\sigma) + \bar{T}(\sigma))\, d\sigma \\ &= \frac{2\pi}{L}(L_0 + \bar{L}_0) - \frac{\pi c}{6L}, \end{aligned} \tag{10.9.3}$$

where we have used eqn (10.9.2) and the definition of the Virasoro generators in the complex plane

$$L_0 = \frac{1}{2\pi i} \oint z T(z)\, dz, \quad \bar{L}_0 = -\frac{1}{2\pi i} \oint \bar{z} \bar{T}(\bar{z})\, d\bar{z}.$$

The theory on the cylinder also has a translation invariance along the σ axis and therefore we can also define the momentum operator P:

$$P = \frac{1}{2\pi} \int_0^L \hat{T}_{\tau\sigma}\, d\sigma = \frac{2\pi}{L}(L_0 - \bar{L}_0). \tag{10.9.4}$$

This operator commutes with H. From the explicit expressions for H and P it can be seen that their eigenvectors are in one-to-one correspondance with the eigenvectors of $L_0 + \bar{L}_0$ and $L_0 - \bar{L}_0$. The minimum value E_0 of the energy is

$$E_0 = -\frac{\pi c_{eff}}{6L}, \tag{10.9.5}$$

where

$$c_{eff} \;=\; c - 24\Delta_{min}, \tag{10.9.6}$$

is the *effective central charge*, given by the central charge c and the minimum eigenvalue Δ_{min} of L_0. For unitary theories $\Delta_{min} = 0$ and therefore $c_{eff} = c$. Futhermore, for unitary theories $c > 0$. For non-unitary theories, Δ_{min} is generically negative as well as the central charge c. However, as we shall see in the next chapter, an interesting observation is that the effective central charge of all minimal conformal models (either unitary or not) is always positive:

$$c_{eff} \;=\; c - 24\Delta_{min} > 0. \tag{10.9.7}$$

The finite expression (10.9.5) of the ground state energy on a cylinder is known as the *Casimir effect*: it depends on its width L and vanishes in the limit $L \to \infty$ when the cylinder reduces to a plane. In addition to its conceptual relevance, this formula is useful to identify which conformal theory is behind the critical behavior of a statistical model defined on a lattice: in fact, it is sufficient to study its ground state energy on a cylinder geometry as a function of L and extract accordingly its effective central charge.

The previous expressions of H and P are also useful to determine the transfer matrix of the conformal models. For simplicity, let's focus attention on a unitary conformal model, with $c > 0$ and $\Delta_i > 0$. In the plane, the two-point correlation function of a primary field is

$$\langle \phi(z,\bar{z})\phi(z',\bar{z}')\rangle \;=\; (z-z')^{-2\Delta}\,(\bar{z}-\bar{z}')^{-2\bar{\Delta}}. \tag{10.9.8}$$

Using the transformation law (10.6.3) of the primary fields under a conformal transformation and the map (10.9.1), we can immediately write down the correlation function on the cylinder

$$\langle \phi(w,\bar{w})\phi(w',\bar{w}')\rangle \;=\; \left(\frac{\pi}{L}\right)^{2(\Delta+\bar{\Delta})} \frac{1}{(\sinh \pi(w-w)/L)^{2\Delta}\,(\sinh \pi(\bar{w}-\bar{w}')/L)^{2\bar{\Delta}}}.$$

Imposing $w = \tau + i\sigma$ and $w' = \tau' + i\sigma'$, for $\tau > \tau'$ this expression can be expanded as

$$\begin{aligned}\left(\frac{\pi}{L}\right)^{2x} \sum_{N,\bar{N}=0}^{\infty} a_N\, a_{\bar{N}} \exp[-2\pi(x+N+\bar{N})(\tau-\tau')/L] \\ \times\, \exp[2\pi i(s+N+\bar{N})(\sigma-\sigma')/L],\end{aligned} \tag{10.9.9}$$

where $x = \Delta + \bar{\Delta}$ is the scaling dimension of the operator, $s = \Delta - \bar{\Delta}$ is its spin, and the coefficients a_N are given by

$$a_N \;=\; \frac{\Gamma(x+N)}{\Gamma(x)N!}.$$

The expression above can be compared with the one computed using the transfer matrix. In the transfer matrix approach, the conformal fields $\phi(u,v)$ are regarded

as operators that act on states of the Hilbert space on the cylinder, with the time evolution provided by eqn (10.8.6). Hence

$$\langle \phi(w,\bar{w})\phi(w',\bar{w}') \equiv \langle 0 \mid e^{H\tau}\phi(0,\sigma)e^{-H\tau}e^{H\tau'}\phi(0,\sigma')e^{-H\tau'} \mid 0\rangle.$$

On the other hand, we can use the momentum operator P to express $\phi(0,\sigma)$ as

$$\phi(0,\sigma) = e^{-iP\sigma}\phi(0,0)e^{iP\sigma}.$$

Inserting now in the expression of the correlation function the completeness relation of the eigenstates $\mid n,k\rangle$ of the energy and the momentum, one has

$$\langle \phi(w,\bar{w})\phi(w',\bar{w}') = \sum_{n,k} \mid \langle 0 \mid \phi(0,0) \mid n,k\rangle \mid^2 \; e^{-(E_n-E_0)(\tau-\tau')+ik(\sigma-\sigma')}. \quad (10.9.10)$$

Comparing with (10.9.9), one derives that the energy and the momentum of these states are given, as expected, by

$$E_n = E_0 + 2\pi(x+N+\bar{N})/L, \quad p_n = 2\pi(s+N+\bar{N})/L \quad (10.9.11)$$

with $E_0 = -\pi c/(6L)$. The matrix element of the operator ϕ on the ground state, here denoted by $\mid \phi\rangle$, is

$$\langle\, 0 \mid \phi(0,0) \mid \phi\rangle = \left(\frac{2\pi}{L}\right)^x, \quad (10.9.12)$$

while the matrix elements on the descendant states are given by

$$\mid \langle\, 0 \mid \phi(0,0) \mid \phi,N,\bar{N}\rangle \mid^2 = \left(\frac{2\pi}{L}\right)^{2x} a_N\, a_{\bar{N}}. \quad (10.9.13)$$

The same considerations can be made for the three-point functions of the primary fields. Trasforming their expression from the plane to the cylinder and expanding it for $\tau_1 \gg \tau_2 \gg \tau_3$, we have

$$\langle \phi_i(\tau_1,\sigma_1)\phi_j(\tau_2,\sigma_2)\phi_k(\tau_3,\sigma_3)\rangle = C_{ijk}\left(\frac{2\pi}{L}\right)^{x_i+x_j+x_k} e^{-2\pi x_i(\tau_1-\tau_2)/L}\, e^{-2\pi x_k(\tau_2-\tau_3)/L} \times e^{2\pi i s_i(\sigma_1-\sigma_2)/L}\, e^{2\pi i s_k(\sigma_2-\sigma_3)/L}. \quad (10.9.14)$$

Comparing this expression with the one obtained by the operatorial formalism, one can conclude that the structure constants C_{ijk} of the primary fields is given by the matrix element of the lowest energy states of the conformal families

$$\langle \phi_i \mid \phi(0,0) \mid \phi_k\rangle = c_{ijk}\left(\frac{2\pi}{L}\right)^{x_j}. \quad (10.9.15)$$

Its derivation is left as an exercise.

Appendix 10A. Moebius Transformations

In this appendix we discuss some important aspects of the Moebius trasformations. They are closely related to the group of isometries of the hyperbolic plane and three-dimensional hyperbolic surfaces. An important subgroup of these transformations is given by the *modular group* that plays an important role in the classification of the partition functions of the conformal theories.

As discussed in the text, the Moebius transformations are given by

$$w(z) \,=\, \frac{az+b}{cz+d}, \tag{10.A.1}$$

with a, b, c, d complex numbers that satisfy $ad - bc \neq 0$. Since multiplying all these numbers by a common factor does not alter the mapping (10.A.1), we can always assume that they satisfy the condition

$$ad - bc \,=\, 1. \tag{10.A.2}$$

Any Moebius trasformation, which is not simply a linear function, can be obtained as the composition of two linear transformations and one inversion. In fact, if $c = 0$, the map is linear. If, on the contrary, $c \neq 0$, it can be written

$$w(z) \,=\, \frac{a}{c} + \frac{bc - ad}{c(cz+d)}. \tag{10.A.3}$$

This expression shows that the original mapping can be decomposed into a sequence of the three transformations

$$z_1 \,=\, cz + d, z_2 \,=\, \frac{1}{z_1}, w \,=\, \frac{a}{c} + \frac{bc - ad}{c} z_2. \tag{10.A.4}$$

Group structure. The Moebius transformations form a group. This means that the class of these functions contains the identity and the inverse transformations and, furthermore, that the product of two Moebius transformations belongs to the same set. It is easy to prove this statement. With the choice $b = c = 0$, $a = d = 1$, we obtain the identity transformation $w(z) = z$. To determine the inverse, we need to solve the equation $w(z) = f(z)$ for the variable z in terms of in w, with the final result (expressed in the variable z) given by

$$\frac{dz - b}{-cz + a}. \tag{10.A.5}$$

This corresponds to the substitutions $a \to d$, $b \to -b$, $c \to -c$ and $d \to a$. As a by-product of this computation, one obtains that the combination $ad - bc$ is an invariant quantity. Consider now the product of two transformations: let $z_2 = f_2(z)$

and $w = f_1(z_2)$ be the two transformations with parameters a_i, b_i, c_i, d_i $(i = 1, 2)$. The final result is given by

$$f_3(z) \;=\; \frac{a_3 z + b_3}{c_3 z + d_3}, \tag{10.A.6}$$

with

$$\begin{array}{l} a_3 = a_1 a_2 + b_1 c_2 \;\;,\;\; b_3 = a_1 b_2 + b_1 d_2 \\ c_3 = c_1 a_2 + d_1 c_2 \;\;,\;\; d_3 = c_1 b_2 + d_1 d_2. \end{array} \tag{10.A.7}$$

These composition laws can be elegantly expressed in terms of a matrix algebra, associating to the transformation (10.A.1) the matrix

$$W \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{10.A.8}$$

The condition (10.A.2) becomes $\det \mathrm{W} = 1$. Hence the inverse matrix exists and is given by

$$W^{-1} \;=\; \begin{pmatrix} d & -b \\ -c & d \end{pmatrix}, \tag{10.A.9}$$

which corresponds to (10.A.5). It is also simple to see that the composition law (10.A.7) corresponds to the usual matrix multiplication law. The decomposition (10.A.4) implies that any Moebius transformation is either linear or it can be decomposed as $W_1 W_2 W_3$, where W_i are expressed by the matrices

$$W_1 \;=\; \begin{pmatrix} a_1 & b_1 \\ 0 & 1 \end{pmatrix}, \quad W_2 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad W_3 \;=\; \begin{pmatrix} a_3 & b_3 \\ 0 & 1 \end{pmatrix}. \tag{10.A.10}$$

Harmonic ratio. It is immediate to show that the harmonic ratio of four distinct points $z_1, \ldots, z_4$ is invariant under a Moebius map, namely

$$\frac{(w_1 - w_4)(w_3 - w_2)}{(w_1 - w_2)(w_3 - w_4)} \;=\; \frac{(z_1 - z_4)(z_3 - z_2)}{(z_1 - z_2)(z_3 - z_4)}, \tag{10.A.11}$$

where the w_i are the images of the points z_i under the mapping (10.A.1). Note that this equation has an important consequence. Namely, imposing $w_4 = w$ e $z_4 = z$, we have

$$\frac{(w_1 - w)(w_3 - w_2)}{(w_1 - w_2)(w_3 - w)} \;=\; \frac{(z_1 - z)(z_3 - z_2)}{(z_1 - z_2)(z_3 - z)}, \tag{10.A.12}$$

which can be written in the form (10.A.1), where the coefficients a, b, c, d are uniquely fixed in terms of the points z_i and w_i. This means that a Moebius trasformation is uniquely determined once we fix the mapping of *three* different points in the complex plane. A close look at eqn (10.A.1) shows that the point $z = -b/a$ is mapped onto the point $w = 0$, the point $z = -d/c$ onto the point at infinity $w = \infty$ and, finally, the point at infinity of the z-plane onto the point $w = a/c$.

Circles onto circles. An important geometrical property of the Moebius transformations is that they map circles onto circles, including in this terminology also straight

lines, regarded as circles of infinite radius.[13] To prove this, it is sufficient to show that each of the three elementary transformations (10.A.4) in which any Moebius transformation can be decomposed, has this property. Let's write initially the general expression of a line and a circle in complex coordinates: for a straight line we have

$$ax + by + c \,=\, 0, a, b, c \in \Re \tag{10.A.13}$$

and using $z = x + iy$, $\bar{z} = x - iy$, it reads

$$Az + \bar{A}\,\bar{z} + c \,=\, 0, A \,=\, \frac{a - ib}{2}. \tag{10.A.14}$$

For a circle of radius r, whose center is in z_0, we have $(z - z_0)(\bar{z} - \bar{z}_0) \,=\, r^2$, namely

$$z\,\bar{z} + B\bar{z} + \bar{B}z + C \,=\, 0, B \,=\, -z_0, C \,=\, |B|^2 - r^2. \tag{10.A.15}$$

Under a translation and a rotation, expressed generally by the transformation $z \rightarrow az + b$, both (10.A.14) and (10.A.15) keep their form. Under the inversion $z = 1/w$, $\bar{z} = 1/\bar{w}$, eqn (10.A.14) becomes

$$cw\bar{w} + A\bar{w} + \bar{A}w \,=\, 0. \tag{10.A.16}$$

If $c \,=\, 0$ (the original line passes through the origin), this equation defines a new straight line that passes through the origin. Vice versa, if $c \neq 0$, the equation above defines a circle of radius $|A|/|c|$, centered at $-A/c$. Acting now with an inversion transformation on (10.A.15), it becomes

$$Cw\,\bar{w} + Bw + \bar{B}\bar{w} + 1 \,=\, 0. \tag{10.A.17}$$

If $C \,=\, 0$ (this corresponds to the original circle that passes through the origin) we have a straigh line. Otherwise, it defines a new circle, with center at $z_0 = -\bar{B}/C$ and radius $r^2 = |B|^2/|C|^2 - 1/C$.

Closely related to the property discussed above, there is the transformation law that involves the internal and external points of the circles. Let $\mathcal{D}_i$ be the set of internal points of the circle C in the z-plane and $\mathcal{D}_e$ the set of its external points, with an analogous definition of $\mathcal{D}'_i$ and $\mathcal{D}'_e$ for the points relative to the circle C' in the w-plane, in which the circle C is mapped. There can be only two cases: (i) the first, in which $\mathcal{D}_i$ is mapped onto $\mathcal{D}'_i$ and correspondingly $\mathcal{D}_e$ onto $\mathcal{D}'_e$; (ii) the second, in which $\mathcal{D}_i$ is mapped onto $\mathcal{D}'_e$ while $\mathcal{D}_e$ onto $\mathcal{D}'_i$. The proof is left as an exercise.

Symmetric points. We also mention, without proof, another characteristic property of the Moebius map: it transforms symmetric points with respect to a circle onto symmetric points of the image circle. Two points p and q are symmetric with respect to a circle of center z_0 and radius r if z_0, p and q are aligned in the given order, with the distances $|z_0 - p|$ and $|z_0 - q|$ that satisfy the condition (see Fig. 10.12)

$$|z_0 - p|\,|z_0 - q| \,=\, r^2. \tag{10.A.18}$$

[13] This is a natural assumption on the Riemann sphere associated to the complex plane.

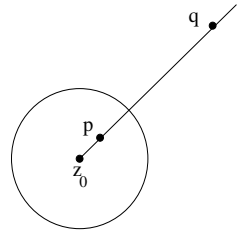


Fig. 10.12 *Symmetric points p and q with respect to a circle of radius r.*

Denoting by w_0 the center of the image circle, R its radius, and p' and q' the image points of p and q, one finds that

$$|w_0 - p'|\,|w_0 - q'| \;=\; R^2. \tag{10.A.19}$$

Fixed points. It is interesting to observe that the Moebius transformations can also be characterized by the properties of their fixed points. These are the points left invariant by the map (10.A.1)

$$z \;=\; w(z). \tag{10.A.20}$$

They can be of four different types: parabolic, elliptic, hyperbolic, and lossodromic. This classification has both a geometrical and algebraic meaning, as shown by the figures given below. The different classes can be distinguished by the trace $\mathrm{Tr}M = a+d$ of the matrix M. In more detail, the Moebius transformations are

- *parabolic*, if $a + d \;=\; \pm 2$;
- *elliptic*, if $a + d$ is a real number, with $\mid a + d \mid \leq 2$;
- *hyperbolic*, if $a + d$ is a real number, with $\mid a + d \mid \geq 2$;
- *lossodromic*, if $a + d$ is a complex number.

Since the trace of a matrix is invariant under a conjugation transformation

$$M \to U^{-1} M U, \tag{10.A.21}$$

where U is also a Moebius transformation, all members of a conjugate class are of the same type.

Solving the second-order algebraic equation (10.A.20) and denoting the two roots as $\gamma_{1,2}$, we have

$$\gamma_{1,2} \;=\; \frac{(a-d) \pm \sqrt{(a-d)^2 + 4bc}}{2} \;=\; \frac{(a-d) \pm \sqrt{(a+d)^2 - 4}}{2}, \tag{10.A.22}$$

where we have used the relation $ad - bc = 1$. Except for the trivial cases $c = 0$, and $a = d$, or $b = c = 0$, in which there is an infinite number of fixed points (since the transformation is the identity), there are in general two distinct fixed points. However they coalesce when

$$(\mathrm{Tr}\, M)^2 = (a+d)^2 \;=\; 4. \tag{10.A.23}$$

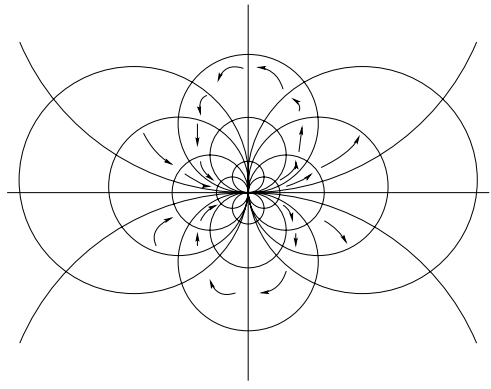


Fig. 10.13 *Transformation of the complex plane under a Moebius map of parabolic type.*

Let us consider the two cases separately. When the two fixed points coincide, we are in the presence of *parabolic transformations*. All these transformations are conjugated to the matrix

$$M_p = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{10.A.24}$$

If γ denotes the only fixed point, their general form is

$$\frac{1}{w-\gamma} = \frac{1}{z-\gamma} + \beta, \tag{10.A.25}$$

where β is a free parameter related to the translations. In the parabolic case we have that: (i) any circle that passes through the fixed point is transformed onto a tangent circle that passes through the fixed point; (ii) any family of tangent circles is then transformed into itself; (iii) the internal region of each circle is transformed onto itself. Under this class of transformations the way in which the plane changes is shown in Fig. 10.13.

When there are two distinct fixed points, eqn (10.A.12) implies

$$\frac{w-\gamma_1}{w-\gamma_2} = \kappa \frac{z-\gamma_1}{z-\gamma_2}, \tag{10.A.26}$$

where κ is a constant that depends on γ_1, γ_2, z_2, and w_2. Hence, the general expression of a Moebius trasformation with two distinct fixed points depends on an additional constant κ. Using the conjugation transformation, the two points $\gamma_{1,2}$ can be mapped one at 0 and the other to ∞, Consequently, all these transformations are conjugated to the matrix

$$M = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \tag{10.A.27}$$

with $\lambda^2 = \kappa$. This matrix corresponds to the mapping $w = \kappa z$. For this reason, the constant κ is called the *multiplier* of the transformation. We have an *elliptic*

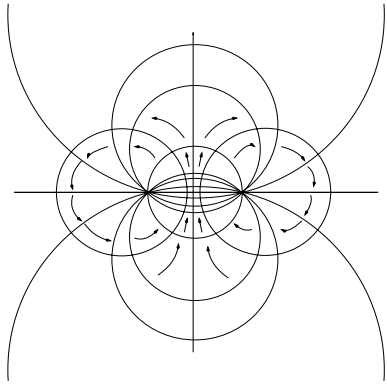


Fig. 10.14 *Transformation of the complex plane under a Moebius map of elliptic type.*

transformation when

$$0 \leq (a+d) \leq 4. \tag{10.A.28}$$

This condition is equivalent to $|\kappa| = 1$, namely $\kappa = e^{i\alpha}$, with α a real parameter.[14] In this case we have the following properties: (i) an arc of a circle passing through the fixed points is transformed to another arc of a circle passing through them but rotated by an angle α with respect to the original one; (ii) each circle orthogonal to the circles passing through the fixed points is transformed onto itself and the same holds for its internal region. The nature of this transformation is shown in Fig. 10.14.

We have a *hyperbolic transformation* when

$$(a+d)^2 \geq 4, \tag{10.A.29}$$

namely when κ is a real number. Note that $w'(\gamma_1) = \kappa$ whereas $w'(\gamma_2) = \kappa^{-1}$ so that, if $\kappa > 1$, γ_1 is a repulsive point, whereas γ_2 is an attractive point. Their role is swapped if $\kappa < 1$. For the hyperbolic transformations we have: (i) each circle that passes through the fixed points is transformed onto itself, namely each of the two arcs of which the circle is composed is mapped on itself; (ii) the internal region of a circle passing through the fixed points is mapped onto itself; (iii) each circle that is orthogonal to a circle passing through the fixed points is transformed to an analogous circle. The way the hyperbolic transformations act is shown in Fig. 10.15.

Finally, we have a *lossodromic transformation* in the remaining cases, namely when $(\mathrm{Tr}M)^2$ does not belong to the interval $[0, 4]$. In this case the multiplier is given by $\kappa = Ae^{i\alpha}$. Hence its action is a combination of the motions shown in Figs 10.14 and 10.15. Each arc passing through the fixed points is transformed to a similar arc but rotated by α, while a circle orthogonal to the circles passing through the fixed points is transformed onto another orthogonal circle. The lossodromic transformations do not have, in general, fixed circles expect in the case in which $\alpha = \pi$.

[14]Since the multiplier of M^n is κ^n, the only Moebius transformations of finite order are elliptic and they correspond to rational values of α.

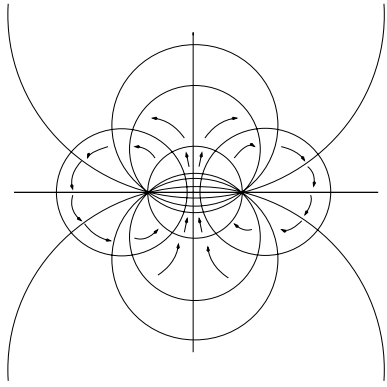


Fig. 10.15 *Transformation of the complex plane under a Moebius map of hyperbolic type.*

Let's now discuss two particular examples of Moebius transformations that may be useful later on. The first is the transformation that maps the upper half-plane $\mathrm{Im}\, z > 0$ in the internal region of the circle $|w| < 1$. Its general expression is

$$w(z) = \lambda \frac{z-\alpha}{z-\bar{\alpha}}, \quad |\lambda| = 1, \quad \mathrm{Im}\, \alpha > 0. \tag{10.A.30}$$

To prove that the upper half-plane is mapped onto the internal points of the circle, consider the points along the real axis. For those points we have $|z-\alpha| = |z-\bar{\alpha}|$, and therefore they are mapped to the points of the circle $|w| = 1$. On the other hand, the point $z = \alpha$ is transformed onto the origin $w = 0$. For the properties of the Moebius map discussed above, this is sufficient to conclude that any other point of the domain $\mathrm{Im}\, z > 0$ is mapped inside the circle. Note that the point $z = \bar{\alpha}$ is mapped onto $w = \infty$ and this is enough to conclude that the lower half-plane is transformed onto the external region of the circle $|w| = 1$.

The second map we consider is the one that maps the disk $|z| < 1$ onto the disk $|w| < 1$. Its general expression is

$$w(z) = \lambda \frac{z-\alpha}{\bar{\alpha} z - 1}, \quad |\lambda| = 1, \quad |\alpha| < 0. \tag{10.A.31}$$

Note, in fact, that the points of the circle in the z-plane are expressed by $z = e^{i\phi}$ and for those points we have

$$|w| = |\lambda| \left| \frac{e^{i\phi} - \alpha}{\bar{\alpha} e^{i\phi} - 1} \right| = \frac{|\alpha - e^{i\phi}|}{|\bar{\alpha} - e^{-i\phi}|} = 1. \tag{10.A.32}$$

Since $z = 0$ is mapped onto the point $\lambda\alpha$, with $|\lambda\alpha| < 1$, this is sufficient to conclude that all internal points of the circle in the z-plane are mapped onto the internal point of the circle in the w-plane.

References and Further Reading

For the mathematical part of this chapter, a superb text on complex analysis is:

M. Ablowitz, A. Fokas, *Complex Variables. Introduction and Applications*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1977.

There are several review articles on conformal field theories. Here we mention:

P. Ginsparg, *Applied Conformal Field Theory*, Les Houches, Session XLIX, 1988, *Field, Strings and Critical Phenomena*.

J. L. Cardy, *Conformal Invariance and Statistical Mechanics*, Les Houches, Session XLIX, 1988, *Field, Strings and Critical Phenomena*.

A.B. Zamolodchikov, Al.B. Zamolodchikov, *Conformal field theory and critical phenomena in two-dimensional systems*, Sov. Sci. Rev. A. Physics, 10 (1989), 269.

For an exhaustive and pedagogical analysis of conformal field theory, and in particular for their mathematical aspects, we refer to:

P. Di Francesco, P. Mathieu, D. Senechal, *Conformal Field Theory*, Springer-Verlag, New York, 1977.

Moreover, it is mandatory to mention the original article in which the two-dimensional conformal field theories were firstly introduced:

A. Belavin, A. Polyakov, A.B. Zamolodchikov, *Infinite conformal symmetry in two-dimensional quantum field theory*, Nucl. Phys. B 241 (1984), 333.

Problems

1. Operatorial identities

There is a simple example that shows the necessity of considering the operatorial identities only in a weak sense, i.e. true only for the matrix elements. Consider an interacting scalar field $\varphi(x)$ and suppose that for $x_0 \to -\infty$ its interactions vanish. In this case it seems natural to impose the operatorial identity

$$\lim_{x_0 \to -\infty} \varphi(x) = \varphi_{in}(x)$$

where $\varphi_{in}(x)$ is a free bosonic field. However, this leads to a contradiction. In fact, if the relation above were true, we would have

$$\lim_{x_0 \to -\infty} \lim_{y_0 \to -\infty} \langle 0 \mid \phi(x)\phi(y) \mid 0 \rangle \; = \; \langle 0 \mid \varphi_{in}(x)\varphi_{in}(y) \mid 0 \rangle .$$

Since $\varphi_{in}(x)$ is a free field, the right hand side is the usual progagator $G_{free}(x-y)$ of a scalar free field

$$G_{free}(x-y) \,=\, \int \frac{d^dp}{(2\pi)^d}\frac{1}{p^2-m^2}e^{ip\dot{(}x-y)}.$$

a Use the Lorentz invariance to fix the dependence on the coordinates of the propagator $G(x-y)$ of the interacting field $\varphi(x)$.

b Argue that the propagator does not coincide in the limit $x_0 \,\to\, -\infty$ with $G_{free}(x-y)$.

2. Correlation functions

Assuming the validity of the operator product expansion, show that all the correlation functions of a massless field theory can be expressed in terms of the propagators and the structure constants.

3. Laplace equation and conjugate harmonic functions

1. Show that the real and imaginary parts of an analytic function of a complex variable z

$$f(z) \,=\, \Omega(x,y) + i\,\Psi(x,y)$$

are both harmonic functions, i.e. they satisfy the Laplace equation

$$\nabla^2\Omega \,=\, \nabla^2\Psi \,=\, 0.$$

2. Vice versa, use the Cauchy–Riemann equations to show that if $\Omega(x,y)$ is a function that satisfies the Laplace equation, then there exists another harmonic function $\Psi(x,y)$ (called the conjugate function of Ω) such that $f(z)=\Omega+i\Psi$ is an analytic function of complex variable.

4. Hydrodynamics of an ideal fluid in two dimensions

Consider the stationary motion of an incompressible and irrotational fluid in two dimensions. Denoting by $\vec{v}(x,y)=(v_1,v_2)$ the vector field of its velocity at the point (x,y) of the plane, it satisfies

$$\vec{\nabla}\dot{\vec{v}} \,=\, 0, \quad \vec{\nabla} \,\wedge\, \vec{v} \,=\, 0.$$

1. Show that these conditions imply the existence of a potential Ω that satisfies the Laplace equation. Moreover, show that, introducing the conjugate function Ψ and defining $f(z)=\Omega+i\Psi$ (the so-called *complex potential*), one has

$$\frac{df}{dz} \,=\, \frac{\partial\Omega}{\partial x} + i\frac{\partial\Psi}{\partial x} \,=\, \frac{\partial\Omega}{\partial x} - i\frac{\partial\Omega}{\partial y} \,=\, v_1 - iv_2 \,=\, \bar{v}.$$

The complex vector field of the velocity is then given by

$$v \,=\, \overline{\left(\frac{df}{dz}\right)}.$$

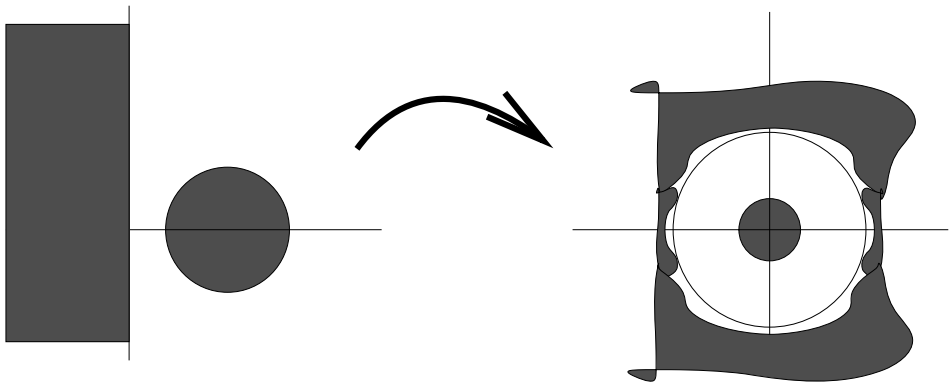


Fig. 10.16 *Conformal map of two domains.*

2. Study the flux lines of the velocity associated to the analytic function

$$f(z) = \frac{i\gamma}{2\pi} \ln z$$

and show that the vector field of the velocity corresponds to a vortex, localized at the origin. Give an interpretation of the parameter γ.

3. Study the flux lines of the velocity relative to the potential

$$f(z) = v_0 \left(z + \frac{a^2}{z} \right) + \frac{i\gamma}{2\pi} \ln z.$$

Determine the points where the velocity vanishes and study their location by varying the parameter γ.

5. Moebius transformations

Show that the transformation $w = (z-a)/(z+a)$, $a = \sqrt{c^2 - \rho^2}$ with c and ρ real and $0 < \rho < c$, maps the domain delimited by the circle $|z - c| = \rho$ and the imaginary axis, onto the annulus domain delimited by $|w| = 1$ and a concentric circle of radius δ, as shown in Fig. 10.16. Find, in particular, the value of δ.

6. Operatorial expansion in the channel of the identity operator

Let $\phi_i(z)$ a primary field of a conformal field theory with central charge c. Let Δ_i be its conformal weight. Prove that the Ward identity uniquely fixes the first terms of the operator expansion in the channel of the identity operator, namely

$$\phi_i(z)\phi_i(w) \frac{1}{(z-w)^{2\Delta_i}} \left[\mathbf{I} + \frac{2\Delta_i}{c} T(w) + \cdots \right].$$

7. Casimir effect

Consider two parallel horizontal planes, separated by a distance **a** along the axis z. Suppose that a massless field theory is defined between the two planes, with boundary

conditions that ensure a non-zero value of the expectation value of the stress–energy tensor $T_{\mu\nu}$, $t_{\mu\nu}(t,\vec{x}) \equiv \langle 0 \mid T_{\mu\nu}(t,\vec{x}) \mid 0\rangle$. The system is assumed to be time invariant. Thanks to the symmetry of the problem, $t_{\mu\nu}$ can be written in terms of the metric tensor $g_{\mu\nu}$ and the tensors made up of the unit vector $\hat{z}_\mu = (0,0,0,1)$.

1. Write the most general expression of $t_{\mu\nu}$ based on the considerations given above.
2. Show that the conservation law $\partial^\mu T_{\mu\nu}(t,\vec{x}) = 0$ and the zero trace condition of $T_{\mu\nu}$ uniquely determine $t_{\mu\nu}$ up to a constant. Use dimensional analysis to fix this constant (up to a numerical coefficient) in terms of the only dimensional parameter of the problem.
3. Use the final form of $t_{\mu\nu}$ to compute the force per unit area between the two planes.

11
Minimal Conformal Models

Small is beautiful.

Anonymous

11.1 Introduction

In this chapter we discuss a particular class of conformal theories, the so-called *minimal models*. The peculiarity of these models consists in the *finite* number of their conformal families that close an OPE. The anomalous dimensions of the conformal fields and the central charge of these theories can be computed exactly and, in particular, assume rational values. Furthermore, of the minimal models we can explicitly compute both the correlation functions of the order parameters and the partition function on a torus, i.e on a cylinder with periodic boundary conditions on both directions. Their mathematical elegance is accompanied by an important physical interpretation: as discussed in more detail in Chapter 14, the conformal minimal models describe the scaling limit of an infinite number of statistical models with a discrete symmetry, among which we find the Ising model, the tricritical Ising model, the Potts model, the Yang–Lee edge singularity, etc. In addition, the unitary minimal models can be put in correspondence with the critical Landau–Ginzburg theories with power interaction $\phi^{2(p-1)}$ ($p = 3, 4, \ldots$): as a matter of fact, they provide the exact solution of these theories at their multicritical point. For all these reasons, the minimal conformal models play a crucial role in the modern understanding of critical phenomena.

This chapter focuses on the general discussion of the minimal models of the Virasoro algebra. We will initially highlight the presence of null vectors in the representations of the Virasoro algebra corresponding to discrete values of the conformal dimensions and the central charge, encoded in the Kac determinant of the so-called *degenerate fields*. Later we will discuss the fusion rules that derive from the particular structure of the Verma modulus of the degenerate fields and the Coulomb gas formalism that allows us to compute the exact expressions of the correlation functions. Finally, we will study the modular invariance of these models and the partition functions compatible with this symmetry. Further aspects of these models will be addressed in the following chapters.

11.2 Null Vectors and Kac Determinant

The starting point in the study of minimal models is the presence of particular null-vectors inside the conformal families. This circumstance is of utmost importance not

only for the study of conformal theories at the critical point but also for their off-critical deformations. For this reason, it deserves to be investigated in detail.

From Section 10.8, we know that a conformal family[1] is identified by the vector $\mid \phi_\Delta \rangle$ associated to the primary field ϕ_Δ. This vector satisfies the conditions

$$\begin{aligned} &L_0 \mid \Delta\rangle \,=\, \Delta \mid \Delta\rangle \\ &L_n \mid \Delta\rangle = 0 \quad n = 1, 2, \ldots \\ &\langle \Delta \mid \Delta\rangle \,=\, 1. \end{aligned} \tag{11.2.1}$$

A conformal family is built on such a vector and on all its descendants obtained by applying to it an ordered string of creation operators L_{-n}. The vector $\mid \Delta\rangle$, as already noticed in the previous chapter, plays the role of *highest weight vector* of the Virasoro algebra. For arbitrary values of Δ and c, all the descendant vectors are linearly independent and the set of all these vectors form then an irreducible representation of the Virasoro algebra. However, for particular values of Δ and c, there are some null-vectors: in such a case, to have an irreducible representation we have to factorize with respect to these states. Before we describe the general case, it is convenient to familiarize ourselves with some simple examples of null-vectors at the lowest levels of the conformal families.

Let's start from the level $N = 1$. Given the primary state $\mid \Delta\rangle$, at this level there is only one descendant state given by $\mid X_1\rangle = L_{-1} \mid \Delta\rangle$. If we request that this is a null-vector, its norm must vanish

$$\begin{aligned} \langle X_1 \mid X_1\rangle &= \langle \Delta \mid L_1 L_{-1} \mid \Delta\rangle \\ &= 2\langle \Delta \mid L_0 \mid \Delta\rangle \,=\, 2\Delta\langle \Delta \mid \Delta\rangle \,=\, 0. \end{aligned}$$

This equation has the only solution $\Delta = 0$. In other words, the only conformal family that has a null-vector at level $N = 1$ is the family of the identity operator $\mathbf{I}$.

A more interesting situation occurs at the level $N = 2$. In this case there are two possible descendant states, the first given by $L_{-1}^2 \mid \Delta\rangle$ and the second by $L_{-2} \mid \Delta\rangle$. Let's determine the conditions for which a linear combination of these states

$$\mid X_2\rangle \,=\, (L_{-2} + \alpha L_{-1}^2) \mid \Delta \,\rangle \tag{11.2.2}$$

gives rise to a null-vector. If $\mid X_2\rangle = 0$, the same is true for the vectors obtained by applying to it either L_1 or L_2. In the first case, using the commutation relations of the Virasoro modes and the properties of the primary state $\mid \Delta\rangle$, we have

$$\begin{aligned} L_1 \mid X_2\rangle &= (L_{-2}L_1 + 3L_{-1} + 2a(L_{-1}L_0 + L_0 L_{-1})) \mid \Delta\rangle \\ &= (3 + 2a(2\Delta + 1))L_{-1} \mid \Delta\rangle \,=\, 0. \end{aligned}$$

This condition then fixes the coefficient a of the linear combination (11.2.2)

$$a \,=\, -\frac{3}{2}\,\frac{1}{2\Delta + 1}. \tag{11.2.3}$$

[1] In this section we focus our attention only on the analytic sector of the theory. As usual, analogous considerations can be done for the anti-analytic sector.

Now applying L_2 to $\mid X_2\rangle$ and again making use of the commutation relations of the Virasoro modes and the conditions of $\mid \Delta\rangle$, we get

$$[L_2, L_{-2}] \mid \Delta\rangle + a[L_2, L_{-1}^2] \mid \Delta\rangle = \left(4L_0 + \frac{c}{2}\right) \mid \Delta\rangle + 3aL_1L_{-1} \mid \Delta\rangle$$
$$= \left(4L_0 + \frac{c}{2} + 6aL_0\right) \mid \Delta\rangle = \left(4\Delta + \frac{c}{2} + 6a\Delta\right) \mid \Delta\rangle = 0$$

namely

$$c = -4\Delta(2 + 3a) \;=\; \frac{2\Delta(5 - 8\Delta)}{2\Delta + 1}. \tag{11.2.4}$$

Summarizing the result of this computation, if the central charge c of the conformal model and the conformal weight Δ of the field under scrutiny are related by the condition (11.2.4), then there exists a linear combination of the descendants at the level $N = 2$ of this primary field ϕ_Δ that leads to a null-vector.

It is worth pointing out that there is a general way to determine whether or not a null-vector at the level N of a conformal family exists. It consists of computing the zeros of the determinant of the Gram matrix at level N (see Section 10.8.1). For $N = 2$, the Gram matrix is given by

$$M^{(2)} \;=\; \begin{pmatrix} 4\Delta(2\Delta + 1) & 6\Delta \\ 6\Delta & 4\Delta + c/2 \end{pmatrix}$$

and its determinant can be written as

$$||M^{(2)}|| \;=\; 2(16\Delta^3 - 10\Delta^2 + 2c\Delta^2 + c\Delta) \;=\; 32(\Delta - \Delta_{1,1})(\Delta - \Delta_{1,2})(\Delta - \Delta_{2,1})$$

where

$$\Delta_{1,1} \;=\; 0, \qquad \Delta_{(1,2),(2,1)} \;=\; \frac{1}{16}(5 - c) \pm \sqrt{(1 - c)(25 - c)}. \tag{11.2.5}$$

Note the appearance of the solution $\Delta_{1,1} = 0$, whose presence was expected. It corresponds to the possibility of having a null-vector at level $N = 1$ that, obviously, will also give rise to a null-vector at level $N = 2$, if we act on it by the arising operator L_{-1}. The other two zeros $\Delta_{1,2}$ and $\Delta_{2,1}$ correspond to the condition (11.2.4) previously derived.

Kac determinant. Remarkably, the zeros of the Gram matrix of level N can be computed exactly. This important mathematical result, due to M. Kac, is a crucial step in the development of two-dimensional conformal theories. The corresponding formula, the so-called *Kac determinant*, is given by

$$\det M^{(N)} \;=\; A_N \prod_{r,s \geq 1; rs \leq N} [\Delta - \Delta_{r,s}]^{P(N - rs)}, \tag{11.2.6}$$

where $P(N - rs)$ is the number of partitions of the integer number $(N - rs)$ and A_N is a positive constant that is not important for the discussion that follows. The zeros $\Delta_{r,s}$ can be parameterized in different ways. One of them, particularly useful for

the Coulomb gas formulation that we will discuss later, is expressed in terms of two parameters, called *charges* $\alpha_\pm$

$$\begin{aligned}\Delta_{r,s}(c) &= \Delta_0 + \frac{1}{4}(r\alpha_+ + s\alpha_-)^2,\\ \Delta_0 &= \frac{1}{24}(c-1),\\ \alpha_\pm &= \frac{\sqrt{1-c} \pm \sqrt{25-c}}{\sqrt{24}}.\end{aligned} \tag{11.2.7}$$

Equivalently, imposing

$$\alpha_+ = \sqrt{t}, \quad \alpha_- = -\frac{1}{\sqrt{t}}$$

the previous conformal quantities can be expressed as

$$\begin{aligned}c &= 13 - 6\left(t + \frac{1}{t}\right)\\ \Delta_{r,s} &= \frac{1}{4}(r^2-1)t + \frac{1}{4t}(s^2-1) - \frac{1}{2}(rs-1).\end{aligned} \tag{11.2.8}$$

Note that, fixing the value of the central charge, there are two possible values of t:

$$t = 1 + \frac{1}{12}\left[1 - c \pm \sqrt{(1-c)(25-c)}\right],$$

but we can choose any of the two, since this does not change the Kac determinant. The parameter t is real only in the cases $c < 1$ or $c > 25$, while it is generally complex in the interval $1 < c < 25$.

A third way to write the conformal data consists of the parameterization of the central charge and the zeros of the Kac determinant given by

$$\begin{aligned}c &= 1 - \frac{6}{q(q+1)}\\ \Delta_{r,s} &= \frac{[(q+1)r - qs]^2 - 1}{4q(q+1)}\end{aligned} \tag{11.2.9}$$

where the real parameter q is related to the central charge c by

$$q = -\frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{25-c}{1-c}}. \tag{11.2.10}$$

Note that the Kac formula does not predict the eigenvalues of the matrix $M^{(N)}$ but only their product. In fact, at each level N, the number of roots $\Delta_{r,s}$ is larger than the number $P(N)$ of its eigenvalues. Another important observation is that the first null-vector of the conformal family $V(c, \Delta_{r,s})$ occurs at the level $N = rs$, since the combinatoric function $P(N - rs)$ vanishes, by definition, for $N < rs$. The multiplicity of the zeros, given by $P(N - rs)$, has the same origin as the one previously pointed out in the explicit computation of the null-vectors at level $N = 2$: namely, among the zeros

of the polynomial at level N, there are also those corresponding to the null-vectors of lower levels. At level N, there are in fact the null-vectors generated by the $P(N-rs)$ combinations of $L_{-n_1} \dots L_{-n_k}$, with $\sum n_i = N - rs$, applied to the null-vectors of level rs.

11.3 Unitary Representations

With the explicit formula of the Kac determinant, one can identify the values of c and Δ that give rise to the unitary irreducible representantions of the Virasoro algebra, in which there are no states with negative norm. Before proceeding with the mathematical analysis of this problem, it should be said that, strictly speaking, the unitary condition is not necessary in statistical mechanics: many non-unitary models find their applications in the discussion of interesting statistical mechanics, also providing a useful generalization of ordinary quantum field theories. In the sections to come, we will see that there are certain statistical models that require the presence of negative anomalous dimensions.

Coming back to the problem of determining the unitary representations, from the mathematical point of view we have to initially determine those regions of c and Δ in which the Kac determinant is negative: in these regions there are definitely states whose norm is negative and the corresponding representations are not unitary. Vice versa, in the regions where the determinant is positive, a further analysis is needed to exclude the presence of such negative norm states, since an even number of them ends up in a positive value of the determinant.

It is easy to see that for $c < 0$ the corresponding conformal theories are non-unitary: in fact it is sufficient to consider the stress–energy tensor of these theories, associated to the descendant $L_{-2} \mid 0\rangle$ of the identity family, to see that the norm of this state is given by

$$\langle 0 \mid L_2 L_{-2} \mid 0 \rangle = \frac{c}{2} \tag{11.3.1}$$

and, for $c < 0$, this is a negative quantity.

For $c > 1$, all representations with $\Delta > 0$ are unitary. It is necessary to distinguish two cases: (i) $1 < c < 25$ and (ii) $c > 25$. In the first case, expressing $\Delta_{r,s}(c)$ as

$$\Delta_{r,s} = \frac{1-c}{96} \left[\left((r+s) + (r-s) \sqrt{\frac{25-c}{1-c}} \right)^2 - 4 \right],$$

one can see that $\Delta_{r,s}$ either has an imaginary part or, for $r = s$, is a negative quantity. For $c > 25$, they are instead all negative. The non-zero value of the Kac determinant in the region $\{c > 1; \Delta > 0\}$ implies that all eigenvalues of $M^{(N)}$ are positive. In fact, for large values of Δ, the Gram matrix is dominated by its diagonal elements, i.e. those with the higher powers of Δ. Since these elements are all positive in this region, this shows that the eigenvalues of $M^{(N)}$ are all positive for large values of Δ. Moreover, the determinant never vanishes in the region $c > 1$ and $\Delta > 0$, implying that all its eigenvalues remain positive in the entire region.

For $c = 1$, the Kac determinant vanishes at $\Delta_n = \frac{n^2}{4}$, with n an integer number, but otherwise it is never negative; even in this case, there is no problem in having unitary representations for $\Delta > 0$.

Hence, the only subtle case is posed by the analysis of the region $0 < c < 1$ and $\Delta > 0$. This problem has been studied by D. Friedan, Z. Qiu, and S. Shenker,[2] and their results can be summarized as follows: all point of the region $R : \{(c, \Delta) \mid 0 < c < 1; \Delta > 0\}$ correspond to non-unitary representations, except the *discrete* series associated to these values of the central charge and the conformal weights

$$c = c(q) = 1 - \frac{6}{m(m+1)}, \quad q = 2, 3, 4, \ldots \tag{11.3.2}$$
$$\Delta = \Delta_{r,s}(q) = \frac{[(q+1)r - qs]^2 - 1}{4q(q+1)}, \quad (1 \le r \le q,\ 1 \le s \le q+1)$$

where m is an integer number. These discrete values of the central charges and conformal weights define the so-called *conformal minimal unitary models*, in the following denoted by $\mathcal{M}_m$.

11.4 Minimal Models

In the interval $0 < c < 1$, the unitary condition selects the discrete set of values (11.3.2). In this section we shall see that is possible to introduce a more general class of minimal models, from now on denoted by $\mathcal{M}_{p,q}$, whose central charge and conformal weights are expressed by the rational values

$$c = 1 - 6\frac{(p-q)^2}{pq}, \quad (p, q) = 1 \tag{11.4.1}$$
$$\Delta_{r,s} = \frac{[(pr - qs]^2 - (p-q)^2]}{4pq}, \quad (1 \le r \le q-1,\ 1 \le s \le p-1)$$

where p and q are two coprime integers, i.e. without common divisors. Note that in all these models we have $\Delta_{1,1} = 0$ and this conformal weight corresponds to the identity operator $\mathbf{I}$. The unitary minimal models are recovered by the choice $p = q+1$ in eqn (11.4.1). In all other cases, the minimal conformal theories are non-unitary, characterized by a negative value of the central charge and some of its conformal weights. The lowest negative conformal weight is given by

$$\Delta_{min} = \Delta_{1,n} = \Delta_{q-1,p-n} = \frac{1 - (p-q)^2}{4pq}. \tag{11.4.2}$$

Note that, even though the central charge of these minimal models is negative, their *effective central charge*

$$c_{eff} = c - 24\Delta_{min} = 1 - \frac{6}{pq} \tag{11.4.3}$$

is instead always a positive quantity.

[2] D. Friedan, Z. Qiu, S. Shenker, *Conformal invariance, unitarity and two-dimensional critical exponents*, Phys. Rev. Lett. 52 (1984), 1575.

As anticipated in the introduction to this chapter, the conformal minimal models satisfy a series of important properties and they are nowadays the most studied and understood conformal theories. In particular, they play an essential role both in the qualitative and quantitative analysis of the phase transitions that take place in two-dimensional systems. To orientate the reader in the discussion to come, it is convenient to briefly summarize their main features:

1. in the minimal models, the number of conformal families is finite and the conformal weights are expressed by the rational numbers given in eqn (11.4.1);
2. the operator product expansion of any pair of conformal fields of these theories involves only a finite number of the operators of the same minimal model;
3. the correlation function of all the conformal fields satisfies a set of linear differential equations that can be exactly solved;
4. the structure constants of the conformal algebra can be exactly computed;
5. their partition functions on a torus geometry can be exactly determined.

Finally, these minimal conformal models provide the exact solution, at criticality, of a significant series of statistical models, such as the Ising model, the tricritical Ising model, the Potts model, etc., and among the non-unitary models, the Yang–Lee edge singularity, self-avoiding random walks, percolation, etc. Thanks to them, there has been a great advance in the comprehension of the classes of universality. Let's now go on with the detailed discussion of the aspects summerized above.

11.4.1 Kac Table

The zeros of the Kac determinant, expressed for instance by eqn (11.2.7), can be graphically associated to a set of points with integer coordinates (r, s) of the first quadrant of a cartesian plane. For the nature of these points, it is naturally to define a lattice on this plane, as in Fig. 11.1. This graphical representation is extremely useful to illustrate some remarkable properties of the Kac formula of the minimal models.

The dashed line in Fig. 11.1 has a slope $\tan\theta = -\alpha_+/\alpha_-$. If $\delta_{r,s}$ stands for the distance of a point (r, s) of the lattice from this straight line, the zeros of the Kac

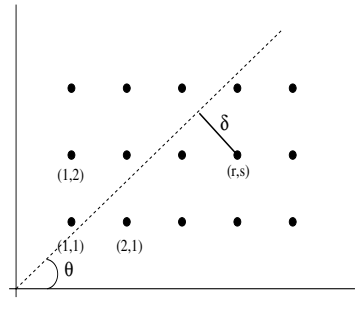


Fig. 11.1 *Kac table.*

determinant can be written as

$$\Delta_{r,s} \,=\, \Delta_0 + \frac{1}{4}(\alpha_+ + \alpha_-)^2\, \delta_{r,s}^2. \tag{11.4.4}$$

When the slope is irrational, the line obviously never meets a point of the lattice. On the contrary, if it is rational, there exist two coprime integers p and q, with $p > q$, such that

$$p\alpha_- + q\alpha_+ \,=\, 0. \tag{11.4.5}$$

In this case, the line passes through the point (q, p). In the rational case, it is easy to see that the zeros of the Kac determinant satisfy the properties

$$\begin{aligned} \Delta_{r,s} \,&=\, \Delta_{r+q,s+p}, \\ \Delta_{r,s} \,&=\, \Delta_{q-r,p-s}. \end{aligned} \tag{11.4.6}$$

These relations can be easily interpreted from a geometrical point of view: the point (r, s) of the lattice has the same distance from the line of slope p/q of the infinite series of points $(r + nq, s + np)$ obtained by reflection with respect to the same line.

We can express the central charge and conformal weights according to the formula (11.4.1) that identifies the most general minimal models. Note that, in addition to eqns (11.4.6), there are also the relations

$$\begin{aligned} \Delta_{r,s} + rs \,&=\, \Delta_{q+r,p-s} \,=\, \Delta_{q-r,p+s} \\ \Delta_{r,s} + (q-r)(p-s) \,&=\, \Delta_{r,2p-s} \,=\, \Delta_{2q-r,s}. \end{aligned} \tag{11.4.7}$$

These expressions imply that the null-vector at the level $N = rs$ of the conformal family $V_{r,s}$ is itself a *highest weight vector* of the Virasoro algebra, because its conformal weight is also expressed in terms of the Kac table! Moreover, besides the null-vector at the level rs, the conformal family $V_{r,s}$ also contains another null-vector at the level $(q - r)(p - s)$. In turn, these two null-vectors generate additional null-vectors and so on. Therefore, inside the conformal family of the primary field $\phi_{r,s}$, there is an infinite nested structure of null-vectors. This null-vector hierarchy deeply influences both the correlation functions and the characters of such primary operators.

11.4.2 Differential Equations

For the minimal models, either unitary or non-unitary, the conformal weights coincide with the zeros of the Kac determinant. Let's study how this circumstance leads to a result of great relevance for the correlation functions of their primary fields.

The primary field associated to $\Delta_{r,s}$ has its first null-vector at the level $N = rs$: this vector is expressed by a particular linear combination of the $P(rs)$ descendant states $\phi_{r,s}^{(n_1,n_2,\ldots)}$ of $\phi_{r,s}$ present at that level. Denoting the null-vector by $\phi_{r,s}^{null}$, its general expression is

$$\phi_{r,s}^{null}(z) \,=\, [a_1 L_{-1}^{rs} + a_2 L_{-1}^{rs-2} L_{-2} + \cdots a_{rs} L_{-rs}]\phi_{r,s}(z) \,=\, \sum a_i \phi_{r,s}^{(n_1,n_2,\ldots)}, \tag{11.4.8}$$

where all the coefficients a_i can be fixed by imposing the linear dependence of the vectors involved in the expression above. Any correlation functions in which such a

null-vector enters obviously vanishes

$$\langle \phi_{r,s}^{null}(z)\phi_1(z_1)\dots\phi_n(z_n)\rangle \;=\; 0. \tag{11.4.9}$$

On the other hand, we have seen in Section (10.8.2) that the correlation functions of the descendant fields $\phi_\Delta^{(n_1,n_2,\dots)}$ at the level N of a primary field are obtained applying a linear differential operator of order N to the correlation functions of the primary fields alone. Since the null-vector is also expressed by a linear combination of descendant field, once we identify the linear differential operator associated to each of them and collect all the terms, we arrive at the following important conclusion: by virtue of the null-vector at level rs, the correlation functions of the primary field $\phi_{r,s}(z)$ are solutions of a *linear differential equation* of order rs:

$$\mathcal{D}_{rs}\langle \phi_{r,s}(z)\phi_1(z_1)\dots\phi_n(z_n)\rangle \;=\; 0. \tag{11.4.10}$$

We have previously observed that the null-vectors of the conformal family $V_{r,s}$ are infinite in number and organized in a nested structure. There is, for instance, another null-vector at the level $(q-r)(p-s)$ and this implies that the correlation functions of the field $\phi_{r,s}$ are also solutions of another linear differential equation of order $(q-r)(p-s)$. All other null-vectors lead to an infinite hierarchy of linear differential equations satisfied by these correlators

$$\mathcal{D}_a\langle \phi_{r,s}(z)\phi_1(z_1)\dots\phi_n(z_n)\rangle \;=\; 0, \tag{11.4.11}$$

whose order a is equal to the level a of the various null vectors: the explicit form can be determined making use of the linear combination of the null-vector in terms of the descendant fields at the level a.

For this underlying structure of linear differential operators, not surprisingly the OPE of the primary fields of the minimal models are severely constrained.

11.4.3 Operator Product Expansion and Fusion Rules

Let's initially focus our attention on the conformal field $\phi_{1,2}$ of the minimal models. Its first null vector occurs at the level $N=2$ and its explicit expression is

$$\phi_{1,2}^{null}(z) \;=\; \left[L_{-2} - \frac{3}{2(2\Delta_{1,2}+1)}L_{-1}^2\right]\phi_{1,2}(z). \tag{11.4.12}$$

Hence the correlation functions of this field satisfy the linear differential equation

$$\left\{\frac{3}{2(2\Delta_{1,2}+1)}\frac{\partial^2}{\partial z^2} - \sum_{i=1}^{n}\left[\frac{\Delta_i}{(z-z_i)^2} + \frac{1}{z-z_i}\frac{\partial}{\partial z_i}\right]\right\}\langle \phi_{1,2}(z)\phi_1(z_1)\dots\phi_n(z_n)\rangle \;=\; 0. \tag{11.4.13}$$

Consider now the operator product expansion of the field $\phi_{1,2}(z)$ with any other primary field $\phi_\Delta(z_1)$

$$\phi_{1,2}(z)\phi_\Delta(z_1) \;=\; \sum_{\Delta'} C_{(1,2),\Delta}^{\Delta'}\,(z-z_1)^{\Delta'-\Delta-\Delta_{1,2}}\,[\phi_{\Delta'}(z_1)+\cdots]\,. \tag{11.4.14}$$

This expansion has to be compatible with the differential equation satisfied by the field $\phi_{1,2}(z)$. Inserting this operator expansion in eqn (11.4.13) and the most singular term

in this expression equal to zero, we arrive at the characteristic equation associated to the differential equation

$$\frac{3x(x-1)}{2(2\Delta_{1,2}+1} - \Delta + x \;=\; 0, \tag{11.4.15}$$

where $x = \Delta' - \Delta - \Delta_{1,2}$. This is a second-order algebraic equation that shows that the OPE of $\phi_{1,2}$ with any other conformal field cannot have more than two conformal families. Furthermore, it permits us to determine the conformal weight of the primary field $\phi_{\Delta'}$ that is generated by the operator expansion with ϕ_Δ. Remarkably enough, if Δ is expressed by one value of the Kac formula, i.e. $\Delta = \Delta_{r,s}$, then the solutions of the characteristic equation also belong to the set of values of the Kac table! Namely, if $\Delta = \Delta_{r,s}$, the two solutions of the quadratic equations are given by

$$\Delta' \;=\; \{\Delta_{r,s-1}, \Delta_{r,s+1}\}\,. \tag{11.4.16}$$

Simplifying the notation of the OPE to its skeleton form, we can write

$$\phi_{1,2} \times \phi_{r,s} \;=\; [\phi_{r,s-1}] + [\phi_{r,s+1}], \tag{11.4.17}$$

and, in particular

$$\phi_{1,2} \times \phi_{1,2} \;=\; [\mathbf{I}] + [\phi_{1,3}]. \tag{11.4.18}$$

In other words, only degenerate fields enter the OPE of $\phi_{1,2}$ with any of the degenerate field $\phi_{r,s}$. It must be stressed, though, that the result above does not take into account the actual value of the structure constant: as it is, it only states which conformal families may possibly enter the OPE. As we will see later, the vanishing of one or more of the structure constants further reduces the number of conformal families. In this case we say that a *truncation* of the OPE has occurred.

Repeating the same analysis for the degenerate field $\phi_{2,1}$ the same conclusions are reached, the only difference is the swapping of the relative indices. With the same notation introduced above, we have in fact

$$\begin{aligned} \phi_{2,1} \times \phi_{r,s} \;&=\; [\phi_{r-1,s}] + [\phi_{r+1,s}] \\ \phi_{2,1} \times \phi_{2,1} \;&=\; [\mathbf{I}] + [\phi_{3,1}]. \end{aligned} \tag{11.4.19}$$

The graphical interpretation of these results is immediate: by using iteratively the operator product expansion of the operators $\phi_{1,2}$ and $\phi_{2,1}$ it is possible to generate all the other degenerate fields of the minimal models, i.e. we can move horizontally and vertically along the Kac lattice, visiting all its points, as shown in Fig. 11.2.

An explicit example of the phenomenon of truncation is provided by the OPE of the fields $\phi_{1,2}$ and $\phi_{2,1}$. Using the formulas above, either with respect to the first field and the second one, we have

$$\begin{aligned} \phi_{1,2} \times \phi_{2,1} \;&=[\phi_{0,2}] + [\phi_{2,2}] \\ \phi_{1,2} \times \phi_{2,1} \;&= [\phi_{2,0}] + [\phi_{2,2}]. \end{aligned} \tag{11.4.20}$$

Since the two different ways should lead to the same result, the structure constants that involve both the fields $\phi_{0,2}$ and $\phi_{2,0}$ must vanish. So, we are in the presence of a

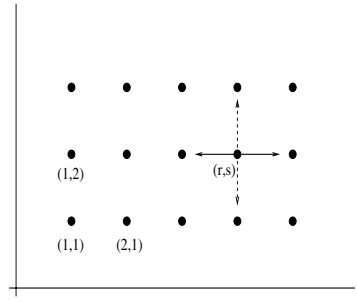


Fig. 11.2 *Action of the operators $\phi_{1,2}$ (dashed line) and $\phi_{2,1}$ (continous line).*

truncation of the operator product expansion of $\phi_{1,2} \times \phi_{2,1}$ that reduces then to the expression

$$\phi_{1,2} \times \phi_{2,1} \; = \; [\phi_{2,2}]. \tag{11.4.21}$$

We can now iteratively insert the operators $\phi_{1,2}$ and $\phi_{2,1}$, using at the same time the associativity of the operator algebra, to compute the fusion rules of the other degenerate fields. Consider, for instance, the product of three fields $\phi_{2,1} \times \phi_{2,1} \times \phi_{r,s}$. Applying the fusion rules (11.4.19) twice, we get

$$\phi_{3,1} \times \phi_{r,s} \; = \; [\phi_{r+2,s}] + [\phi_{r,s}] + [\phi_{r-2,s}]. \tag{11.4.22}$$

Analogously, consider $\phi_{1,2} \times \phi_{1,2}\phi_{r,s}$. Using eqn (11.4.17) twice, we arrive at

$$\phi_{1,3} \times \phi_{r,s} \; = \; [\phi_{r,s+2}] + [\phi_{r,s}] + [\phi_{r,s-2}]. \tag{11.4.23}$$

It is easy to check that these are precisely the fusion rules that are compatible with the linear differential equations of the third-order satisfied by the fields $\phi_{3,1}$ and $\phi_{1,3}$, as proposed in Problem 1.

Fusion rule. Carrying on a similar analysis for the other fields, one can reach the general formula of the *fusion rules* relative to two arbitrary degenerate fields of the Kac table

$$\phi_{r_1,s_1} \times \phi_{r_2,s_2} \; = \sum_{r_3=|r_1-r_2|+1}^{\min(r_1+r_2-1,2q-1-r_1-r_2)} \; \sum_{s_3=|s_1-s_2|+1}^{\min(s_1+s_2-1,2p-1-s_1-s_2)} [\phi_{r_3,s_3}] \tag{11.4.24}$$

where both indices are summed in steps of 2. These fusion rules can be written in a more transparent form noting that they are similar to the fusion rules of two irreducible representations of spins j and j' of $SU(2)$. To this end, it is useful to use as indices $r_i = 2j_1 + 1$ and $r_i = 2j_i' + 1$. This similarity explains the null values of the structure constants for all odd values of r (corresponding to the vector representations of $SU(2)$) as well as their vanishing when there are two even indices and one odd (corresponding to two spinor representations and one vector representation). However, there is an important difference between the fusion rules of conformal field theory and those of $SU(2)$, as clearly shown by the upper restriction of the two sums that involve the

parameters q and p. In fact, the fusion rules of the minimal models are not those of $SU(2)$ but those of the *quantum group* $SU_q(2)$. In the minimal models there are two quantum groups:[3] the first $SU_{q_1}(2)$ with $q_1 = \exp(i\pi q/p)$ acting on the rows, the second $SU_{q_2}(2)$ with $q_2 = \exp(i\pi p/q)$ acting on the column. Since q_1 and q_2 are both roots of unity, the representations of the corresponding quantum groups get restricted and their composition laws are expressed by the fusion rules given above.

11.4.4 Verlinde Algebra

It is important to formulate in a more abstract way the fusion rules for better analyzing their properties. Denoting by ϕ_i a generic primary field, the algebraic structure of the fusion rules can be expressed by simply putting to 1 all the non-zero structure constants by

$$\phi_i \times \phi_j \;=\; \sum_k N_{ij}^k \, \phi_k, \tag{11.4.25}$$

where N_{ij}^k is a set of integers that express the number of independent fusions that relate ϕ_i and ϕ_j to the field ϕ_k. For the minimal models, these numbers can only be 0 and 1, but for conformal theories with an extended algebra they can be generically integers.

From their definition, the quantities N_{ij}^k are symmetric with respect to the indices i and j. The associativity condition of the algebra (11.4.25) leads to a quadratic condition for the quantities N_{ij}^k: this can be derived by the two possible ways of applying eqn (11.4.25) to the product of three fields

$$\phi_i \times \phi_j \times \phi_l \;=\; \begin{cases} \sum_k N_{ij}^k \phi_k \times \phi_l \;=\; \sum_{k,p} N_{ij}^k N_{kl}^p \phi_p \\ \phi_i \times \sum_k N_{jl}^k \phi_k \;=\; \sum_{k,p} N_{jl}^k N_{ik}^p \phi_p. \end{cases} \tag{11.4.26}$$

Using the matrix notation $(N_i)_j^k = N_{ij}^k$ and the symmetry with respect to the indices i, j, the identity of the two expressions above reads

$$N_i \, N_l \;=\; N_l \, N_i. \tag{11.4.27}$$

This condition can also be expressed as

$$N_j \, N_l \;=\; \sum_k N_{jl}^k \, N_k. \tag{11.4.28}$$

The commutativity of the matrices N_i, shown in eqn (11.4.27), implies that all these matrices can be diagonalized simultaneously and their eigenvalues $\lambda_i^{(n)}$ form a one-dimensional representation of the fusion rules. Note that the algebra (11.4.25), known as the *Verlinde algebra*, is very similar to the formula that appears in the theory of finite groups and that rules the composition law of their irreducible representations. Further properties of the Verlinde algebra can be found in Problem 6 at the end of the chapter.

[3] For the notation and the theory of quantum groups, see Section 18.9.

11.5 Coulomb Gas

Above we have seen that the correlation functions of the primary fields of the minimal models, for their null-vectors, satisfy a series of linear differential equations. Hence, to determine the correlators explicitly, one can adopt the following strategy:

1. find the explicit expression of the null-vectors;
2. translate this expression into the corresponding linear differential equation;
3. find its solutions.

All these steps can be explicitly implemented for the minimal models. However, there exists a more efficient way to find the final expressions of the correlators. The method has been proposed by Dotsenko and Fateev and it enables us to write down directly the final expression of the correlation functions without passing through the three steps given above. It is based on a modified version of the Coulomb gas in two dimensions. To explain what it consists of, it is necessary to discuss first the conformal field theory associated to a free massless bosonic field. Further details on this theory will be given in Section 12.4 of the next chapter.

11.5.1 Free Theory of a Bosonic Field

Consider the action of a free massless scalar field in two dimensions

$$\mathcal{S} \,=\, \frac{g}{16\pi}\int d^2x\,\partial_\mu\varphi\,\partial^\mu\varphi. \tag{11.5.1}$$

The propagator of this theory needs both an ultraviolet and an infrared cut-off, given respectively by a and R, and it can be written as

$$G(z,\bar z) \,=\, \langle\varphi(z,\bar z)\varphi(0,0)\rangle \,=\, \begin{cases} -\frac{2}{g}\,\log\frac{z}{a}-\frac{2}{g}\,\log\frac{\bar z}{a}, & z,\bar z\neq 0 \\ -\frac{4}{g}\,\log\frac{R}{a}, & z=\bar z=0. \end{cases} \tag{11.5.2}$$

Note that this correlator is also the Green function of a two-dimensional electrostatic problem, and for this reason, the formalism we are going to present is also known as the Coulomb gas approach. To simplify the formula to come, in this section we choose for the coupling constant the value $g=1$.

As is evident from the form of its propagator, $\varphi(x)$ is not a conformal field. However, conformal fields can be constructed in terms of some of its composite operators, as for instance all derivative fields $\partial_z^n\partial_{\bar z}^m\varphi$ or the exponential operators $\tilde V_\alpha = e^{i\alpha\varphi}$, also known as *vertex operators*. The quantity α entering the exponential is also called the *charge parameter*. Let's focus attention on the analytic part of the theory. It is easy to see that the n-point correlation functions of the vertex operators can be computed by means of Wick's theorem or directly by the functional integral, for the action is quadratic,

$$\prod_{i=1}^n \langle\tilde V_{\alpha_i}(z_i)\,\rangle \,=\, \Big(\frac{a}{R}\Big)^{(\sum_i^n\alpha_i)^2}\,\prod_{i<j}\Big(\frac{z_{ij}}{a}\Big)^{-2\alpha_i\alpha_j}. \tag{11.5.3}$$

We need to get rid of the two cut-offs. To eliminate the dependence from the ultraviolet cut-off, it is sufficient to subtract all the tadpole contributions coming from the

contractions of the the field $\varphi(z)$ with itself. This can be implemented defining the renormalized vertex operator

$$\tilde{V}_\alpha(z) \to V_\alpha = \lim_{a\to 0} a^{-\alpha^2/2}\, e^{i\alpha\varphi} \equiv : e^{i\alpha\varphi} : . \tag{11.5.4}$$

To eliminate the dependence on the infrared cut-off, so as to have a non-zero limit of the correlation functions when $R \to \infty$, it is necessary to impose the neutrality conditions of all the charges

$$\sum_{i=1}^{n} \alpha_i \,=\, 0. \tag{11.5.5}$$

This is in agreement with the well-known result of statistical mechanics that a system of electric charges is unstable unless it has a zero total charge.

Looking at the two-point correlation functions of the renormalized vertex operators satisfying the neutrality condition

$$\langle V_\alpha(z) V_{-\alpha}(w)\rangle \,=\, \frac{1}{(z-w)^{2\alpha^2}}, \tag{11.5.6}$$

we can extract the conformal weight of the two vertex operators $V_{\pm\alpha}(z)$, given by

$$\Delta_\alpha \,=\, \Delta_{-\alpha} \,=\, \alpha^2. \tag{11.5.7}$$

The analytic component of the stress–energy tensor associated to the action (11.5.1) is

$$T(z) \,=\, -\frac{1}{4} : (\partial_z \varphi)^2 :, \tag{11.5.8}$$

and, as we have previously seen, the central charge of this system is $C = 1$.

One may wonder if it would be possible to modify the Coulomb gas in such a way as to have values of the central charge different from $C = 1$ and conformal weights equal to those of the Kac table of the minimal models. This is indeed possible, as discussed in the next section.

11.5.2 Modified Coulomb Gas

Consider a vertex operator with charge $-2\alpha_0$ and suppose we insert it in a correlation function, moving its position to infinity (i.e. on the north pole of the Riemann sphere associated to the complex plane) using the prescription (10.8.10). With this procedure we can define a new set of correlators given by

$$\langle\langle V_{\alpha_1}(z_1) \dots V_{\alpha_n}(z_n)\rangle\rangle \,\equiv\, \lim_{R\to\infty} R^{8\alpha_0^2}\, \langle V_{-2\alpha_0}(R)\, V_{\alpha_1}(z_1) \dots V_{\alpha_n}(z_n)\rangle. \tag{11.5.9}$$

Note that to recover the translation invariant of these quantities (which is expressed by the dependence on the coordinates only through the differences $z_i - z_j$), it is necessary

to place the vertex operator $V_{-2\alpha_0}$ just at infinity. This new set of correlation functions for the vertex operators $V_{\alpha_1} \dots V_{\alpha_n}$ satisfies a different neutrality condition

$$\sum_{i=1}^{n} \alpha_i \;=\; 2\alpha_0. \tag{11.5.10}$$

In agreement with that, the two-point correlation function of the vertex operators is now given by[4]

$$\langle V_\alpha(z) V_{2\alpha_0-\alpha}(w)\rangle \;=\; \frac{1}{(z-w)^{2\alpha(\alpha-2\alpha_0)}}. \tag{11.5.11}$$

The new conformal weights are then

$$\Delta_\alpha \;=\; \Delta_{2\alpha_0-\alpha} \;=\; \alpha(\alpha-2\alpha_0). \tag{11.5.12}$$

This result can be directly confirmed by the operator product expansions of the vertex operators with the new expression of the stress–energy tensor. To derive the new stress–energy tensor, one should observe that placing a charge at infinity is equivalent to modifying the original action (11.5.1) in such a way as to make anomalous the original $U(1)$ symmetry implemented by $\varphi \to \varphi + \eta$. In a generalized system of coordinates, this can be realized by coupling the field φ to the scalar curvature R of the space manifold

$$\mathcal{S} \;=\; \frac{1}{8\pi} \int d^2x\, \sqrt{g}(\partial_\mu\varphi\partial^\mu\varphi + 2i\alpha_0 R\varphi). \tag{11.5.13}$$

This new action is no longer invariant under a shift of the φ and its variation becomes

$$\delta\mathcal{S} \;=\; i\,\frac{\alpha_0}{4\pi} \int d^2x\sqrt{g}R. \tag{11.5.14}$$

In two dimensions this is just a topological term that can be computed by the Gauss–Bonnet theorem

$$\int d^2\sqrt{g}R \;=\; 8\pi(1-h), \tag{11.5.15}$$

where h is the number of handles of the Riemann surface on which is defined the field theory (11.5.13) and for a sphere we have $h = 0$. Corresponding to the new action (11.5.13), there is a new version of the stress–energy tensor given by Noether's theorem and its analytic component reads

$$T(z) \;=\; -\frac{1}{4}(\partial\varphi)^2 + i\alpha_0\,\partial^2\varphi. \tag{11.5.16}$$

Its two-point function is given by

$$\langle T(z)T(w)\rangle \;=\; \frac{1-24\alpha_0^2}{2(z-w)^4}. \tag{11.5.17}$$

[4] In the following instead of using the notation $\langle\langle \dots \rangle\rangle$ we switch back to the simpler notation $\langle \dots \rangle$ to denote the correlation functions also in the modified Coulomb gas system.

In conclusion, with a charge at infinity, the central charge of the theory assumes a different value from the original $C = 1$ and is now

$$C \;=\; 1 - 24\alpha_0^2. \tag{11.5.18}$$

The operator product expansion of the new stress–energy tensor with the vertex operator $V_\alpha(w)$ becomes

$$\begin{aligned} T(z) : e^{i\alpha\phi(w)} : &= \frac{\alpha^2 - 2\alpha_0\alpha}{(z-w)^2} : e^{i\alpha\phi} : + \frac{i\alpha}{(z-w)} : \partial\varphi\, e^{i\alpha\varphi} : + : T(z) e^{i\alpha\varphi} : \\ &= \frac{\alpha(\alpha - 2\alpha_0)}{(z-w)^2} V_\alpha(w) + \frac{1}{z-w}\partial V_\alpha(w) + \dots \end{aligned} \tag{11.5.19}$$

This formula clearly shows that, in the presence of the charge at infinity, the conformal weight of the vertex operator V_α is effectively given by eqn (11.5.12).

It is also important to discuss the conformal trasformation of the scalar field. In the absence of the charge at infinity, under the transformation $z \to f(z)$, $\phi(z)$ trasforms as $\phi(z) \to \phi(f(z))$. But with the charge at infinity, there is a change of the boundary conditions and the field transforms instead as

$$\phi(z) \to \phi(f(z)) + 2i\alpha_0 \ln f'(z), \tag{11.5.20}$$

whose infinitesimal form is

$$\delta\phi(z) \;=\; \epsilon(z)\partial_z\phi(z) + 2i\alpha_0\epsilon'(z). \tag{11.5.21}$$

11.5.3 Screening Operators

The modified Coulom gas formalism allows us to describe the conformal models with central charge less that 1 and, in particular, the minimal models. In this approach the primary fields are associated to the vertex operators. Notice that, assigning the conformal weight Δ of the primary field, there are however *two* different vertex operators V_α and $V_{2\alpha_0-\alpha}$ that can be put in correspondence with it, because the charges satisfy the quadratic condition $\Delta = \alpha(\alpha - 2\alpha_0)$. The two-point correlation function of the primary field ϕ_Δ is different from zero but, in the formalism of the Coulomb gas, it can be computed in four different ways, namely

$$\langle \phi_\Delta(z)\phi_\Delta(0)\rangle \to \begin{cases} \langle V_\alpha(z) V_{2\alpha_0-\alpha}(w)\rangle \\ \langle V_{2\alpha_0-\alpha}(z) V_\alpha(w)\rangle \\ \langle V_\alpha(z) V_\alpha(w)\rangle \\ \langle V_{2\alpha_0-\alpha}(z) V_{2\alpha_0-\alpha}(w)\rangle. \end{cases} \tag{11.5.22}$$

The first two expressions automatically satisfy the neutrality condition (11.5.10): hence they are different from zero and give rise to the usual expression $\langle \phi_\Delta(z)\phi_\Delta(w)\rangle$. On the contrary, the last two expressions do not fulfill the neutrality condition (11.5.10) and are therefore zero. There is then the problem of correcting this drawback, in such a way that one can equivalently use either $V_\alpha(z)$ or $V_{2\alpha_0-\alpha}(z)$ to represent the primary field $\phi_\Delta(z)$.

The solution to this problem consists of the introduction of the so-called *screening operators*. Such operators, once inserted in the correlators, should be able to absorb the extra charge that spoils the neutrality condition without altering, though, the conformal properties of the correlators. To satisfy such conditions, the screening operators must have zero conformal weight but non-zero charge. It is impossible to fulfill these conditions in terms of local operators but there is no problem in finding them in terms of a closed contour integral of an operator of conformal weight equal to 1. Hence, denoting by $V_{\alpha_\pm}$ the vertex operators of these fields with conformal weight $\Delta = 1$, we can unite

$$Q_\pm \,=\, \oint dz\, V_{\alpha_\pm}(z), \tag{11.5.23}$$

where the charges $\alpha_\pm$ satisfy the equation

$$\alpha_\pm(\alpha_\pm - 2\alpha_0) \,=\, 1. \tag{11.5.24}$$

The solutions of this equation are

$$\alpha_\pm \,=\, \alpha_0 \pm \sqrt{\alpha_0^2 + 1}. \tag{11.5.25}$$

Note that

$$\begin{aligned} \alpha_+ + \alpha_- &= 2\alpha_0 \\ \alpha_+\, \alpha_- &= -1. \end{aligned} \tag{11.5.26}$$

Inserting an integer number of these operators $Q_\pm$ in the correlation functions, we can therefore screen the extra charge present in their vertex operator representation. Obviously this cancellation mechanism takes place only if the extra charge is expressible in terms of integer multiplies of $Q_\pm$. Consider, for instance, the third expression in (11.5.22): inserting now an integer number of screening operators, it becomes

$$\langle V_\alpha(z) V_\alpha(w) Q_+^r\, Q_-^s \rangle \tag{11.5.27}$$

and the neutrality conditions translates into the condition

$$2\alpha + r\alpha_+ + s\alpha_- \,=\, 2\alpha_0 \,=\, \alpha_+ + \alpha_-. \tag{11.5.28}$$

Hence we can screen the extra charge of the original expression only if the charges α present in the system satisfy the quantization condition

$$\alpha \,=\, \alpha_{r,s} \,=\, \frac{1}{2}(1-r)\alpha_+ + \frac{1}{2}(1-s)\alpha_-. \tag{11.5.29}$$

In this case, there is complete equivalence of the operators $V_{\alpha_{r,s}}$ and $V_{2\alpha_0 - \alpha_{r,s}}$, after all, the logic consistency of the modified Coulomb gas. Corresponding to the values (11.5.29) the conformal weights of the fields are given by

$$\Delta_{r,s} \,=\, \frac{1}{4}(r\alpha_+ + s\alpha_-)^2 - \alpha_0^2. \tag{11.5.30}$$

They assume the form (11.2.7) given by the Kac table and the symmetry $\alpha \to 2\alpha_0 - \alpha$ translates into the transformation $(r, s) \to (-r, -s)$. To recover the minimal models,

we need, however, to impose an additional quantization condition on the charges $\alpha_\pm$ $(p > q)$

$$q\alpha_+ + p\alpha_- \;=\; 0, \tag{11.5.31}$$

where p and q are two coprime integers. With this last condition, it is easy to see that eqns (11.5.18) and (11.5.30) reproduce the central charge and the conformal weights of the minimal models, eqn (11.4.1), and we have moreover the periodicity relation

$$\alpha_{r,s} \;=\; \alpha_{r+q,s+p}. \tag{11.5.32}$$

In the next section we discuss how to compute the correlation functions of the minimal models using the Coulomb gas formalism.

11.5.4 Correlation Functions

The correlation functions of the primary fields $\phi_{r,s}$ satisfy an infinite number of linear differential equations in coincidence of their null-vector hierarchy. The main advantage of the Coulomb gas formalism is to provide the solutions of the differential equations directly in terms of their integral representation. In this section we initially discuss the implementation of this formalism for the simplest cases of the correlators of the fields $\phi_{1,2}$ and $\phi_{2,1}$.

Consider the holomorphic part of the four-point correlation function of the primary field

$$G(z_1, z_2, z_3, z_4) \;=\; \langle \phi_{n,m}(z_1)\phi_{1,2}(z_2)\phi_{1,2}(z_3)\phi_{n,m}(z_4)\rangle. \tag{11.5.33}$$

This quantity is surely different from zero since there exists a common conformal channel – given by the family of the identity operator – in the operator product expansion of $\phi_{1,2} \times \phi_{1,2}$ and $\phi_{n,m} \times \phi_{n,m}$. Since the primary fields $\phi_{r,s}$ can be associated either to $V_{\alpha_{r,s}}$ or $V_{2\alpha_0-\alpha_{r,s}}$, the correlation function (11.5.33) admits 16 different expressions in terms of the vertex operators of the Coulomb gas. Out of these expressions, the one that needs the least number of screening operators is the following[5]

$$\langle V_{\alpha_{n,m}}(z_1)V_{\alpha_{1,2}}(z_2)V_{\alpha_{1,2}}(z_3)V_{2\alpha_0-\alpha_{n,m}}(z_4)\rangle.$$

The extra charge present in this representation is $2\alpha_{1,2}$, thus its screening requires only one operator Q_-. This leads to the integral representation

$$\begin{aligned}&\langle \phi_{n,m}(z_1)\phi_{1,2}(z_2)\phi_{1,2}(z_3)\phi_{n,m}(z_4)\rangle \\ &\quad= \oint_C dv\, \langle V_{\alpha_{1,2}}(z_1)V_{\alpha_{1,2}}(z_2)V_{\alpha_{n,m}}(z_3)V_{2\alpha_0-\alpha_{n,m}}(z_4)V_{\alpha_-}(v)\rangle.\end{aligned} \tag{11.5.34}$$

From the analytic nature of the integrand as a function of v, the integral does not depend on the precise shape of the contour, although it must be chosen to enclose the points $z_1, \ldots, z_4$ otherwise it could be shrunk to a point, with a vanishing result.

[5]The other expressions lead to the integral representation of the solutions of the higher order differential equations satisfied by the same correlator.

Performing the expectation values of the vertex operators in the integrand by using Wick's theorem

$$\prod_{i=1}^{n}\langle V_{\alpha_i}(z_i)\rangle \;=\; \prod_{i<j}^{n}(z_{ij})^{2\alpha_i\alpha_j}$$

one has

$$\begin{aligned}
&\oint_{C_1} du_1 \langle V_{\alpha_{n,m}}(z_1)V_{\alpha_{1,2}}(z_2)V_{\alpha_{1,2}}(z_3)V_{2\alpha_0-\alpha_{n,m}}(z_4)V_{\alpha_+}(u_1)\rangle \\
&= (z_{12}z_{13})^{2\alpha_{1,2}\alpha_{n,m}}(z_{23})^{2\alpha_{1,2}^2}(z_{14})^{2\alpha_{n,m}(2\alpha_0-\alpha_{n,m})}(z_{24}z_{34})^{2\alpha_{1,2}(2\alpha_0-\alpha_{n,m})} \\
&\quad\times \oint dv (v-z_1)^{2\alpha_--\alpha_{n,m}}[(v-z_2)(v-z_3)]^{2\alpha_--\alpha_{1,2}}(v-z_4)^{2\alpha_-(2\alpha_0-\alpha_{n,m})}.
\end{aligned}$$

To simplify the expressions, let's use the Moebius invariance to fix three out of the four points of the correlator: we then impose $z_1 = 1$, $z_3 = 0$ and $z_4 = \infty$, leaving as a free position $z_2 = z$. In this way we arrive at

$$\begin{aligned}
\langle\phi_{n,m}(1)\phi_{1,2}(z)\phi_{1,2}(0)\phi_{n,m}(\infty)\rangle &= z^{2\alpha_{1,2}^2}(1-z)^{2\alpha_{1,2}\alpha_{n,m}} \\
&\quad\times \oint_C dv\, v^{2\alpha_--\alpha_{1,2}}(v-1)^{2\alpha_--\alpha_{n,m}}(v-z)^{2\alpha_--\alpha_{1,2}}.
\end{aligned}$$

Consider now the integral

$$F(z,a,b,c) \;=\; \oint dv\, v^a(v-1)^b(v-z)^c, \tag{11.5.35}$$

with

$$a \;=\; 2\alpha_-\alpha_{1,2} \quad b \;=\; 2\alpha_-\alpha_{n,m} \quad c \;=\; 2\alpha_-\alpha_{1,2}. \tag{11.5.36}$$

The integrand, as a function of the complex variable v, has branch cuts at the points $v = 0, 1, z$. To be closed, the integration contour must cross each cut twice. There are several ways to choose such a contour, although only two of them are independent: the most convenient choice consists of the paths shown in Fig. 11.3. If the integral along these paths converges, the first contour can be restricted to the interval $C_1 = [1,\infty]$ while the second to the interval $C_2 = [0,z]$. With this choice of the paths of integration, we have defined two different functions

$$\begin{aligned}
I_1(z,a,b,c) \;&=\; \int_1^\infty dv\, v^a(v-1)^b(v-z)^c \\
&=\; \frac{\Gamma(-a-b-c-1)\Gamma(b+1)}{\Gamma(-a-c)}\,F(z;-c,-a-b-c-1,-a-c)
\end{aligned} \tag{11.5.37}$$

$$\begin{aligned}
I_2(z,a,b,c) \;&=\; \int_0^z dv\, v^a(v-1)^b(z-v)^c \\
&=\; \frac{\Gamma(a+1)\Gamma(c+1)}{\Gamma(a+c+2)}\,z^{1+a+c}\,F(z;-b,a+1,a+c+2)
\end{aligned}$$

where $F(z,\alpha,\beta,\gamma)$ is a hypergeometric function (see Appendix 11A for its properties). The two functions above set up the vector space of the solutions of the second-order

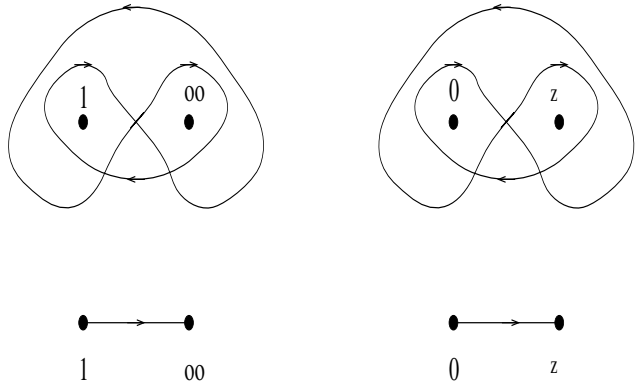


Fig. 11.3 *Independent contours.*

differential equation satisfied by the correlation functions of the primary field $\phi_{1,2}$. An analogous result is obtained if we consider the anti-analytic part of the correlator, so that the general form of the correlation function in the physical plane is expressed as a linear combination of the analytic and anti-analytic solutions of the differential equation

$$\begin{aligned} G(z,\bar{z}) &= \langle \phi_{n,m}(1,1)\phi_{1,2}(z,\bar{z})\phi_{1,2}(0,0)\phi_{n,m}(\infty)\rangle \\ &= |z|^{4\alpha_{1,2}}|1-z|^{4\alpha_{1,2}\alpha_{n,m}}\, Y(z,\bar{z}), \end{aligned} \tag{11.5.38}$$

where

$$Y(z,\bar{z}) = \sum_{i,j=1}^{2} \mathcal{X}_{ij} I_i(z) I_j(\bar{z}).$$

Monodromy invariance. The explicit expression of the coefficients $\mathcal{X}_{ij}$ can be obtained by the condition that the correlation functions is an unambiguous function of the points in the plane, i.e. independent of the paths by which we reach the points. To implement such a condition, we have to analyze the monodromy group associated to the functions $I_i(z)$. They have the singular points $z = 0, 1, \infty$. If we make an analytic continuation along a close contour that encloses one of these points, as in Fig. 11.4, they do not return to their original value, instead the new function – which is still a solution of the linear differential equation – is expressed as a linear combination of the I_i's

$$\begin{aligned} I_i(z) &\to (g_0)_{ij} I_j(z) \\ I_i(z) &\to (g_1)_{ij} I_j(z). \end{aligned} \tag{11.5.39}$$

Note that it is sufficient to consider the monodromy properties only around the points $z = 0, 1$, since those around the point $z = \infty$ follow from them. The monodromy matrix $(g_0)_{ij}$, in our case, is diagonal:

$$g_0 = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i(a+c+1)} \end{pmatrix}. \tag{11.5.40}$$

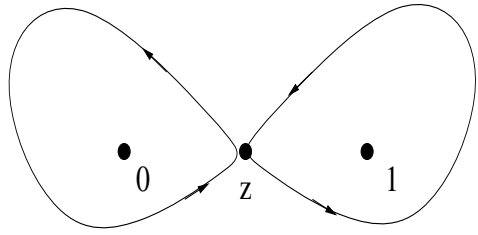


Fig. 11.4 *Analytic continuation of the functions $I_i(z)$ around the singular points $z = 0, 1$.*

Imposing the invariance of the correlation function (11.5.38) under this transformation immediatety leads to the conditions $\mathcal{X}_{12} = \mathcal{X}_{21} = 0$ and, after that, the function $Y(z, \bar{z})$ reduces to diagonal form:

$$Y(z, \bar{z}) \; = \; \sum_{i=1}^{2} \mathcal{X}_{ii} I_i(z) I_i(\bar{z}). \tag{11.5.41}$$

To determine the remaining coefficients $\mathcal{X}_{11}$ e $\mathcal{X}_{22}$ we must impose the invariance under the monodromy transformation g_1. The simplest way to do that is to express initially the functions $I_i(z)$ in terms of another basis that has the series expansion in the variable $(1 - z)$

$$I_i(z) \; = \; \sum_{j=1}^{2} a_{ij} \hat{I}_j(1 - z). \tag{11.5.42}$$

For the hypergeometric functions this can be done using the Gauss formulas: with $s(x) \equiv \sin(\pi x)$ we have[6]

$$\begin{aligned} I_1(z; a, b, c) \; &= \; \frac{s(a)}{s(b+c)} \hat{I}_1(1 - z; b, a, c) - \frac{s(c)}{s(b+c)} \hat{I}_2(1 - z; b, a, c) \\ I_2(z; a, b, c) \; &= \; -\frac{s(a+b+c)}{s(b+c)} \hat{I}_1(1 - z; b, a, c) - \frac{s(b)}{s(b+c)} \hat{I}_2(1 - z; b, a, c). \end{aligned} \tag{11.5.43}$$

Substituting eqn (11.5.42) in (11.5.41) we get

$$Y(z, \bar{z}) \; = \; \sum_{i,j,k=1,2} \mathcal{X}_{ii} a_{ij} a_{ik} \hat{I}_j(1 - z) \hat{I}_k(1 - \bar{z}). \tag{11.5.44}$$

Since the monodromy matrix of the functions $\hat{I}_i(1 - z)$ around the point $z = 1$ is diagonal, the monodromy invariance implies that the quadratic form in $\hat{I}_i$ must be diagonal as well. Hence

$$\frac{\mathcal{X}_{11}}{\mathcal{X}_{22}} \; = \; -\frac{a_{21} a_{22}}{a_{12} a_{11}} \; = \; \frac{s(a+b+c) s(b)}{s(a) s(c)}. \tag{11.5.45}$$

[6]Note the exchange of the indices a and b in the functions on the right-hand side of these expressions.

It is clear that, besides an overall constant λ (simply related to the normalization of the operators and that cannot be fixed by the monodromy invariance), the correlation function is given by

$$G(z,\bar{z}) = \langle \phi_{n,m}(1,1)\phi_{1,2}(z,\bar{z})\phi_{1,2}(0,0)\phi_{n,m}(\infty)\rangle \;=\; \lambda\,|z|^{4\alpha_{1,2}}|1-z|^{4\alpha_{1,2}\alpha_{n,m}} \times \left[\frac{s(b)s(a+b+c)}{s(a+c)}|I_1(z;a,b,c)|^2 + \frac{s(a)s(c)}{s(a+c)}|I_2(z;a,b,c)|^2\right]. \quad (11.5.46)$$

λ can be fixed once we specify the normalization of the conformal fields that we choose to be

$$\langle \phi_\Delta(z,\bar{z})\phi_\Delta(w,\bar{w})\rangle = 1/|z-w|^{4\Delta}. \quad (11.5.47)$$

With the values of the parameters a, b, c of the hypergeometric functions given in eqn (11.5.36), it is now easy to see that the function $I_2(z;a,b,c)$ corresponds to the channel of the conformal family of the identity operator present in this correlator. This means that in the limit $z \to 0$ the singularity coming from this term is precisely $1/|z|^{4\Delta_{1,2}}$. Hence, using the numerical factor in the function $I_2(z;a,b,c)$, the value of λ that implements the normalization condition (11.5.47) is given by

$$\lambda \;=\; \left(\frac{\Gamma(a+c+2)}{\Gamma(a+1)\Gamma(c+1)}\right)^2 \frac{s(a+c)}{s(a)s(c)}. \quad (11.5.48)$$

Structure constants. We can now extract the exact value of some of the structure constants of the conformal operator algebra. The operator product expansion of the field $\phi_{1,2}$ with itself is

$$\phi_{1,2}(z,\bar{z})\phi_{1,2}(0,0) \;=\; \frac{1}{|z|^{4\Delta_{1,2}}}\{\mathbf{I}+\ldots\} + C^{(13)}_{(12,12)}\frac{1}{|z|^{2(2\Delta_{1,2}-\Delta_{1,3})}}\{\phi_{1,3}+\ldots\}. \quad (11.5.49)$$

Substituting this expression on the right-hand side of (11.5.46), in the limit $z \to 0$ we have

$$\frac{1}{|z|^{4\Delta_{1,2}}}\langle \phi_{n,m}(1)\phi_{n,m}(\infty)\rangle + C^{(13)}_{(12,12)}\frac{1}{|z|^{2(2\Delta_{1,2}-\Delta_{1,3})}}\langle \phi_{n,m}(1)\phi_{1,3}(0)\phi_{n,m}(\infty)\rangle. \quad (11.5.50)$$

The three-point correlator $\langle \phi_{n,m}(1)\phi_{1,3}(0)\phi_{n,m}(\infty)\rangle$ in this case is precisely equal to the structure constant $C^{(13)}_{(nm),(nm)}$. Comparing now (11.5.50) with the right-hand side of eqn (11.5.46) in the limit $z \to 0$, the singularity with power law $1/|z|^{(2(2\Delta_{1,2}-\Delta_{1,3})}$ is reproduced by the function $|I_1(z;a,b,c)|^2$. Taking into account the value of λ and the numerical factor in the definition of this function we thus arrive at the formula

$$C^{(13)}_{(12,12)}\,C^{(13)}_{(nm,nm)} \;=\; \frac{s(a+b+c)s(b)}{s(a)s(c)}\left(\frac{\Gamma(a+c+2)\,\Gamma(-a-b-c-1)\,\Gamma(b+1)}{\Gamma(a+1)\,\Gamma(c+1)\,\Gamma(-a-c)}\right)^2. \quad (11.5.51)$$

It is worth stressing that the analysis of the singularities of the four-point correlation functions allows us to determine only the product of the structure constants, in agreement with the associative nature of the operator product expansion discussed in

the previous chapter (see eqn (10.2.15) and Fig. 10.2). Plugging in the formula above, $n = 1$ and $m = 2$, we can determine (up to a sign that we choose to be positive) the structure constant $C^{(13)}_{(12,12)}$ is

$$C^{(13)}_{(12,12)} \;=\; \frac{\Gamma(2-2\rho)}{\Gamma(2\rho)}\,\left[-\frac{\gamma^3(\rho)}{\gamma(3\rho-1)}\right]^{1/2}\,\frac{\gamma(1-\rho)}{\gamma(2-3\rho)}, \tag{11.5.52}$$

where we introduce the notation

$$\rho \equiv \alpha_-^2, \quad \gamma[x] \equiv \Gamma(x)/\Gamma(1-x).$$

Once $C^{(13)}_{(12,12)}$ is known, we can now use eqn (11.5.51) to determine the other structure constant

$$C^{(13)}_{(nm,nm)} \;=\; \frac{\Gamma(2-2\rho)}{\Gamma(2\rho)}\,\left[-\frac{\gamma^3(\rho)}{\gamma(3\rho-1)}\right]^{1/2}\,\frac{\gamma(n+(1-m)\rho)}{\gamma(1+n-(1+m)\rho)}. \tag{11.5.53}$$

The exact expression of the correlator $\langle\phi_{nm}(1,1)\phi_{12}(z,\bar z)\phi_{12}(0,0)\phi_{nm}(\infty)\rangle$ permits us to easily derive other structure constants. In fact, going in the dual channel and studying the limit $z \to 1$, one can extract the structure constants $C^{(n,m\pm1)}_{(12,nm)}$. The simplest way to do such a computation is to express the functions $I_i(z)$ in terms of the functions $\hat I_i(1-z)$ and write the correlator as

$$\begin{aligned} G(z,\bar z) &= \langle\phi_{n,m}(1,1)\phi_{1,2}(z,\bar z)\phi_{1,2}(0,0)\phi_{n,m}(\infty)\rangle \\ &= \lambda\,\frac{s(c)s(a+b+c)+s(a)s(b)}{s(a+c)s^2(b+c)} \\ &\quad\times|z|^{4\alpha_{1,2}}|1-z|^{4\alpha_{1,2}\alpha_{n,m}}\Big[s(a)s(a+b+c)\,|\hat I_1(1-z;b,a,c)|^2 \\ &\quad+s(b)s(c)\,|\hat I_2(1-z;b,a,c)|^2\Big]. \end{aligned} \tag{11.5.54}$$

It is now necessary to use the OPE

$$\begin{aligned} \phi_{1,2}(z,\bar z)\phi_{n,m}(1,1) = C^{(n,m+1)}_{(12,nm)}\frac{1}{|z-1|^{2\gamma_+}}\{\phi_{n,m+1}+\ldots\} \\ +C^{(n,m-1)}_{(12,nm)}\frac{1}{|z-1|^{2\gamma_-}}\{\phi_{n,m-1}+\ldots\} \end{aligned}$$

with

$$\gamma_\pm \;=\; \Delta_{1,2}+\Delta_{n,m}-\Delta_{n,m\pm1}.$$

Substituting this formula on the left-hand side of (11.5.54), the first function $|\hat I_1(1-z,b,a,c)|^2$ is easily identified with the intermediate states coming from the conformal family $\phi_{n,m+1}$, whereas the second function $|\hat I_2(1-z;a,b,c)|^2$ is associated to the

intermediate states of the family of $\phi_{n,m-1}$. Using the value of λ given above and the normalization of the functions $\hat{I}_i$, one obtains the structure constants

$$C_{(12,nm)}^{(n,m+1)} = \left[\frac{\gamma(2-2\rho)\gamma(n-m\rho)}{\gamma(1-\rho)\gamma(1+n-(1+m)\rho)}\right]^{1/2} \tag{11.5.55}$$

$$C_{(12,nm)}^{(n,m-1)} = C_{(12,nm)}^{(n,m+1)} \sqrt{-\frac{\gamma(n+(1-m)\rho)}{\gamma(-n-(m+1)\rho)} \frac{\Gamma^2(m\rho+1-n)}{\Gamma^2(n+1-m\rho)}}. \tag{11.5.56}$$

Other correlators. Let's briefly comment on the computation of the other correlators, referring the reader to the original articles for their derivation. Suppose, for instance, we wish to compute the correlator of the primary field $\phi_{n,m}(z,\bar{z})$

$$G = \langle \phi_{n,m}(z_1,\bar{z}_1)\phi_{n,m}(z_2,\bar{z}_2)\phi_{n,m}(z_3,\bar{z}_3)\phi_{n,m}(z_4,\bar{z}_4)\rangle.$$

The first step consists of expressing the analytic and anti-analyitic parts in terms of the vertex operator representation that needs the smallest number possible of screening operators. A possible choice is

$$\oint_{C_1} du_1 \ldots \oint_{C_{n-1}} du_{n-1} \oint_{S_1} dv_1 \ldots \oint_{S_{m-1}} dv_{m-1}$$

$$\langle V_{\alpha_{n,m}}(z_1) V_{\alpha_{n,m}}(z_2) V_{\alpha_{n,m}}(z_3) V_{2\alpha_0-\alpha_{n,m}}(z_4) V_{\alpha_+}(u_1)\ldots$$
$$V_{\alpha_+}(u_{n-1}) V_{\alpha_-}(v_1)\ldots V_{\alpha_-}(v_{m-1})\rangle.$$

In this expression there are $n \times m$ independent contours that, correspondingly, define a similar number of independent functions $I_i(z)$ $(i = 1, 2, \ldots, nm)$. These functions span the vector space of the solutions of the linear differential equation of order $n \times m$ satisfied by the correlation function. Together with the anti-analytic part, we arrive at a linear combination of these functions that provides the most general solution

$$G = \sum_{i,j=1}^{nm} \mathcal{X}_{ij} I_i(z) I_j(\bar{z}). \tag{11.5.57}$$

The coefficients $\mathcal{X}_{ij}$ can be determined by imposing the monodromy invariance under the monodromy group identified by the functions fi$I_i(z)$ and the normalization of the two-point functions. After all these steps, we can obtain the complete determination of all structure constants of the conformal theory. As a further example of their expressions, we report here the value

$$C_{(13,nm)}^{(n,m+2)} = \frac{2\rho-1}{(m+1)\rho-n} \left[\frac{\gamma(2-3\rho)\gamma(n-m\rho)}{\gamma(1-\rho)\gamma(1+n-(m+2)\rho)}\right]^{1/2}, \tag{11.5.58}$$

with the notation previously introduced.

11.6 Landau–Ginzburg Formulation

The aim of this section is to show that the minimal unitary models $\mathcal{M}_q$ describe the dynamics of the multicritical points of a quantum field theory of a massless scalar field φ with polynomial interaction of highest power $\varphi^{2(q-1)}$. This is the so-called Landau–Ginzburg theory, where the euclidean action is given by

$$\mathcal{S} = \int d^2x \left[\frac{1}{2}(\partial_\mu \varphi)^2 + V(\varphi)\right]$$

and the general form of the potential is expressed by the normal order powers of the field φ

$$V(\varphi) = g_1\varphi + g_2 : \varphi^2 : + \cdots + g_{2(q-2)} : \varphi^{2(q-2)} : + g : \varphi^{2(q-1)} : . \tag{11.6.1}$$

Note the absence of the term $\varphi^{2(q-1)-1}$: this term can always be removed by a shift of the field $\varphi \to \varphi + const$ and absorbed in the linear term $g_1\varphi$. Below we assume that the higher coupling constant g is fixed to a positive value. By varying the different parameters g_i, the shape of $V(\varphi)$ can greatly vary. Nevertheless, there always exists a set of values of the coupling constants for which the potential presents $(q-1)$ degenerate ground states: for instance in the φ^4 theory, this situation is realized by a family of curves associated to the parameter a, and the potential, shown in Fig. 11.5, given by

$$V(\varphi) = g(\varphi^2 - a^2)^2. \tag{11.6.2}$$

Analogously, for the ϕ^6 theory, there is a family of potentials that presents three degenerate vacua by varying the parameter b, as shown in Fig. 11.6

$$V(\varphi) = g\varphi^2\,(\varphi^2 - b^2)^2. \tag{11.6.3}$$

In the general case, these vacua correspond to the $(q-1)$ different phases of the model. In the space of the coupling constants, the point at which all the coupling constants but g vanish is then a multicritical point: at this point, characterized by the vanishing of the first $(2q-3)$ derivatives, there is a coalescence of $(q-1)$ different phases.

The operator content of the Landau–Ginzburg theory with potential (11.6.1) consists of $2(q-2)$ scalar relevant fields, associated to the various powers $: \varphi^k :$

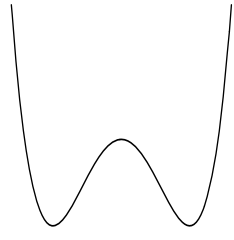


Fig. 11.5 *Potential with two degenerate vacua for the φ^4 theory.*

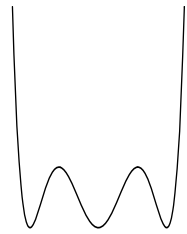


Fig. 11.6 *Potential in the φ^6 theory with three degenerate vacua.*

($k = 1, 2, \ldots, 2(q-2)$), and the irrelevant operators given by all their other derivative fields, as $:\varphi^k\,\partial_\mu\varphi\partial_\mu\varphi:$. At the multicritical point, the equation of motion in complex coordinates is given by

$$\partial_z\,\partial_{\bar{z}}\varphi \;\sim :\,\varphi^{2q-3}\,: . \tag{11.6.4}$$

Such an equation has to be understood as an operator identity once inserted in the correlation functions. Namely, each time that in a correlation function the field $:\varphi^{2q-3}:$ appears, it can be replaced by $\partial_z\,\partial_{\bar{z}}\varphi$:

$$\langle \cdots :\varphi^{2q-3}: \cdots \rangle \;=\; \langle \cdots \partial_z\,\partial_{\bar{z}}\varphi \cdots \rangle.$$

Let's now show how these features of the Landau–Ginzburg theory are implemented by the minimal unitary models $\mathcal{M}_q$.

Counting of the operators. To start with, notice that these models have a number of relevant fields precisely equal to $2(q-2)$: they correspond to the scalar conformal fields $\phi_{\Delta,\Delta}(z,\bar{z})$ with conformal weight $\Delta < 1$. As already noticed, in these models the formula of the conformal weights

$$\Delta_{r,s} = \frac{((q+1)r - qs)^2 - 1}{4q(q+1)} \tag{11.6.5}$$

is proportional to the distance of the points of the Kac lattice to the straight line of slope $q/(q+1)$ and therefore a such a counting problem simply reduces to determining how many points of the lattice fall inside the strip identified by the condition $\Delta_{r,s} < 1$, as shown in Fig. 11.7.

Identification of the operators. The first thing to do is to identify the most relevant field of the conformal model $\mathcal{M}_p$ with the scalar field φ that enters the Landau–Ginzburg lagrangian. It is easy to check that such a conformal field is the one placed at the position $(2,2)$ of the Kac table, with conformal weight

$$\Delta \;=\; \Delta_{2,2} \;=\; \frac{3}{4q(q+1)}. \tag{11.6.6}$$

With the position $\varphi \equiv \phi_{2,2}$, let's now proceed toward the recursive definition of the normal product of the higher powers of φ by means of the OPE provided by the conformal theory. Remember that, in conformal field theory, to define the composite

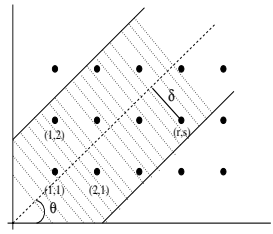


Fig. 11.7 *The relevant operators correspond to the lattice points that satisfy* $\Delta_{r,s} < 1$*, i.e. those inside the strip shown in the figure.*

operator $:A^2:(x)$ of an operator $A(x)$ (with anomalous dimension η_1), one has to consider the OPE of $A(x)$ with itself

$$A(x)A(0) - \langle A(x)A(0)\rangle \;=\; |x|^{\eta_2 - 2\eta_1}\, A_2(0) + \cdots \tag{11.6.7}$$

Once the most singular terms of this expression have been subtracted, the identification of the operator $:A^2:$ is made through the limit

$$:A^2(0):\equiv\; A_2 \;=\; \lim_{x\to 0} |x|^{2\eta_1 - \eta_2}\; \left(A(x)A(0) - \langle A(x)A(0)\rangle\right). \tag{11.6.8}$$

The higher-power composite operators $:A^{k+1}:$ are defined in an analogous way, and they coincide with those conformal fields selected by the limit

$$:A^{k+1}:(0) \;=\; \lim_{x\to 0} |x|^{\eta_1+\eta_k-\eta_{k+1}} \left[A(x)\;:A^k(0): - \sum_{l=1}^{k/2} C_l\, |x|^{\eta_{k-2l}-\eta_1-\eta_k}\;:A^{k-2l}:(0)\right]. \tag{11.6.9}$$

Note that the most singular terms of this expansion come from the previous operators $:A^{k-2l}:$, with $l = 1, 2, \ldots, < k/2$, and therefore they must be subtracted. In this expression η_p are the anomalous dimensions of the composite fields $:A^p:$ whereas the coefficients C_p are the relative structure constants that ensure the existence of the limit.

Given the above definition of the normal order and the identification of φ with $\phi_{2,2}$, it is easy to see that the composite operator $:\varphi^2:$ ends up to be the conformal field $\phi_{3,3}$. In fact, the fusion rules of $\phi_{2,2}$ are

$$\phi_{2,2}\times\phi_{2,2} \;=\; [\mathbf{1}] + [\phi_{3,3}] + [\phi_{1,3}] + [\phi_{3,1}] \tag{11.6.10}$$

and, once we have subtracted the contibution coming from the identity family, the most singular term in the expansion and the one that survives in the limit (11.6.8) is given by the conformal field $\phi_{3,3}$. One can then proceed to identitfy the higher powers of φ, with the final result given by (see Fig. 11.8)

$$:\varphi^k: \;=\; \begin{cases} \phi_{k+1,k+1} & k = 0, 1, \ldots q-2 \\ \phi_{k-q+3,k-q+2} & k = q-1, q, q+1, 2q-4. \end{cases} \tag{11.6.11}$$

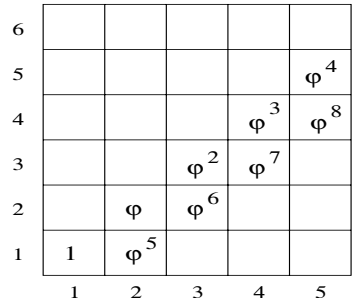


Fig. 11.8 *The correspondence between the conformal fields and the composite operators* φ^k *for the minimal model* $\mathcal{M}_6$.

The anomalous dimension $\eta_k = 2\Delta_k$ of these composite operators is obtained using eqn (11.6.5)

$$\eta_k = \begin{cases} \frac{(k+1)^2-1}{2q(q+1)} & k = 0, 1, 2, \ldots, q-2 \\ \frac{(k+3)^2-1}{2q(q+1)} & k = q-1, q, \ldots, 2q-4. \end{cases} \tag{11.6.12}$$

The key point in the route to establish the identification of the conformal model $\mathcal{M}_q$ with the $\varphi^{2(q-2)}$ Landau–Ginzburg theory is met when we consider the operator expansion $\varphi : \varphi^{2(q-2)} := \phi_{2,2}\, \phi_{q-2,q-1}$. The most singular contribution in this product comes from the conformal fields $\phi_{q-3,q-2} \equiv: \varphi^{2q-5}$ and $\phi_{q-1,q_2} = \phi_{2,2}$, and both must be subtracted in the proper definition of the composite operator φ^{2q-3}. After these subtractions, the first term that remains is the first descendant of the conformal field $\phi_{2,2}$, that is nothing else but $\partial_z\, \partial_{\bar{z}} \varphi$. In this way we arrive at the operator identity

$$: \varphi^{2q-3} := \partial_z\, \partial_{\bar{z}}\, \varphi \tag{11.6.13}$$

which coincides with the equation of motion (11.6.4) of the Landau–Ginzburg theory. Finally, the other conformal fields that enter the Kac table of the minimal model $\mathcal{M}_p$ can be identified with the irrelevant composite fields $: \varphi^k\, \partial_\mu \varphi \partial_\mu \varphi$.

11.7 Modular Invariance

In the two-dimensional conformal theories there is a natural splitting of the analytic and anti-analytic parts of the fields. This algebraic separation is stated by the Ward identity and is quite useful in many contexts (like finding the irreducible representations, the linear differential equations, etc.). However, to recover the real physical situation we have to combine together the analytic and the anti-analytic parts.

In the previous section we have seen that, in an infinite plane, a way to establish the correct combination of the two sectors is given by the condition of the monodromy invariance of the correlation functions. There is, however, another approach to finding the physical content of the conformal theories. Proposed originally by J.L. Cardy, this approach consists of studying the properties of a conformal field theory defined on a torus, i.e. a cylinder with periodic boundary conditions along both directions. From the geometrical symmetry of this problem, there are quite severe constraints

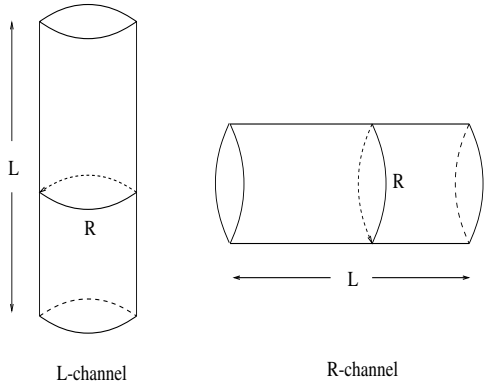


Fig. 11.9 *A torus geometry, i.e. a cylinder with periodic boundary conditions along both directions, and the relative quantization channels.*

on the operator content of the theories. Before discussing in detail the mathematical formalism of this approach, it is useful to present its main idea by means of a simple example. Consider a rectangle of sides L along the vertical direction and R along the horizontal one. For computing the partition function of a conformal field theory defined on such a geometry there are two possible ways (see Fig. 11.9):

- In the first approach, one considers the vertical axis as the time direction, with the time propagation ruled by a hamiltonian H_R. This quantization scheme defines the so-called *L-channel* of the theory. In this case the partition function can be expressed as
$$Z_1(L,R) \;=\; \mathrm{Tr}\, e^{-LH_R}. \tag{11.7.1}$$
- In the second approach, one considers instead the horizontal axis as the time direction, with a time evolution implemented by a hamiltonian H_L. This quantization scheme defines the so-called *R-channel* of the theory and, correspondingly, the partition function is given by
$$Z_2(R,L) \;=\; \mathrm{Tr}\, e^{-RH_L}. \tag{11.7.2}$$

The two ways of computing the partition function are obviously equivalent and this leads to the identity
$$Z_1(L,R) \;=\; Z_2(R,L), \tag{11.7.3}$$
which expresses the *modular invariance* of the theory. As we are going to see in the next section, by enforcing the validity of eqn (11.7.3) we can characterize the operator content of the conformal theories, given by an appropriate combination of their analytic and anti-analytic sectors.

11.7.1 Torus Geometry

Let's now refine the previous considerations by studying the mathematical properties that are relevant for modular invariance. A torus is defined by specifying two independent vectors in the plane and identifying the points that differ by a linear combination

of them with integer coefficients. In this way, the plane takes the periodic structure of a lattice. In the complex plane, such vectors can be specified by two complex numbers ω_1 and ω_2, the so-called *periods* of the lattice, and a torus is therefore defined by the equivalence relation

$$z \equiv z + n\omega_1 + m\omega_2 \qquad n, m \in N. \tag{11.7.4}$$

Such a tiling of the plane is not unique. Let $\omega'_{1,2}$ be in fact two new periods. Since they give to the same lattice structure, they can be expressed as a linear combination with integer coefficients of the original periods ω_1 and ω_2:

$$\begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \qquad \begin{matrix} a, b, c, d \in Z \\ ad - bc = 1. \end{matrix} \tag{11.7.5}$$

The determinant of such a transformation should not vanish since, by the symmetrical role of the two sets of periods, the linear combination should be invertible. Moreover its normalization is fixed by the condition that the area of the elementary cell of the lattice is the same, if expressed on the basis given by $\omega'_{1,2}$ or $\omega_{1,2}$.

In the complex plane, two lattices correspond to two conformally equivalent torus geometries if they differ just by a rotation and a dilatation. We can use this freedom to reduce the pair of periods (ω_1, ω_2) to the values $(1, \tau)$, where the complex number $\tau = \omega_2/\omega_1$, with $\mathrm{Im}\,\tau > 0$ is the *modular parameter*. Without losing generality, we can choose as vertices of the torus the points $\{0, 1, \tau, (1+\tau)\}$. The physical request is that the conformal theories defined on such a geometry should not depend either on the scale or on the orientation of the lattice, i.e. the condition that the theory presents a *modular invariance*. Since under the change (11.7.5) the modular parameter transforms according to the Moebius map

$$\tau \to \frac{a\tau + b}{c\tau + d} \qquad ad - bc = 1 \tag{11.7.6}$$

the corresponding symmetry group coincides with the 2×2 linear transformations with integer coefficients and determinant equal to 1. Furthermore, since all parameters a, b, c, d can be changed by sign without affecting the final transformation, the modular group Γ is given by $SL(2, Z)/Z_2$ and consists of the group of discontinuous diffeomorphisms of the torus, i.e. the set of all those transformations of the torus that cannot be obtained adiabatically starting from the identity transformation. Such a discrete group can be generated by the repeated action of the operators (see Problem 5)

$$\begin{aligned} &\mathcal{T} : \tau \to \tau + 1 \ \text{ alias } \ \mathcal{T} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \\ &\mathcal{S} : \tau \to -1/\tau \ \text{ alias } \ \mathcal{S} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned} \tag{11.7.7}$$

whose graphical representation is shown in Fig. 11.10. These transformations satisfy

$$\mathcal{S}^2 = (\mathcal{S}\mathcal{T})^3 = 1. \tag{11.7.8}$$

The *fundamental domain* of the modular group is defined as that region of the upper half complex plane for which any pair of its points cannot be related by a modular

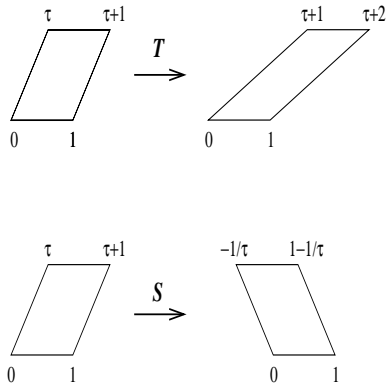


Fig. 11.10 *Transformation of the lattice under the action of the generators $\mathcal{T}$ and $\mathcal{S}$.*

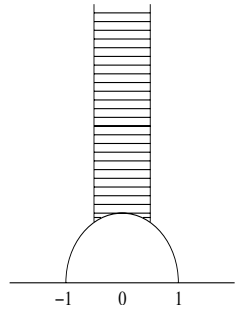


Fig. 11.11 *Fundamental domain of the modular group Γ (dashed area).*

transformation, whereas any other external point can be reached from one of its interior points by a modular transformation. The usual choice of the fundamental domain is the following: $-\frac{1}{2} < \mathrm{Re}\,\tau < \frac{1}{2}, |\,\tau\,| \geq 1$ (see Fig. 11.11).

11.7.2 Partition Function and Characters

In order to define the partition function on a torus, it is necessary to specify the time and space directions of the lattice. Let's initially choose as space direction that along the real axis of the complex plane, while the time direction is the imaginary axis. This choice introduces the L-channel, according to the terminology above. The translations along these axes are implemented by the momentum operator P and by the hamiltonian H, respectively. The partition function is then

$$Z(\tau,\bar{\tau}) \;=\; \mathrm{Tr}\,\exp\{-H\mathrm{Im}\,\tau - i\,P\,\mathrm{Re}\,\tau\}. \tag{11.7.9}$$

For H and P we can use the expression previously derived for a cylinder of width R (in the present case $R = 1$), namely

$$H \;=\; \frac{2\pi}{R}\left(L_0 + \bar{L}_0 - \frac{c}{12}\right), \quad P \;=\; \frac{2\pi}{R}(L_0 - \bar{L}_0) \tag{11.7.10}$$

where L_0 and $\bar{L}_0$ are the generators of the Virasoro algebra. Substituting these expressions in (11.7.9) and collecting the various terms we get

$$Z(\tau,\bar{\tau}) \;=\; \mathrm{Tr}\,\exp\left\{2\pi i\left[\tau(L_0-\frac{c}{24})\right]-\left[\bar{\tau}(\bar{L}_0-\frac{c}{24})\right]\right\}. \tag{11.7.11}$$

Defining the parameters

$$q \;\equiv\; e^{2\pi i\tau}, \qquad \bar{q} \;\equiv\; e^{-2\pi i\bar{\tau}} \tag{11.7.12}$$

the partition functions can be expressed as

$$Z(q,\bar{q}) \;=\; \mathrm{Tr}\,\left(q^{L_0-c/24}\,\bar{q}^{\bar{L}_0-c/24}\right). \tag{11.7.13}$$

The eigenstates of $(L_0,\bar{L}_0)$ are organized in terms of the irreducible representations given by the Verma modules of the direct sum of the two Virasoro algebras. Hence we can decompose the trace on these states into the sum of these representations and write then

$$Z(q,\bar{q}) \;=\; \sum_{\Delta,\bar{\Delta}} N_{\Delta,\bar{\Delta}}\,\chi_{\Delta}(q)\,\chi_{\bar{\Delta}}(\bar{q}), \tag{11.7.14}$$

where the non-negative integers $N_{\Delta,\bar{\Delta}}$ represent the number of times the representations associated to the conformal weights $(\Delta,\bar{\Delta})$ enter the trace, whereas $\chi_{\Delta}(q)$ are the *characters* of the Virasoro algebra, defined by

$$\chi_{\Delta}(q) \;\equiv\; q^{-c/24}\mathrm{Tr}\;q^{L_0}|_{\Delta} \;=\; q^{-(c/24)+\Delta}\;\sum_{n=0}^{\infty} d_{\Delta}(n)q^n\;. \tag{11.7.15}$$

The coefficients $d_{\Delta}(n)$ are the weights of the vector spaces at the level n in the representation identified by the conformal weight Δ.

The problem is now to determine the set of integers $N_{\Delta,\bar{\Delta}}$ that ensures the modular invariance of the partition function. This means that Z must be a function invariant both under $\mathcal{T}:\tau\to\tau+1$ and $\mathcal{S}:\tau\to-1/\tau$:

$$\begin{aligned} Z(\tau) \;&=\; Z(\tau+1),\\ Z(\tau) \;&=\; Z(-1/\tau). \end{aligned} \tag{11.7.16}$$

For the minimal conformal models, the explicit expression of the characters of the degenerate fields $\varphi_{r,s}$ is provided by the Rocha–Caridi formula

$$\chi_{r,s}(q,c)=\eta^{-1}(q)\;q^{-\frac{(c-1)}{24}+\Delta_{r,s}}\;\sum_{k=-\infty}^{\infty} q^{pp'k^2}\left(q^{k(rp'-sp)}-q^{k(rp'+sp)}\right), \tag{11.7.17}$$

where $\eta(q)$ is the Dedekind function

$$\eta(q)=q^{1/24}\prod_{k=1}^{\infty}\left(1-q^k\right). \tag{11.7.18}$$

To find which integers $N_{\Delta,\bar{\Delta}}$ ensure the validity of eqns (11.7.16) it is necessary to analyze how the characters transform under the action of the two generators of the modular group. $\mathcal{T}$ acts on the characters in a particularly simple way:

$$\mathcal{T}: \quad \chi_\Delta(q) \to e^{2\pi i(\Delta - c/24)}\, \chi_\Delta(q). \tag{11.7.19}$$

The invariance of the partition function under this transformation implies that $N_{\Delta,\bar{\Delta}} = 0$, unless $\Delta - \bar{\Delta} = k$, where k is an integer.

Consider now the action of $\mathcal{S}$. Note that this transformation implements an exchange of the space and time axes of the theory. This implies that if we had computed the partition function swapping the role of the two directions, we would have obtained an expression similar to the previous (11.7.14)

$$Z(\tilde{q}, \tilde{\bar{q}}) \;=\; \sum_{\Delta,\bar{\Delta}} N_{\Delta,\bar{\Delta}}\, \chi_\Delta(\tilde{q})\, \chi_{\bar{\Delta}}(\tilde{\bar{q}}), \tag{11.7.20}$$

but with the fundamental difference given by the presence of the quantity $\tilde{q} = e^{-2\pi i/\tau}$ instead of the original variable q. The equality of this expression with the one in eqn (11.7.14) has two important consequences:

1. there should exist a linear transformation that link the characters expressed in terms of the variables q and $\tilde{q}$;
2. there should exist a stringent condition on the coefficients $N_{\Delta,\bar{\Delta}}$ that ensures the identity of the two expressions of the partition function.

Let's address the first point. Note that the expression for the characters of the degenerate fields in the minimal models is very similar to the infinite series that define the $\theta_i(z)$ functions, given in general by an infinite sum of exponentials, with a quadratic expression for k in the exponent. As for the $\theta_i(z)$ functions, the characters present remarkable properties under the transformation $\tau \to -1/\tau$, whose derivation requires the Poisson resummation formula. Here we only state the final result relative to the minimal models identified by the pair of coprime integer numbers p, p':

$$\chi_{r,s}(\tilde{q}) \;=\; \sum_{r',s'} S_{rs}^{r's'}\, \chi_{r's'}(q), \tag{11.7.21}$$

where

$$S_{rs}^{r's'} \;=\; \left(\frac{8}{pp'}\right)^{1/2} (-1)^{(r+s)(r'+s')}\, \sin\frac{\pi r r'}{p}\, \sin\frac{\pi s s'}{p'}. \tag{11.7.22}$$

This formula shows that the characters $\chi_{r,s}$ change according to a finite-dimensional representation of the modular group Γ. Note that the matrix elements[7] $S_{rs}^{r's'}$ are symmetric and real. Moreover, since the transformation $\mathcal{S}$ is unitary, we have $S^2 = 1$. Denoting by $\mathcal{R}$ this finite-dimensional representation, the combination of the characters in the partition function transforms as $M \equiv R \otimes R^*$. Therefore to find the

[7]The pair of indices (rs), as well as $(r's')$, has to be considered as a single index that identifies the corresponding conformal field.

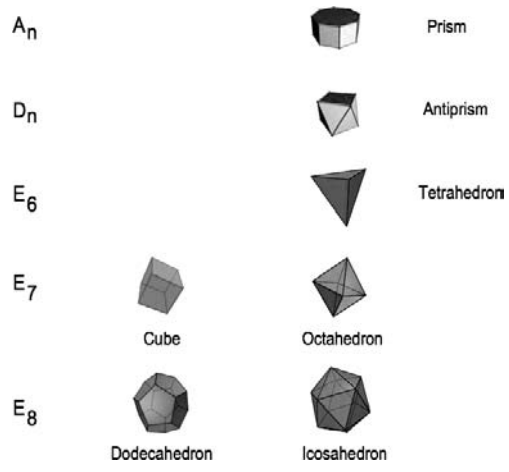


Fig. 11.12 *The ADE classification of regular polyhedra. The polyhedra are convex, with all equivalent vertices. A similar classification holds for convex polyhedra with equivalent faces and the two series are related by face-vertex duality.*

modular invariant expressions of the partition functions we have to determine the integer coefficients $N_{\Delta,\bar{\Delta}}$ that satisfy the condition (expressed in matrix notation)

$$M\,N \;=\; N. \tag{11.7.23}$$

In other words, we shall find the eigenvectors, with non-negative integer components and with eigenvalue equal to 1, of the matrix M. An additional condition is $N_{0,0} = 1$: this enforces the presence of the identity operator in the partition function with multiplicity equal to 1.

The general solution of this mathematical problem has been found by Cappelli, Itzykson, and Zuber. It has a remarkable structure: in fact, the modular invariant partition functions can be put in correspondence with the series ADE that classify the simply laced Lie algebra.[8] At first sight it may seem surprising to see conformal field theories being classified by the ADE Lie algebra but, on the other hand, Lie algebras arise whenever integrability and local symmetries are involved. The classical example is the classification of regular convex polyhedra shown in Fig. 11.12.

The explicit expressions of the partition functions are reported in Table 11.1. Without claiming full justification of these formulas (we refer the reader to the original literature for all details), it is however possible to understand the origin of some of them. For instance, a natural solution of eqn (11.7.23) is provided by the diagonal combination, i.e. by the integers $N_{\Delta,\bar{\Delta}} = \delta_{\Delta,\bar{\Delta}}$. The partition functions associated to this solution involve all the scalar primary fields of the Kac table of a given model. To present another class of solutions, consider the case of unitary minimal models, where $p' = p+1$. We assume that the indices r, s run over all possible values of the Kac table

[8]The discussion of Lie algebras can be found in the Appendix of Chapter 13.

Table 11.1: *Modular invariant partition functions of the conformal minimal models.*

p, p'	$\frac{1}{2}\sum_{r=1}^{p'-1}\sum_{s=1}^{p-1} \mid \chi_{rs} \mid^2$	$(A_{p'-1}, A_{p-1})$
$p' = 4\rho+2$ $p \geq 1$	$\frac{1}{2}\sum_{s=1}^{p-1}\Big\{\sum_{\substack{r \text{ odd}=1 \\ r\neq 2\rho+1}}^{4\rho+1} \mid \chi_{rs} \mid^2 + 2 \mid \chi_{2\rho+1,s} \mid^2 + \sum_{r \text{ odd}=1}^{2\rho-1} (\chi_{rs}\chi^*_{r,p-s} + c.c.)\Big\}$	$(D_{2\rho+2}, A_{p-1})$
$p' = 4\rho$ $p \geq 2$	$\frac{1}{2}\sum_{s=1}^{p-1}\Big\{\sum_{r \text{ odd}=1}^{4\rho-1} \mid \chi_{rs} \mid^2 + \mid \chi_{2\rho,s} \mid^2 + \sum_{r \text{ even}=2}^{2\rho-2} (\chi_{rs}\chi^*_{r,p'-s} + c.c.)\Big\}$	$(D_{2\rho+1}, A_{p-1})$
$p' = 12$	$\frac{1}{2}\sum_{s=1}^{p-1}\big\{\mid \chi_{1s}+\chi_{7s} \mid^2 + \mid \chi_{4s}+\chi_{8s} \mid^2 + \mid \chi_{5s}+\chi_{11s} \mid^2\big\}$	(E_6, A_{p-1})
$p' = 18$	$\frac{1}{2}\sum_{s=1}^{p-1}\big\{\mid \chi_{1s}+\chi_{17s} \mid^2 + \mid \chi_{5s}+\chi_{13s} \mid^2 + \mid \chi_{7s}+\chi_{11s} \mid^2 + \mid \chi_{9s} \mid^2 + [(\chi_{3s}+\chi_{15s})\chi^*_{9s} + cc]\big\}$	$(E_7, A_{p=1})$
$p' = 30$	$\frac{1}{2}\sum_{s=1}^{p-1}\big\{\mid \chi_{1s}+\chi_{11s}+\chi_{19s}+\chi_{29s} \mid^2 + \mid \chi_{7s}+\chi_{13s}+\chi_{17s}+\chi_{23s} \mid^2\big\}$	(E_8, A_{p-1})

$(1 \leq r \leq p;\ 1 \leq s \leq p+1)$ and therefore each primary field appears twice. It is easy to see that if p is an odd number, we have

$$S_{rs}r's' \ = \ (-1)^{s-1}S^{r',p'-s'}_{rs} \ = \ (-1)^{s'-1}S^{r's'}_{r,p'-s}. \tag{11.7.24}$$

This identity implies that the combination made by the characters

$$\chi_{rs} + \chi_{r,p'-s} \qquad (s \,\text{odd}) \tag{11.7.25}$$

defines an invariant subspace. Therefore the partition function given by

$$Z \ = \ \frac{1}{2}\sum_{r}\sum_{s \text{ odd}} \mid \chi_{rs} + \chi_{r,p'-s} \mid^2 \tag{11.7.26}$$

is invariant under the $\mathcal{S}$ transformation. It is also easy to check that $\Delta_{rs} - \Delta_{r,p+1-s}$ is always an integer if $p = 1 \pmod 4$, and in this case, the above partition function is also invariant under $\mathcal{T}$. Similar invariant expressions can be found for all values of $p \geq 5$.

In addition to these two infinite series of solutions, there are others that are relative to particular values of p', given by $p' = 12, 18, 30$. As we mentioned above, all the modular invariant solutions can be put in correspondence with the ADE algebras that appear in so many branches of mathematics, as in the classification of finite subgroups of the group of rotations or in the classification of the critical points in the theory of catastrophies. V. Pasquier has also shown that the modular invariant partition functions can be obtained as a continuum limit of certain discrete lattice statistical models defined in terms of Dynkin diagrams. The modular invariant partition functions of the conformal minimal models are reported in Table 11.1.

Appendix 11A. Hypergeometric Functions

Let a, b, and c be complex numbers. The hypergeometric differential equation

$$z(z-1)\frac{d^2w}{dz^2} + [(a+b+1)z - c]\frac{dw}{dz} + abw = 0, \tag{11.A.1}$$

has three singular regular points at $z = 0, 1, \infty$. When c is different from zero or it is a negative integer, an analytic solution of this equation in the vicinity of $z = 0$ is expressed by the series

$$F(z; a, b, c) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}, c \neq 0, -1, -2, \ldots \tag{11.A.2}$$

where

$$(a)_n \equiv a(a+1)\ldots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)}.$$

If a or b are equal to zero or are negative integers, the series truncates and the hypergeometric function becomes a simple polynomial.

Since the hypergeometric differential equation is of second order, it admits a second solution, usually written in the form

$$w_2(z) = z^{1-c} F(z; a-c+1, b-c+2, 2-c). \tag{11.A.3}$$

It is easy to see that if c is an integer, either the two solutions coincide or one of them diverges. In the second case, the second solution presents a logarithmic contribution.

References and Further Reading

The unitary properties of the minimal models have been studied in the article:

D. Friedan, Z. Qiu, S. Shenker, *Conformal invariance, unitarity and two-dimensional critical exponents*, Phys. Rev. Lett. 52 (1984), 1575.

The modified Coulomb gas method has been proposed by V. Dotsenko and V. Fateev in:

V.S. Dotsenko, V. Fateev, *Conformal algebra and multipoint correlation functions in 2D statistical models*, Nucl. Phys. B 240 (1984), 312; *Four point correlation functions and the operator algebra in the two-dimensional conformal invariant theories with the central charge* $c < 1$, Nucl. Phys. B 251 (1985), 691.

For the quantum group formulation of the minimal models see:

L. Alvarez-Gaume, C. Gomez, G. Sierra, *Quantum group interpretation of some conformal field theories*, Phys. Lett. B 220 (1989), 142.

L. Alvarez-Gaume, C. Gomez and G. Sierra, *Hidden Quantum Symmetries in Rational Conformal Field Theories*, Nucl. Phys. B 319 (1989), 155.

The principle of modular invariance has been stressed by Cardy in:

J.L. Cardy, *Operator content of two-dimensional conformally invariant theories*, Nucl. Phys. B 270 (1986), 186.

The Landau–Ginzburg picture of the unitary minimal models has been proposed by A. Zamolodchikov in:

A.B. Zamolodchikov, *Conformal symmetry and multicritical points in two-dimensional quantum field theory*, Sov. J. Nucl. Phys. 44 (1986), 529.

The full classification of modular invariant partition functions has been obtained in:

A. Cappelli, C. Itzykson, J.B. Zuber, *Modular invariant partition functions in two-dimensions*, Nucl. Phys. B 280 (1987), 445.

The identification of modular invariant partition functions with statistical modes based on Dynkin diagrams has been proposed by Pasquier in the paper:

V. Pasquier, *Two-dimensional critical systems labelled by Dynkin diagrams*, Nucl. Phys. B 285 (1987), 162.

A detailed analysis of the hypergeometric functions can be found in the book:

P.T. Bateman, *Complex Variables. Introduction and Applications*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1977.

For the classification of regular polyhedra and the correspondence with ADE Lie algebras see:

H.S.M. Coxeter, *Regular Polytopes*, Dover, New York, 1973.

P. Slodowy, *Simple singularities and simple algebraic groups*, in Lecture Notes in Mathematics 815, Springer, Berlin, 1980.

The Verlinde algebra is deeply related to the modular group. It has found application in topological quantum computation. The interested reader may consult:

E. Verlinde, *Fusion rules and modular transformations in 2D conformal field theory*, Nucl. Phys. B 300 (1988), 360.

J.L. Cardy, *Boundary conditions, fusion rules and the Verlinde algebra*, Nucl. Phys. B 324 (1989), 324.

S. Das Sarma, M. Freedman, C. Nayak, S.H. Simon, A. Stern, *Non-abelian anyons and topological quantum computation*, Rev. Mod. Phys. 80 (2008), 1083.

Problems

1. Null-vector at the third level

1. Show that the linear combination that gives rise to null-vectors at level $N = 3$ is given by
$$\left[L_{-3} - \frac{2}{\Delta+1} L_{-1}L_{-2} + \frac{1}{(\Delta+1)(\Delta+2)} L_{-1}^3\right] \phi_\Delta \;=\; 0$$
with $\Delta = \Delta_{1,3}$ or $\Delta = \Delta_{3,1}$.
2. Determine the differential equation satisfied by the correlators of the primary fields $\phi_{1,3}$ and $\phi_{3,1}$.
3. Show that the fusion rules of the fields $\phi_{1,3}$ and $\phi_{3,1}$ given in eqns (11.4.22) and (11.4.23) are compatible with the differential equation satisfied by their correlation functions.

2. Structure constant

For the minimal unitary models, identified by the integer p, compute the limiting value of $C^{(13)}_{(13)(13)}$ for $p \to \infty$.

3. Fusion rules

Consider the minimal models $\mathcal{M}_{2,2n+1}$ $(n = 1, 2, \ldots)$. Compute the central charge and the effective central charge by identifying the operator with the lowest conformal weight. Determine the fusion rules of these models.

4. Non-unitarity model $\mathcal{M}_{3,5}$

The non-unitarity model $\mathcal{M}_{3,5}$ has the following Kac table.

$\frac{3}{4}$	0
$\frac{1}{5}$	$-\frac{1}{20}$
$-\frac{1}{20}$	$\frac{1}{5}$
0	$\frac{3}{4}$

With the identification of the fields

$$1 = \Phi_{0,0}, \qquad \sigma = \Phi_{-\frac{1}{20},-\frac{1}{20}};$$
$$\varphi = \Phi_{\frac{1}{5},\frac{1}{5}}, \qquad \psi = \Phi_{\frac{3}{4},\frac{3}{4}}$$

prove that the fusion rules of this model are given by

$$\begin{aligned} &\psi \times \psi = 1, && \psi \times \sigma = \varphi; \\ &\sigma \times \sigma = 1 + \varphi, && \psi \times \varphi = \sigma; \\ &\varphi \times \varphi = 1 + \varphi, && \varphi \times \sigma = \sigma + \psi. \end{aligned}$$

Compute the four-point correlation functions involving ψ and σ and determine the exact expressions of the structure constants of the conformal algebra.

5. Modular group

Prove that any matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with integer coefficients and satisfying $ad - bc = 1$, can be obtained by multiplication of suitable powers of the elementary matrices

$$\mathcal{T} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad \mathcal{S} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

6. Quantum dimensions

Denote the number of linearly independent states having $\mathcal{N}$ fields of type a as $\mathcal{H}_a(\mathcal{N})$. The quantum dimension d_a of the excitations of type a is given by studying the behavior $\mathcal{H}_a(\mathcal{N})$ for large $\mathcal{N}$, which behaves as $\mathcal{H}_a(\mathcal{N}) \simeq d_a^{\mathcal{N}}$. To compute d_a, let's fuse M ϕ_a fields recursively using the Verlinde algebra

$$\phi_a \cdot \phi_a \cdot \ldots \cdot \phi_a = \sum_{\{c_k\}} N_{aa}^{c_1} N_{ac_1}^{c_2} \cdots N_{ac_{M-2}}^{c_{M-1}} \phi_{c_{M-1}}.$$

Observe that this is the product of $(M-1)$ copies of the matrix $(N_a)_b^c = N_{ab}^c$, so that in the limit $M \to \infty$, the product will be dominated by the largest eigenvalue of N_a.

1. If S_a^b is the unitary matrix that simultaneously diagonalizes all the matrices N_a of the Verlinde algebra, show that the fusion coefficients are expressed by

$$N_{ab}^c = \sum_j \frac{S_a^j S_b^j S_j^c}{S_0^j},$$

   where 0 denotes the identity field and the sum is over all fields entering the algebra.
2. Show that the largest eigenvalue (and therefore the quantum dimension d_a) is given by

$$d_a = \frac{S_a^0}{S_0^0}.$$

3. Consider the algebra

$$\mathbf{1} \cdot \mathbf{1}, \qquad \mathbf{1} \cdot \phi = \phi, \qquad \phi \cdot \phi = \mathbf{1} + \phi.$$

   Compute the quantum dimension d_ϕ and show that it is equal to the golden ratio

$$d_\phi = \frac{\sqrt{5}+1}{2}.$$

12
Conformal Field Theory of Free Bosonic and Fermionic Fields

Science is spectral analysis. Art is light synthesis.

Karl Kraus

12.1 Introduction

In this chapter we discuss two explicit examples of conformal field theories. We start our analysis with the free massless bosonic theory that we have already seen in the previous chapter. After, we discuss the conformal field theory of a complex fermion operator (a Dirac fermion) using its decomposition in the real Majorana components. The central charge of both bosonic and fermion theories is $c = 1$ and this suggests the existence of an equivalence between them. The transformation that maps a bosonic into a fermion theory and vice versa is known as *bosonization*: it provides a useful tool both for the comprehension of the conformal theories and for a wide range of applications, in particular in low-dimensional condensed matter systems.

12.2 Conformal Field Theory of a Free Bosonic Field

This section is devoted to the detailed analysis of the conformal field theory of a massless bosonic field that was employed in Chapter 11 for the Coulomb gas approach. Despite the simple form of the lagrangian of this model, it presents a rich operator content and a remarkable duality property of its partition function on a torus. Later we will also established the equivalence of this theory with the theory of massless Dirac fermions.

12.2.1 Quantization of the Bosonic Field

Let $\varphi(x,t)$ be a free bosonic and real field, with action

$$\mathcal{S} = \frac{g}{2}\int d^2x\, \partial_\mu\varphi\partial^\mu\varphi. \tag{12.2.1}$$

Let's assume that the model is defined on a cylinder of width L with periodic boundary conditions $\varphi(z+L,t) = \varphi(x,t)$. Expanding the field in its Fourier modes

$$\varphi(x,t) = \sum_n e^{2\pi n x/L}\,\varphi_n(t)$$
$$\varphi_n = \frac{1}{L}\int_0^L dx\, e^{-2\pi i n x/L}\,\varphi(x,t)$$

and substituting this expression into the action, we have

$$\mathcal{S} = \frac{1}{2} gL \sum_n \left[\dot\varphi_n\,\dot\varphi_{-n} - \left(\frac{2\pi n}{L}\right)^2 \varphi_n\,\varphi_{-n}\right]. \tag{12.2.2}$$

Let $\Pi(x,t) = g\partial_t\,\varphi(x,t)$ be the conjugate momentum of the field, with commutation relations (at a given time t)

$$[\varphi(x,t),\Pi(y,t)] = i\,\delta(x-y) \qquad [\varphi(x,t),\varphi(y,t)] = 0 \qquad [\Pi(x,t),\Pi(y,t)] = 0. \tag{12.2.3}$$

Expanding also $\Pi(x,t)$ in Fourier series and denoting by π_n the conjugate momenta of φ_n, we have

$$\pi_n = g\,L\,\dot\varphi_{-n}, \qquad [\varphi_n,\pi_m] = i\,\delta_{nm} \tag{12.2.4}$$

with $\varphi_n^\dagger = \varphi_{-n}$ and $\pi_n^\dagger = \pi_{-n}$. We can define the hamiltonian of the system by using the Legendre transformation

$$H = \frac{1}{2gL}\sum_n [\pi_n\,\pi_{-n} + (2\pi n g)^2\,\varphi_n\,\varphi_{-n}]. \tag{12.2.5}$$

This is the hamiltonian of a set of decoupled harmonic oscillators of frequencies $\omega_n = 2\pi|n|/L$. Note that the oscillator with $n=0$ has zero frequency. To take into account this feature, it is convenient to explicitly separate the zero mode φ_0 of the field and introduce, for the other modes, the operators a_n and $\bar a_n$ through the formulas

$$\varphi_n = \frac{i}{n\sqrt{4\pi g}}\,(a_n - \bar a_{-n}), \tag{12.2.6}$$
$$\pi_n = \sqrt{\pi g}\,n\,(a_{-n} + \bar a_n).$$

These operators satisfy the commutation relations

$$[a_n,a_m] = n\delta_{n+m,0} \qquad [a_n,\bar a_m] = 0 \qquad [\bar a_n,\bar a_m] = n\delta_{n+m,0}. \tag{12.2.7}$$

For the zero mode we have

$$[\varphi_0,\pi_0] = i. \tag{12.2.8}$$

Substituting these new operators into eqn (12.2.5) we obtain

$$H = \frac{1}{2\pi g L}\pi_0^2 + \frac{2\pi}{L}\sum_{n>0}(a_{-n}a_n + \bar a_{-n}\bar a_n), \tag{12.2.9}$$

and therefore

$$[H, a_{-n}] \,=\, \frac{2\pi}{L} n\, a_{-n} \quad , \quad [H, \bar{a}_{-n}] \,=\, \frac{2\pi}{L} n\, \bar{a}_{-n}, \quad [H, \varphi_0] \,=\, 0. \tag{12.2.10}$$

These expressions show that a_{-n} and $\bar{a}_{-n}$ (with $n > 0$) act as raising operators of the theory: applying one of these operators to an energy eigenstate with eigenvalue E, we obtain another energy eigenstate with eigenvalue $E + 2\pi n/L$. The operators a_n and $\bar{a}_n$ (with $n > 0$) act instead as annihilation operators of the theory: their application to an eigenstate with eigenvalue E gives rise to another energy eigenstate with eigenvalue $E - 2\pi n/L$. Equation (12.2.10) helps us to easily obtain the time evolution of the operators in the Heisenberg representation

$$\begin{array}{l} a_n(t) \,=\, a_n(0)\, e^{-2\pi i n t/L} \\ \bar{a}_n(t) \,=\, \bar{a}_n(0)\, e^{-2\pi i n t/L} \end{array} \qquad \varphi_0(t) \,=\, \varphi_0 + \frac{1}{gL}\pi_0 t. \tag{12.2.11}$$

Hence, the solution of the equation of motion of the field $\varphi(x,t)$ reads

$$\varphi(x,t) \,=\, \varphi_0 + \frac{1}{gL}\pi_0 t + \frac{i}{\sqrt{4\pi g}} \sum_{n\neq 0} \frac{1}{n} \left(a_n\, e^{2\pi i n (x-t)/L} - \bar{a}_{-n}\, e^{2\pi i n (x+t)/L} \right), \tag{12.2.12}$$

where all the operators that appear in this formula are those relative to the time $t = 0$. Equivalently, adopting a euclidean formulation, with $t = -i\tau$ and introducing the coordinates

$$z \,=\, e^{2\pi(\tau - ix)/L}, \qquad \bar{z} \,=\, e^{2\pi(\tau + ix)/L},$$

we have

$$\varphi(z,\bar{z}) \,=\, \varphi_0 - \frac{i}{4\pi g}\pi_0\, \ln(z\bar{z}) + \frac{i}{\sqrt{4\pi g}} \sum_{n\neq 0} \frac{1}{n} \left(a_n\, z^{-n} + \bar{a}_{-n}\, \bar{z}^{-n} \right). \tag{12.2.13}$$

This expression explicitly shows the decoupling of the analytic and anti-analytic components of the field, both due to the the equation of motion $\bar{\partial}\,\partial\varphi = 0$ and the periodic boundary conditions chosen for the field φ

$$\varphi(z,\bar{z}) \,=\, \phi(z) + \bar{\phi}(\bar{z}), \tag{12.2.14}$$

where

$$\begin{aligned} \phi(z) \,&=\, \frac{1}{2}\varphi_0 - \frac{i}{4\pi g}\pi_0\, \ln z + \frac{i}{\sqrt{4\pi g}} \sum_{n\neq 0} \frac{1}{n} a_n\, z^{-n} \\ \bar{\phi}(\bar{z}) \,&=\, \frac{1}{2}\varphi_0 - \frac{i}{4\pi g}\pi_0\, \ln \bar{z} + \frac{i}{\sqrt{4\pi g}} \sum_{n\neq 0} \frac{1}{n} \bar{a}_n\, \bar{z}^{-n}. \end{aligned} \tag{12.2.15}$$

According to the negative or positive value of the index n, the operators a_n create or annihilate the analytic excitation of the field φ, with a similar situation for $\bar{a}_n$ with

respect to the anti-analytic excitations.[1] It is also convenient to define

$$\theta(z,\bar{z}) \;=\; \phi(z) - \bar{\phi}(\bar{z}), \tag{12.2.16}$$

the so-called *dual field* of φ. It satisfies $\partial_\mu \varphi \;=\; -i\,\epsilon_{\mu\nu}\partial_\nu\,\theta$, and in complex coordinates

$$\begin{aligned} \partial_z \varphi \;&=\; \partial_z \theta, \\ \partial_{\bar{z}} \varphi \;&=\; -\partial_{\bar{z}} \theta. \end{aligned} \tag{12.2.17}$$

It is easy to verify that in the original Minkowski space, it satisfies the commutation relation (at fixed time t)

$$[\varphi(x,t), \theta(y,t)] \;=\; -i\,\epsilon(x-y), \tag{12.2.18}$$

where $\epsilon(v)$ is the step function

$$\epsilon(v) \;=\; \begin{cases} 1, & v > 0 \\ 0, & v < 0. \end{cases}$$

Equation (12.2.18) clearly shows the non-local relationship between φ and θ.

We have already noticed that φ is not a scaling field, whereas scaling fields are the two currents $J(z) \;=\; i\partial\varphi$ and $\bar{J}(\bar{z}) \;=\; -i\bar{\partial}\varphi$ that both generate a $U(1)$ symmetry. The expansion of these fields is given by

$$i\,\partial\varphi \;=\; i\,\partial\phi \;=\; \frac{1}{\sqrt{4\pi g}} \sum_{n\neq 0} a_n\, z^{-n} \tag{12.2.19}$$

(with a similar formula for $i\,\bar{\partial}\varphi$), where we have introduced the notation

$$a_0 \;=\; \bar{a}_0 \equiv \frac{\pi_0}{\sqrt{4\pi g}}.$$

We can now use the previous expression to define the analytic part of the stress–energy tensor

$$T(z) \;=\; -2\pi g \;:\partial\phi(z)\,\partial\phi(z):\; =\; \frac{1}{2}\sum_{n,m} z^{-n-m-2} \;:a_n\, a_m: \tag{12.2.20}$$

and extract the modes L_n of the Virasoro algebra of this theory

$$\begin{aligned} L_n &= \frac{1}{2}\sum_{m=-\infty}^{\infty} a_{n-m}\, a_m \quad (n \neq 0) \\ L_0 &= \frac{1}{2}\alpha_0^2 + \sum_{m=1}^{\infty} a_{-m}\, a_m \end{aligned} \tag{12.2.21}$$

[1] It is interesting to observe that the formulas given in the text appear also in string theory, in particular they enter the quantization of the closed string, with the zero mode φ_0 associated to the center of mass of the string and π_0 to its total momentum.

The hamiltonian (12.2.9) can be written as

$$H \;=\; \frac{2\pi}{L}(L_0 + \bar{L}_0) - \frac{\pi}{6L}. \tag{12.2.22}$$

The term $-\pi/6L$ is obtained by the normal order of a_n by using the regularization given by the Riemann function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ for the divergent series

$$\sum_{n>0} n = \zeta(-1) = -1/12.$$

Comparing this formula with the general expression previously derived for H on a cylinder geometry, eqn (10.9.3), we see that the central charge of the free bosonic theory is $c = 1$. Starting from the scalar field φ, in addition to $\partial\varphi$, we can construct an infinite series of scaling operators, given by the vertex operators

$$V_{\alpha,\bar{\alpha}}(z,\bar{z}) \;=:\; e^{i\alpha\phi(z)+i\bar{\alpha}\bar{\phi}(\bar{z})} : \tag{12.2.23}$$

whose conformal weights are

$$\Delta_\alpha \;=\; \frac{\alpha^2}{8\pi g}, \qquad \bar{\Delta}_\alpha \;=\; \frac{\bar{\alpha}^2}{8\pi g}. \tag{12.2.24}$$

They satisfy the operator product expansion

$$V_{\alpha,\bar{\alpha}}(z,\bar{z})\, V_{\beta,\bar{\beta}}(w,\bar{w}) \;=\; (z-w)^{\alpha\beta/4\pi g}\,(\bar{z}-\bar{w})^{\bar{\alpha}\,\bar{\beta}/4\pi g}\, V_{\alpha+\beta,\bar{\alpha}+\bar{\beta}}(w,\bar{w}) + \cdots \tag{12.2.25}$$

An interesting interpretation of the vertex operators is given in the next section.

12.2.2 Vertex Operators

Since the hamiltonian (12.2.5) does not depend on φ_0, it commutes with its conjugate momentum π_0 and this quantity can be used as a quantum number to identify the various eigenstates of H. For the decoupling of the analytic and anti-analytic sectors, we can focus attention on one of them, say the analytic one. Let us consider then the analytic part of the vertex operator (12.2.23), denoting by p_0 the value of the conjugate momentum to the zero mode φ_0 in this sector. Let's introduce the "ground states" $|\,\alpha\,\rangle$, with $\alpha = p_0/\sqrt{4\pi g}$. They are characterized by the algebraic conditions

$$\begin{aligned} a_n \,|\,\alpha\,\rangle \;&=\; 0 \quad (n>0), \\ a_0 \,|\,\alpha\,\rangle \;&=\; \alpha \,|\,\alpha\,\rangle. \end{aligned} \tag{12.2.26}$$

From eqn (12.2.21), it can be easily seen that $|\,\alpha\,\rangle$ has conformal weight $\frac{\alpha^2}{8\pi g}$ and any other states of the Fock space of this theory are obtained by acting on $|\,\alpha\rangle$ with the creation operators a_{-n} $(n>0)$. These ground states are in one-to-one correspondence with the vertex operators. In fact, as shown in Problem 1, the ground state $|\,\alpha\rangle$ comes from the application of the vertex operator $V_\alpha(z) =: e^{i\alpha\phi(z)}$ to the conformal vacuum state $|\,0\,\rangle$

$$|\,\alpha\,\rangle \;=\; V_\alpha(0)\;|\,0\,\rangle. \tag{12.2.27}$$

Up to now, the real parameter α is a free quantity. However, we can constrain the set of its values by noting that the lagrangian of the massless scalar field is invariant

under the transformation $\varphi \to \varphi + \delta$, where δ is a constant. This is the $U(1)$ symmetry generated by the current $i\partial\varphi$ and permits us to identify the field φ with $\varphi + 2\pi R$: this compactification is equivalent to regarding φ as an angular variable along a circle of radius R. In this new interpretation, the most general boundary conditions are given by

$$\varphi(x + L, t) \equiv \varphi(x, t) + 2\pi m\, R, \tag{12.2.28}$$

where $m \in \mathbf{Z}$ is the number of times that φ winds in its internal space when the space coordinate reaches the edge of the cylinder. The compactification of φ induces a quantization in integer multiples of $1/R$ of its conjugate momentum π_0: the operator associated to the zero mode also becomes an angular variable and, with the identification $\varphi_0 \equiv \varphi_0 + 2\pi R$, only the exponentials $e^{ie\varphi_0/R}$, with $e \in \mathbf{Z}$, are well-defined. Since $[\varphi_0, \pi_0] = i$, we then have

$$e^{-ie\varphi_0/R}\, \pi_0\, e^{ie\varphi_0/R} \;=\; \pi_0 + \frac{e}{R}. \tag{12.2.29}$$

In complex coordinates and in terms of the integers e and m introduced above, the new expansion of the field is given by

$$\begin{aligned} \varphi(z, \bar{z}) = \varphi_0 &- i \left(\frac{e}{4\pi g R} + \frac{mR}{2} \right) \ln z + \frac{i}{\sqrt{4\pi g}} \sum_{p \neq 0} \frac{1}{p} a_p\, z^{-p} \\ &- i \left(\frac{e}{4\pi g R} - \frac{mR}{2} \right) \ln \bar{z} + \frac{i}{\sqrt{4\pi g}} \sum_{p \neq 0} \frac{1}{p} \bar{a}_p\, \bar{z}^{-p}. \end{aligned} \tag{12.2.30}$$

For the modes L_0 and $\bar{L}_0$ we have

$$\begin{aligned} L_0 \;&=\; \sum_{p>0} a_{-p} a_p + 2\pi g \left(\frac{mR}{2} + \frac{e}{4\pi g R} \right)^2, \\ \bar{L}_0 \;&=\; \sum_{p>0} \bar{a}_{-p} \bar{a}_p + 2\pi g \left(\frac{mR}{2} - \frac{e}{4\pi g R} \right)^2. \end{aligned} \tag{12.2.31}$$

In terms of the integers e and m we can now define the most general expression of the vertex operator

$$\begin{aligned} V_{e,m}(z, \bar{z}) &= :\, \exp \left[i \left(\frac{e}{4\pi g R} + \frac{mR}{2} \right) \phi(z) + i \left(\frac{e}{4\pi g R} - \frac{mR}{2} \right) \bar{\phi}(\bar{z}) \right] :\\ &= :\, \exp \left[i\, \frac{e}{4\pi g R}\, \varphi(z, \bar{z}) + i\, \frac{mR}{2}\, \theta(z, \bar{z}) \right] : \end{aligned} \tag{12.2.32}$$

whose anomalous dimension and spin are given by

$$\begin{aligned} \eta_{e,m} = \Delta + \bar{\Delta} &= \frac{1}{4\pi g}\left(\frac{e^2}{(4\pi g R)^2} + \frac{m^2 R^2}{4}\right), \\ S_{e,m} = \Delta - \bar{\Delta} &= \frac{e\,m}{(4\pi g)^2}. \end{aligned} \tag{12.2.33}$$

To simplify the formula below, it is convenient to assume $g = 1/4\pi$. With such a choice, the previous expressions become

$$\begin{aligned} \eta_{e,m} &= \left(\frac{e^2}{R^2} + \frac{m^2 R^2}{4}\right), \\ S_{e,m} &= e\,m. \end{aligned} \tag{12.2.34}$$

Note that the simultaneous substitutions $R \leftrightarrow 2/R$ and $m \leftrightarrow e$ leave invariant the spectrum of both the anomalous dimensions and spins. This observation will be useful in the discussion of the partition functions of the next section. In the language of the Coulomb gas, the integers e and m can be identified with the *electric* and *magnetic* charges of the system, respectively. The reason for this interpretation becomes evident if one considers the operator expansion

$$\varphi(z_1,\bar{z}_1)\, V_{e,m}(z_2,\bar{z}_2) \;=\; -\left[\frac{e}{R}\,\ln|z_{12}|^2 \,+\, \frac{mR}{2}\,\ln\frac{z_{12}}{\bar{z}_{12}}\right]\, V_{e,m}(z_2,\bar{z}_2) + \cdots \tag{12.2.35}$$

If we consider only the purely electric vertex operator

$$V_{e,0}(z_2,\bar{z}_2) \;=:\; e^{i\,\frac{e}{R}\,\varphi(z_2,\bar{z}_2)}\;:$$

and we wind $\varphi(z_1,\bar{z}_1)$ around the point $(z_2,\bar{z}_2)$ at which the vertex operator acts (by making the analytic continuation $z_{12} \to e^{2\pi i}\, z_{12}$, $\bar{z}_{12} \to e^{-2\pi i}\,\bar{z}_{12}$), this operation does not induce any discontinuity in the field φ. However, repeating the same operation in the presence of the purely magnetic vertex operator

$$V_{0,m}(z_2,\bar{z}_2) \;=:\; e^{i\,\frac{mR}{2}\,\theta(z_2,\bar{z}_2)}\;:$$

the field $\varphi(z_1,\bar{z}_1)$ has a jump equal to $2\pi m R$. The most general vertex operator $V_{e,m}$ is a combination of electric and magnetic vertex operators and its two-point correlation function is given by

$$\langle V_{e,m}(z_1,\bar{z}_1) V_{-e,-m}(z_2,\bar{z}_2)\rangle \;=\; \frac{1}{|\,z_{12}\,|^{\eta}}\,\left(\frac{z_{12}}{\bar{z}_{12}}\right)^{S}. \tag{12.2.36}$$

Let's now discuss the partition function of the bosonic field on a torus, highlighting its remarkable duality properties.

12.2.3 Free Bosonic Field on a Torus

Let us initially consider the partition function of a gaussian free bosonic theory. In this case the variable φ takes values on all the real axis. In terms of the path integral, its expression would be

$$Z(\tau) \,=\, \int \mathcal{D}\varphi \; e^{-\mathcal{S}}, \tag{12.2.37}$$

with

$$\mathcal{S} \,=\, \frac{g}{2} \int d^2x \, (\partial\varphi)^2 \,=\, -\frac{g}{2} \int d^2x \, \varphi \, \Box \, \varphi, \tag{12.2.38}$$

where we have assumed periodic boundary conditions on both directions of the torus

$$\begin{array}{l} \varphi(z+\tau, \bar{z}+\bar{\tau}) \,=\, \varphi(z,\bar{z}), \\ \varphi(z+1, \bar{z}+1) \,=\, \varphi(z,\bar{z}). \end{array} \tag{12.2.39}$$

The definition (12.2.37) presents certain drawbacks. For the quadratic expression of the action, the functional integral reduces to the product of the eigenvalues $\lambda_{n,m}$ of the laplacian $\Box$ on the torus

$$Z \,\sim\, \prod_{n,m} \left(\frac{1}{\lambda_{n,m}} \right)^{1/2}. \tag{12.2.40}$$

Among the eigenvalues, there is $\lambda_{0,0} = 0$, which corresponds to the zero mode of the field associated to its constant configuration. The original definition of the partition function (12.2.37) is therefore divergent. For the correct definition of this quantity it is necessary to restrict the functional integration only to the non-zero modes of the field φ. To this end, let's define

$$Z_B(\tau) \,=\, 2\pi \int \mathcal{D}\varphi \, \sqrt{A} \, \delta \left(\int d^2x \varphi(x) \, \varphi_0 \right) \, e^{-\mathcal{S}}, \tag{12.2.41}$$

where A is the area of the torus $A = \mathrm{Im}\,\tau$, $\varphi_0 = A^{-1/2}$ is the normalized eigenfunction of the zero mode on this geometry and the prefactor 2π has been inserted for future convenience. The integral $\int d^2x \varphi(x) \, \varphi_0$ obviously filters the zero mode of the field, on which it is no longer necessary to integrate for the delta-function inserted into the functional integral.

To compute (12.2.41), expand φ on the basis of the normalized eigenfunctions $\varphi_{n,m}$ of the operator $\Box$

$$\varphi \,=\, \sum_{n,m} c_{n,m} \, \varphi_{n,m}.$$

The eigenvalues corresponding to the boundary conditions (12.2.39) are given by

$$\lambda_{n,m} \,=\, (2\pi)^2 \, |nk_2 + mk_1|^2,$$

where $k_{1,2}$ are the vectors of the basis of the lattice that is dual to the original lattice defined by the periods $\omega_{1,2}$. With the choice $(\omega_1, \omega_2) = (1, \tau)$, one has

$$k_1 \,=\, -i\,\omega_2/A \,=\, -i\,\tau/A, \qquad k_2 \,=\, i\omega_1/A \,=\, i/A$$

namely

$$\lambda_{n,m} \;=\; \left(\frac{2\pi}{A}\right)^2 |n - m\tau|^2. \tag{12.2.42}$$

For the partition function (12.2.41) we then have

$$Z_B(\tau) = 2\pi\sqrt{A}\int \prod_{n,m}{}' dc_{n,m} e^{-\frac{g}{2}\sum_{n,m}\lambda_{n,m}c_{n,m}^2} = \sqrt{\frac{A}{\det{}'\frac{g}{2\pi}\,\Box}} = \sqrt{A}\prod_{n,m}{}'\left(\frac{1}{\frac{g}{2\pi}\lambda_{n,m}}\right)^{1/2} \tag{12.2.43}$$

where the index in the product means the omission of the term $n = m = 0$. To evaluate this infinite product we use the regularization given by the Riemann zeta function. We recall that, with the usual definition of this function, $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$, we have $\zeta(-1) = -\frac{1}{12}$, $\zeta(0) = -\frac{1}{2}$ and $\zeta'(0) = \frac{d\zeta(0)}{ds} = -\frac{1}{2}\ln 2$. So, with this regularization, we have for instance

$$\prod_{n=1}^{\infty} a \;=\; a^{\zeta(0)} \;=\; a^{-1/2}, \qquad \prod_{n=-\infty}^{\infty} a \;=\; a^{2\zeta(0)+1} \;=\; 1.$$

Other useful formulas are

$$\prod_{n=1}^{\infty} n^{\alpha} \;=\; e^{-\alpha\zeta'(0)} = (2\pi)^{\alpha/2},$$
$$\prod_{n=-\infty}^{\infty} (n+a) \;=\; a\prod_{n=1}^{\infty}(-n^2)\left(1 - \frac{a^2}{n^2}\right) \;=\; 2i\,\sin\pi\,a.$$

Applying these expressions, one has

$$\begin{aligned}
\det{}'\,\frac{g}{2\pi}\,\Box &= \prod_{(n,m)\neq(0,0)}\left(\frac{\sqrt{g}}{A}\right)^2 (n - m\tau)\,(n - m\bar\tau) \\
&= \left(\frac{A}{\sqrt{g}}\right)^2 \left(\prod_{n\neq 0} n^2\right) \prod_{m\neq 0, n\in\mathbf{Z}} (n - m\tau)\,(n - m\bar\tau) \\
&= \left(\frac{A}{\sqrt{g}}\right)^2 (2\pi)^2 \prod_{m>0, n\in\mathbf{Z}} (n - m\tau)\,(n + m\tau)\,(n - m\bar\tau)\,(n + m\bar\tau) \\
&= \frac{(2\pi A)^2}{g} \prod_{m>0} \left(e^{-i\pi m\tau} - e^{i\pi m\tau}\right)^2 \left(e^{-i\pi m\bar\tau} - e^{i\pi m\bar\tau}\right)^2 \\
&= \frac{(2\pi A)^2}{g} \prod_{m>0} (q\,\bar q)^{-m}\,(1 - q^m)^2\,(1 - \bar q^m)^2 \\
&= \frac{(2\pi A)^2}{g}\,(q\,\bar q)^{\frac{1}{12}} \prod_{m>0} (1 - q^m)^2\,(1 - \bar q^m)^2 \;=\; \frac{(2\pi A)^2}{g}\,\eta^2\,\bar\eta^2,
\end{aligned}$$

where $\eta(q)$ is the Dedekind function

$$\eta(q) \;=\; q^{\frac{1}{24}} \prod_{m=1}^{\infty} (1 - q^m). \tag{12.2.44}$$

Substituting in (12.2.43), we arrive at the final expression of the partition function of a gaussian bosonic field on a torus:

$$Z_B(\tau) \;=\; \frac{g^{1/2}}{(\mathrm{Im}\,\tau)^{1/2}\eta(q)\,\eta(\bar{q})}. \tag{12.2.45}$$

To check that this function is invariant under the modular group, we need the transformations of the Dedekink function under $\mathcal{T}$ and $\mathcal{S}$:

$$\begin{aligned} \eta(\tau+1) \;&=\; e^{i\pi/12}\,\eta(\tau), \\ \eta(-1/\tau) \;&=\; \sqrt{-i\tau}\,\eta(\tau). \end{aligned} \tag{12.2.46}$$

These formulas are derived by using the identity

$$\eta(\tau) \;=\; \frac{1}{2}\theta_2(\tau)\,\theta_3(\tau)\,\theta_4(\tau), \tag{12.2.47}$$

and the modular tranformations of the Jacobi $\theta_i(\tau)$ functions, discussed in Problem 2 at the end of the chapter.

Let's now generalize the previous result (12.2.45) when the scalar field φ is compactified on a circle of radius R. The equations of motion are obviously the same as for the gaussian case but the boundary conditions are different. Instead of those expressed by (12.2.39), we have in fact

$$\begin{aligned} \varphi(z+\omega_1, \bar{z}+\bar{\omega}_1) \;&=\; \varphi(z,\bar{z}) + 2\pi R\, m, \\ \varphi(z+\omega_2, \bar{z}+\bar{\omega}_2) \;&=\; \varphi(z,\bar{z}) + 2\pi R\, n. \end{aligned} \tag{12.2.48}$$

The integers (m, n) identify a specific topological class of field configurations of φ and, integrating over these configurations, we define the corresponding partition function $Z_{m,n}$. To compute such a quantity, let's decompose the field in terms of its classical solution $\varphi^{cl}_{m,n}$ (which satisfies the boundary conditions (12.2.48)) and of its fluctuation $\tilde{\varphi}$, that is a fully periodic function

$$\begin{aligned} \varphi &= \varphi^{cl}_{m,n} + \tilde{\varphi}, \\ \varphi^{cl}_{m,n} &= 2\pi R \left[\frac{z}{\omega_1} \frac{m\bar{\tau} - n}{\bar{\tau} - \tau} - \frac{\bar{z}}{\omega_1 *} \frac{m\tau - n}{\bar{\tau} - \tau} \right]. \end{aligned} \tag{12.2.49}$$

Substituting this expression in the action (12.2.38), we can decompose this quantity into the action $\mathcal{S}[\tilde{\varphi}]$ of the periodic field and into the action $\mathcal{S}[\varphi^{cl}_{m,n}]$ relative to the

classical configuration of the field. The latter quantity is expressed by

$$\begin{aligned}\mathcal{S}[\varphi^{cl}_{m,n}] &= \frac{g}{2}\int d^2x\,(\nabla\,\varphi^{cl}_{m,n})^2,\\ &= 2g\int dz\,d\bar{z}\,\partial\varphi^{cl}_{m,n}\,\bar{\partial}\varphi^{cl}_{m,n}\\ &= 8\pi^2\,g\,R^2\,A\frac{1}{|\omega|^2}\left|\frac{m\tau-n}{\tau-\bar{\tau}}\right| \qquad (12.2.50)\\ &= 2\pi^2\,g\,R^2\,\frac{|m\tau-n|^2}{\mathrm{Im}\,\tau}.\end{aligned}$$

The functional integral on the periodic term $\tilde{\varphi}$ of the field gives rise to the prefactor $Z_B(\tau)$ previously computed and therefore

$$Z_{m,n}(\tau)\;=\;Z_B(\tau)\,\exp\left[-2g\pi^2R^2\frac{|m\tau-n|^2}{\mathrm{Im}\,\tau}\right]. \qquad (12.2.51)$$

Let's determine the transformation properties of this expression under the modular group. For a generic modular transformation, the parameter τ changes as

$$\tau\to(a\tau+b)/(c\tau+d)$$

and therefore

$$\begin{aligned}\frac{|m\tau-n|^2}{\mathrm{Im}\,\tau}&\to\frac{|(ma\tau+bm)/(c\tau+d)-n|^2\,|c\tau+d|^2}{\mathrm{Im}\,[(a\tau+b)(c\tau+d)]}\\ &=\frac{|(ma-nc)\tau+bm-dn|^2}{\mathrm{Im}\,\tau},\end{aligned}$$

where we use the formula

$$\mathrm{Im}\,[(a\tau+b)(c\tau+d)]\;=\;\mathrm{Im}[(ad-bc)\tau]\;=\;\mathrm{Im}\,\tau\qquad(ad-bc=1).$$

Hence, under a modular transformation, the indices (m,n) transform with the matrix

$$\begin{pmatrix}m\\n\end{pmatrix}\to\begin{pmatrix}a&-c\\-b&d\end{pmatrix}\begin{pmatrix}m\\n\end{pmatrix}, \qquad (12.2.52)$$

so that

$$\begin{aligned}Z_{m,n}(\tau+1)\;&=\;Z_{m,n-m},\\ Z_{m,n}(-1/\tau)\;&=\;Z_{-n,m}.\end{aligned} \qquad (12.2.53)$$

To have a modular invariant partition function one simply needs to sum over all sectors relative to the different boundary conditions. We arrive then at the final expression

$$Z(R)\;=\;\sqrt{2\pi g}\,R\,\frac{1}{\mathrm{Im}\,\tau|\eta(\tau)|^2}\sum_{m,n}\exp\left[-2g\pi^2R^2\frac{|m\tau-n|^2}{\mathrm{Im}\,\tau}\right], \qquad (12.2.54)$$

where the prefactor $\sqrt{2\pi g}\, R$ comes from integration over the zero mode of the field. This term can also be justified in a different way, i.e. transforming the previous expression with the Poisson resummation formula

$$\sum_{n=-\infty}^{\infty} \exp[-\pi a n^2 + bn] \;=\; \frac{1}{\sqrt{a}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{\pi}{a}\left(k + \frac{b}{2\pi i}\right)^2\right]. \tag{12.2.55}$$

Imposing for simplicity $g = 1/4\pi$ and

$$a = R^2/2\tau_2 \qquad b = \pi m R^2 \tau_1/\tau_2 \qquad \tau \;=\; \tau_1 + i\tau_2$$

we arrive at

$$Z(R) \;=\; \frac{1}{|\eta(\tau)|^2} \sum_{e,m\in\mathbf{Z}} q^{(e/R+mR/2)^2/2}\, \bar{q}^{(e/R-mR/2)^2/2}. \tag{12.2.56}$$

Comparing with eqn (11.7.13), it is easy to see that the expressions for L_0 and $\bar{L}_0$ coincide with those given in (12.2.31). The spectrum of the anomalous dimensions and spins, given in eqn (12.2.34), shows that the partition function is symmetric under the simultaneous change of $e \leftrightarrow m$ and $R \leftrightarrow 2/R$. This leads to the duality relation of the partition function (12.2.56)

$$Z(R) \;=\; Z(2/R). \tag{12.2.57}$$

The computation of the partition function for the self-dual value $R = \sqrt{2}$ is proposed in Problem 4 at the end of the chapter.

12.3 Conformal Field Theory of a Free Fermionic Field

In this section we discuss the conformal theory of the complex fermion field (Dirac field)

$$\Psi(z,\bar{z}) \;=\; \begin{pmatrix} \chi(z,\bar{z}) \\ \bar{\chi}(z,\bar{z}) \end{pmatrix}, \tag{12.3.1}$$

in euclidean space. The action is

$$\mathcal{S} \;=\; \frac{\lambda}{2\pi} \int d^2x\, \bar{\Psi}\, \gamma^\mu\, \partial_\mu\, \Psi, \tag{12.3.2}$$

where $\bar{\Psi} = \Psi^\dagger\, \gamma^0$, while the euclidean Dirac matrices γ^μ satisfy the algebra

$$\{\gamma^\mu, \gamma^\nu\} \;=\; 2\delta^{\mu,\nu}. \tag{12.3.3}$$

We choose as representation of the γ^μ matrices

$$\gamma^0 \;=\; \sigma_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^1 \;=\; \sigma_2 \;=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{12.3.4}$$

where σ_i are the usual Pauli matrices. The two-dimensional analog of the γ^5 matrix is here given by σ_3. In complex coordinates, the euclidean Dirac operator associated

to this fermion is

$$\mathcal{D} = \gamma^0 \partial_\tau + \gamma^1 \partial_x = \begin{pmatrix} 0 & 2\partial_z \\ 2\partial_{\bar{z}} & 0 \end{pmatrix}, \tag{12.3.5}$$

and the equations of motion are

$$\begin{aligned} \partial_{\bar{z}} \chi(z, \bar{z}) &= 0, \\ \partial_z \bar{\chi}(z, \bar{z}) &= 0. \end{aligned} \tag{12.3.6}$$

They show that $\chi(z, \bar{z}) = \chi(z)$ is a purely analytic field whereas $\bar{\chi}(z, \bar{z}) = \bar{\chi}(\bar{z})$ is purely anti-analytic. The two-point correlation functions are

$$\begin{aligned} \langle \chi^\dagger(z_1)\, \chi(z_2) \rangle &= \frac{1}{\lambda} \frac{1}{z_1 - z_2}, \\ \langle \bar{\chi}^\dagger(\bar{z}_1)\, \bar{\chi}(\bar{z}_2) \rangle &= \frac{1}{\lambda} \frac{1}{\bar{z}_1 - \bar{z}_2}, \\ \langle \chi^\dagger(z_1)\, \bar{\chi}(\bar{z}_2) \rangle &= \langle \bar{\chi}^\dagger(\bar{z}_1)\, \chi(z_2) \rangle = 0. \end{aligned} \tag{12.3.7}$$

It should be noticed that the complex fermion field Ψ can be written in terms of the two real Majorana fermions ψ_1 and ψ_2, with $\psi_i = \psi_i^\dagger$

$$\Psi(z, \bar{z}) = \begin{pmatrix} \chi(z) \\ \bar{\chi}(\bar{z}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \psi_1 + i\psi_2 \\ \bar{\psi}_1 + i\bar{\psi}_2 \end{pmatrix}. \tag{12.3.8}$$

Since $\chi^\dagger = (\psi_1 - i\psi_2)/\sqrt{2}$, the analytic component of the stress-energy tensor of this theory is

$$T(z) = \frac{\lambda}{2} : (\partial \Psi^\dagger\, \Psi - \Psi^\dagger\, \partial \Psi) := -\frac{\lambda}{2} : (\psi_1 \partial \psi_1 + \psi_2 \partial \psi_2) : \tag{12.3.9}$$

and is given by the sum of the stress–energy tensors relative to the two real fermions ψ_1 and ψ_2. From the correlator $\langle T(z_1) T(z_2) \rangle$ (which can be computed using the results of Chapter 11 for the Majorana fermion), one obtains the value of the central charge, $c = 1$. Analogous formulas hold for the anti-analytic component of the fermion field.

To study the quantization of Ψ it is sufficient to consider the quantization of its Majorana components. We deal with this problem in the next section.

12.3.1 Quantization of the Free Majorana Fermion

In this section and in the next ones, we denote by $\psi(z)$ and $\bar{\psi}(\bar{z})$ the analytic and anti-analytic components of the Majorana fermion, with action[2]

$$\mathcal{S} = \frac{1}{2\pi} \int d^2x \left[\psi\, \partial_{\bar{z}}\, \psi + \bar{\psi}\, \partial_z\, \bar{\psi} \right]. \tag{12.3.10}$$

The equations of motion

$$\begin{aligned} \partial_z\, \bar{\psi} &= 0, \\ \partial_{\bar{z}} \psi &= 0, \end{aligned} \tag{12.3.11}$$

[2]To simplify the notation from now on we take $\lambda = 1$.

show that the two components are decoupled. Moreover, ψ depends only on z, whereas $\bar{\psi}$ depends only on $\bar{z}$. The conformal weights of the two fields are

$$\psi \to \left(\frac{1}{2}, 0\right) \qquad \bar{\psi} \to \left(0, \frac{1}{2}\right).$$

As seen previously, the analytic and anti-analytic components of the stress–energy tensor associated to the action (12.3.10) are

$$T \,=\, -\frac{1}{2} : \psi \partial_z \psi :, \quad \bar{T} \,=\, -\frac{1}{2} : \bar{\psi} \partial_{\bar{z}} \bar{\psi} : . \tag{12.3.12}$$

Let's now focus attention on the analytic sector, since analogous considerations can be applied to the anti-analytic one.

Given the conformal weight of $\psi(z)$, the operator product expansion with itself is

$$\psi(z_1)\psi(z_2) \,=\, \frac{1}{z_1 - z_2} + \cdots \tag{12.3.13}$$

In the complex plane, the mode expansion of the Taylor–Laurent series reads

$$\psi(z) \,=\, \sum_{n=-\infty}^{\infty} \frac{\psi_n}{z^{n+1/2}}, \tag{12.3.14}$$

where

$$\psi_n \,=\, \oint_C \frac{dz}{2\pi i} z^{n-1/2}\, \psi(z), \tag{12.3.15}$$

with C a closed contour around the origin. Using eqn (12.3.13), we can derive the anticommutation relations of the modes: we simply need to use the operator product expansion and exchange, as usual, the order of the contours around the origin

$$\begin{aligned} \{\psi_n, \psi_m\} &= \left[\oint \frac{dz}{2\pi i}, \oint \frac{dw}{2\pi i}\right] z^{n-1/2}\, w^{m-1/2}\, \psi(z)\psi(w) \\ &= \oint \frac{dw}{2\pi i} w^{m-1/2} \oint \frac{dz}{2\pi i} z^{n-1/2}\, \frac{1}{z-w} \\ &= \oint \frac{dw}{2\pi i} w^{m+n-1} \,=\, \delta_{n+m,0}. \end{aligned} \tag{12.3.16}$$

Neveu–Schwarz and Ramond sectors. It is worth stressing that we can choose two different monodromy properties of the field $\psi(z)$. In fact, the fermion field is naturally defined on the double covering of the complex plane: with a branch cut that starts from the origin, when the coordinate z goes around the origin

$$\psi(e^{2\pi i}\, z) \,=\, \pm\, \psi(z) \tag{12.3.17}$$

we can adopt either periodic (P) or antiperiodic (A) boundary conditions. The first case defines the so-called Neveu–Schwarz (NS) sector, while the second defines the so-called Ramond (R) sector. In the Neveu–Schwarz sector, the mode expansion of the

field is given in terms of half-integer indices, while in the Ramond sector the indices n of the (12.3.14) are instead integers

$$\begin{array}{lll} \psi(e^{2\pi i}\, z) \,=\, \psi(z), & n \in \mathbf{Z} + \frac{1}{2}, & (NS) \\ \psi(e^{2\pi i}\, z) \,=\, -\psi(z), & n \in \mathbf{Z}, & (R). \end{array} \tag{12.3.18}$$

It is also convenient to introduce the operator $(-1)^F$, where F is the fermionic number, defined in terms of its anticommutation with the field ψ

$$(-1)^F\, \psi(z) \,=\, -\psi(z)\, (-1)^F.$$

This operator satisfies $\left((-1)^F\right)^2 = 1$ and

$$\left\{(-1)^F, \psi_n\right\} \,=\, 0, \quad \forall n. \tag{12.3.19}$$

There are some interesting consequences of the integer or half-integer mode expansion of the field both for its correlation functions and for the operator content. Let's analyze first the periodic case: to compute its two-point function of the vacuum state, we can use the anticommutations of its modes, keeping in mind that

$$\begin{array}{l} \psi_n \mid 0\rangle \,=\, 0, \quad n > 0 \\ \langle 0 \mid \psi_n \,=\, 0, \quad n < 0. \end{array} \tag{12.3.20}$$

Hence, we have

$$\begin{aligned} \langle 0 \mid \psi(z)\psi(w) \mid 0\rangle &= \langle 0 \mid \sum_{n=1/2}^{\infty} \psi_n\, z^{-n-1/2} \sum_{m=-1/2}^{-\infty} \psi_m\, w^{-m-1/2} \mid 0\rangle \\ &= \sum_{n=1/2}^{\infty} z^{-n-1/2}\, w^{n-1/2} \,=\, \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{w}{z}\right)^n \,=\, \frac{1}{z-w}. \end{aligned} \tag{12.3.21}$$

Let's consider now the two-point correlation function when the field satisfies the antiperiodic boundary conditions. In such a case, we have to take into account the presence of the zero mode of the field that satisfies

$$\{\psi_0, \psi_0\} \,=\, 1, \qquad \{(-1)^F, \psi_0\} \,=\, 0. \tag{12.3.22}$$

Applying ψ_0 to an eigenstate of L_0 does not change its eigenvalue. This means that the ground state of the Ramond sector must realize a representation of the two-dimensional algebra given by ψ_0 and $(-1)^F$. The smallest irreducible representation consists of a doublet of operators σ and μ, the so-called *order and disorder operators*, with the same conformal weight. In this space, a 2×2 matrix representation of ψ_0 and $(-1)^F$ is given by

$$\psi_0 \,=\, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad (-1)^F \,=\, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{12.3.23}$$

In this representation the fields σ and μ are eigenvectors of $(-1)^F$ with eigenvalue $+1$ and -1, respectively.

In the presence of the order/disorder fields, the OPE of the fermionic field is

$$\psi(z)\sigma(w) \sim (z-w)^{-1/2}\,\mu(w) + \cdots ; \qquad \psi(z)\mu(w) \sim (z-w)^{-1/2}\,\sigma(w) + \cdots \tag{12.3.24}$$

Therefore we can interpret the two-point correlation function of the field $\psi(z)$ with antiperiodic boundary conditions as their correlation in the presence of these two fields, placed at the origin and at infinity, respectively:

$$\langle\psi(z)\psi(w)\rangle_A \equiv \langle\, 0 \mid \sigma(\infty)\psi(z)\psi(w)\sigma(0) \mid 0\rangle = \langle\, 0 \mid \mu(\infty)\psi(z)\psi(w)\mu(0) \mid 0\rangle. \tag{12.3.25}$$

To compute these correlators, we can use the expansion in integer modes of ψ: separating the zero mode, and using in this computation simply its vacuum expectation value $\psi_0^2 = \frac{1}{2}$, we obtain

$$\begin{aligned}\langle\psi(z)\psi(w)\rangle_A &= \left\langle \sum_{n=0} \psi_n\, z^{-n-1/2} \sum_{m=0}^{-\infty} \psi_m\, w^{-m-1/2} \right\rangle_A \\ &= \sum_{n=1}^{\infty} z^{-n-1/2}\, w^{n-1/2} + \frac{1}{2}\frac{1}{\sqrt{z\,w}} \\ &= \frac{1}{\sqrt{zw}}\left(\frac{w}{z-w} + \frac{1}{2}\right) = \frac{1}{2}\frac{\sqrt{\frac{z}{w}} + \sqrt{w}z}{z-w}.\end{aligned} \tag{12.3.26}$$

Note that, in the limit $z \to w$, this result correctly reproduces the operator product expansion (12.3.13), as expected, because this relation expresses a *local* property of the field and is insensitive to the boundary conditions chosen for it.

It is now easy to compute the conformal weights of the fields σ and μ. Let's firstly use the Ward identity

$$T(z)\sigma(0) \mid 0\rangle = \frac{\Delta_\sigma}{z^2}\,\sigma(0) \mid 0\rangle + \cdots$$

which leads to

$$\langle T(z)\rangle_A \equiv \langle 0 \mid \sigma(\infty)T(z)\sigma(0) \mid 0\rangle = \frac{\Delta_\sigma}{z^2}. \tag{12.3.27}$$

The left-hand side of this equation can be evaluated using both the definition of the normal order

$$T(z) = \lim_{\eta\to 0}\frac{1}{2}\left(-\psi(z+\eta)\partial_z\psi(z) + \frac{1}{\eta^2}\right) \tag{12.3.28}$$

and the correlation function (12.3.26). Hence,

$$\langle T(z)\rangle_A = \lim_{w\to z}\left[-\frac{1}{4}\,\partial_w\left(\frac{\sqrt{z/w} + \sqrt{w/z}}{z-w}\right) + \frac{1}{2(z-w)^2}\right] = \frac{1}{16z^2} \tag{12.3.29}$$

and so

$$\Delta_\sigma = \Delta_\mu = \frac{1}{16}. \tag{12.3.30}$$

Bosonic order/disorder operators. It is interesting to remark that an analogous result for the periodic and antiperiodic boundary conditions also holds for the bosonic field. In fact, in view of the symmetry of the action under $\varphi \to -\varphi$, also in this theory we can adopt antiperiodic boundary conditions. Consider, for instance, $J(z)$, the analytic component of the current, with expansion

$$J(z) \;=\; i\,\partial_z \Phi(z) \;=\; \sum_n a_n\, z^{-n-1}. \tag{12.3.31}$$

When $J(e^{2\pi i}\, z) = J(z)$, $n \in \mathbf{Z}$, while when $J(e^{2\pi i}\, z) = -J(z)$, we have $n \in \mathbf{Z} + 1/2$. As for the fermions, in the antiperiodic case we can introduce the order/disorder operators ς and τ, with operator expansion

$$\partial\Phi(z)\,\varsigma(w) \;=\; (z-w)^{-1/2}\,\tau(w) + \cdots \tag{12.3.32}$$

Contrary to the fermionic case, in this case the two fields have different conformal weights, related by $\Delta_\tau = \Delta_\varsigma + \frac{1}{2}$. In the presence of antiperiodic boundary conditions, the two-point correlation function of the current is given by

$$\langle\, \partial\Phi(z)\, \partial\Phi(w)\,\rangle_A \;\equiv\; \langle\, 0\,|\,\sigma(\infty)\partial\Phi(z)\,\partial\Phi(w)\,\sigma(0)\,|\,0\,\rangle. \tag{12.3.33}$$

Repeating the same computation as in the fermionic field, we have

$$-\langle\, \partial\Phi(z)\,\partial\Phi(w)\rangle_A \;=\; \frac{\left(\sqrt{\frac{z}{w}} + \sqrt{\frac{w}{z}}\right)}{2(z-w)^2}. \tag{12.3.34}$$

The conformal weight of ς can be derived by the vacuum expectation value of the stress–energy tensor

$$\langle T(z)\rangle_A \;=\; -\frac{1}{2}\lim_{z\to w}\left\langle \partial\Phi(z)\partial\Phi(w) + \frac{1}{(z-w)^2}\right\rangle_A \;=\; \frac{1}{16z^2} \tag{12.3.35}$$

namely $\Delta_\varsigma = \frac{1}{16}$.

Using eqn (12.3.14), the stress–energy tensor becomes

$$T(z) \;=\; \frac{1}{2}\sum_{n,k}\left(k+\frac{1}{2}\right)\, z^{-n-2}\;:\,\psi_{n-k}\,\psi_k\,: \tag{12.3.36}$$

where by the normal order :: we mean an ordering of the operators, with the lowest index placed on the left. Since $T(z) = \sum_n L_n\, z^{-n-2}$, the Virasoro generators are

$$L_n \;=\; \frac{1}{2}\sum_k\left(k+\frac{1}{2}\right)\;:\,\psi_{n-k}\,\psi_k\,:. \tag{12.3.37}$$

For the generators L_n (with $n \neq 0$) there is no problem to implement the normal order, since the operators involved in their definition anticommute. However, we have to pay attention to the definition of L_0, in which there may be an additive constant coming from the anticommutation of the operators ψ_{-k} and ψ_k. This constant can be determined by the vacuum expectation of $T(z)$. However, we have to distinguish the Neveu–Schwarz and the Ramond sectors

$$\begin{aligned} L_0 &= \sum_{k>0} k\,\psi_{-k}\,\psi_k \qquad \left(\mathrm{NS}, k \in \mathbf{Z} + \frac{1}{2}\right) \\ L_0 &= \sum_{k>0} k\,\psi_{-k}\,\psi_k + \frac{1}{16} \quad (\mathrm{R}, k \in \mathbf{Z})\,. \end{aligned} \tag{12.3.38}$$

12.3.2 Fermions on a Torus

To discuss the partition function of the fermionic theory, let's initially consider the transformation that maps the plane in a cylinder geometry of width L. This is given by

$$w = \frac{L}{2\pi}\log z. \tag{12.3.39}$$

Since the fermionic field has conformal weight $1/2$, the field ψ on the cylinder is related to the field ψ_{pl} on the plane by the transformation

$$\psi(w) = \left(\frac{dz}{dw}\right)^{1/2} \psi_{pl}(z) = \sqrt{\frac{2\pi z}{L}}\,\psi_{pl}(z). \tag{12.3.40}$$

Let x be the space coordinate along the cylinder and τ its euclidean time variable, such that $w = \tau - ix$. At a fixed τ, using eqn (12.3.40) and the mode expansion of the field in the plane, we can easily derive the expansion of the field on the cylinder

$$\psi(x) = \sqrt{\frac{2\pi}{L}} \sum_k \psi_k\, e^{2\pi i k x/L}. \tag{12.3.41}$$

Since it is a free theory, the euclidean time evolution of the modes is expressed by

$$\psi_k(t) = \psi_k(0)\, e^{-2\pi k\tau/L}, \tag{12.3.42}$$

and therefore, for any x and τ, we have the expansion

$$\psi(w) = \sqrt{\frac{2\pi}{L}} \sum_k \psi_k\, e^{-2\pi k w/L}, \tag{12.3.43}$$

where $\psi_k = \psi_k(0)$. Notice that, for the transformation law (12.3.40), on the cylinder there is a *swapping* of the boundary conditions with respect to those of the plane: the Ramond field, which has an integer mode expansion, now corresponds to periodic

boundary conditions while the Neveu–Schwarz field, with half-integer modes, satisfies antiperiodic boundary conditions

$$\begin{array}{l} \psi(x+2\pi L) \,=\, \psi \quad \text{(Ramond)} \\ \psi(x+2\pi L) \,=\, -\psi \quad \text{(Neveu}-\text{Schwarz)}. \end{array} \tag{12.3.44}$$

It is interesting to compute L_0 for the two boundary conditions

$$(L_0)_{cyl} \,=\, \frac{1}{2}\sum_n n \,:\, \psi_{-n}\psi_n \,:=\, \sum_{n>0} n\psi_{-n}\psi_n - \frac{1}{2}\sum_{n>0} n. \tag{12.3.45}$$

The last term is obviously divergent but it can be regularized in terms of the Riemann zeta function. In the Ramond case, the sum is over all the integers

$$\sum_{n=1}^{\infty} n \,=\, \zeta(-1) \,=\, -\frac{1}{12}.$$

In the Neveu–Schwarz case, the sum runs over the half-integers $n = (2k+1)/2$. Such a series can be written as a sum over all the integers, minus the sum over the even numbers

$$\frac{1}{2}\sum_{k=0}^{\infty}(2k+1) \,=\, \frac{1}{2}\left[\sum_{m=1}^{\infty} m - \sum_{m=1}^{\infty} 2m\right] \,=\, -\frac{1}{2}\,\zeta(-1) \,=\, \frac{1}{24}.$$

Hence

$$(L_0)_{cyl} \,=\, \sum_{n>0} \psi_{-n}\,\psi_n + \begin{cases} \frac{1}{24} & \text{Ramond} \\ -\frac{1}{48} & \text{Neveu}-\text{Schwarz} \end{cases} \tag{12.3.46}$$

where, for Ramond, the sum is over the integers and for Neveu–Schwarz, the half-integers. To interpret the presence of the additional constant, one needs to recall the transformation law of T in passing from the plane to the cylinder:

$$T_{cyl}(w) \,=\, \left(\frac{dz}{dw}\right)^2 T_{pl}(z) + \frac{c}{12}\,\{z,w\} \,=\, \left(\frac{2\pi}{L}\right)^2 \left[z^2\,T_{pl}(z) - \frac{c}{24}\right]. \tag{12.3.47}$$

Substituting $T(z) = \sum_n L_n z^{-n-2}$, we get

$$T_{cyl}(w) \,=\, \left(\frac{2\pi}{L}\right)^2 \left[\sum_n L_n z^{-n} - \frac{c}{24}\right] \,=\, \left(\frac{2\pi}{L}\right)^2 \sum_n \left(L_n - \frac{c}{24}\,\delta_{n,0}\right) e^{-2\pi n w/L}$$

namely

$$(L_0)_{cyl} \,=\, L_0 - \frac{c}{24}. \tag{12.3.48}$$

Since the Majorana fermion has a central charge $c = \frac{1}{2}$, in the Neveu–Schwarz sector we correctly recover the ground state energy given by $-\frac{1}{48}$. In the Ramond sector, we have to take into account the conformal weight of the ground states of this sector, equal to $\frac{1}{16}$: the difference $\frac{1}{24} - (-\frac{1}{48}) = \frac{1}{16}$ is precisely the conformal weight of this state.

Calculus for anticommuting quantities. To proceed, it is necessary to briefly recall the mathematical properties of the anticommuting variables. Let α_i $(i = 1, \ldots, n)$ be a set of anticommuting variables $\{\alpha_i, \alpha_j\} = 0$. Since $\alpha_i^2 = 0$, any function $f(\alpha_1, \ldots, \alpha_n)$ of these variables, once expanded in series, is at most a polynomial of first order α_i. Moreover, for anticommuting variables the integration rules are

$$\int d\alpha_i \, = \, 0 \qquad \int d\alpha_i \, \alpha_j \, = \, \delta_{ij} \tag{12.3.49}$$

namely, the integration corresponds to taking the derivative. Consider the quantity

$$I \, = \, \int d\alpha_1 \, \ldots \, d\alpha_n \, \exp\left[-\alpha_i A_{ij} \alpha_j\right], \tag{12.3.50}$$

where A_{ij} is an antisymmetric matrix of dimension n, where n is an even number. Expanding the exponential, in power series by the nature of the variables α_i, there is only a finite number of terms

$$I \, = \, \int d\alpha_1 \, \ldots \, d\alpha_n \prod_{i<j} (1 - \alpha_i A_{ij} \alpha_j). \tag{12.3.51}$$

Now expanding the product, the only terms that survive the integration are those in which each variable appears only once, and therefore there are $n/2$ matrix elements of A_{ij}, with the result

$$I \, = \, \sum_{p \in S_n} \epsilon(p) A_{p(1)p(2)} \, A_{p(3)p(4)} \cdots A_{p(n-1)p(n)}, \tag{12.3.52}$$

where $\epsilon(p)$ is the sign of the permutation p. The expression above is the "Pfaffian" of the matrix A, $Pf(A)$, a quantity that was introduced in Chapter 5. It satisfies the identity

$$(Pf(A))^2 \, = \, \det A. \tag{12.3.53}$$

Partition function. Based on the previous formulas, let's now compute the partition function on a torus of the fermion system with action $\mathcal{S}$ given by (12.3.10)

$$Z \, = \, \int \mathcal{D}\psi \, \mathcal{D}\bar{\psi} \, e^{-\mathcal{S}} \, = \, (\det \partial)^{1/2} \, (\det \bar{\partial})^{1/2} \, = \, (\det \Box)^{1/2}. \tag{12.3.54}$$

It is necessary to specify the boundary conditions on the torus, along both the horizontal and vertical directions

$$\begin{aligned} \psi(z + \omega_1) \, &= \, e^{2\pi i \nu} \, \psi(z) \\ \psi(z + \omega_2) \, &= \, e^{2\pi i \mu} \, \psi(z). \end{aligned} \tag{12.3.55}$$

We assume that the same conditions are imposed for the anti-analytic component. Requiring that the action is periodic on a torus under a shift of any of its periods,

there are only four possibilities:

$$\begin{array}{ll} (\mu,\nu) = (0,0) & (P,P) \\ (\mu,\nu) = \left(0,\frac{1}{2}\right) & (P,A) \\ (\mu,\nu) = \left(\frac{1}{2},0\right) & (A,P) \\ (\mu,\nu) = \left(\frac{1}{2},\frac{1}{2}\right) & (A,A) \end{array} \tag{12.3.56}$$

where P and A denote the *periodic* and the *antiperiodic* boundary conditions, respectively. If $z_{\mu,\nu}$ denotes the functional integral on the chiral component with boundary condition of type μ along the vertical axis (the time direction) and type ν along the horizontal axis (the space direction), for the partition function we have

$$Z_{\mu,\nu} = \mid z_{\mu,\nu} \mid^2 . \tag{12.3.57}$$

The quantities $z_{\mu,\nu}$ can be easily computed in terms of the characters of the fermionic theory. Let's consider, firstly, the simplest case (A,A), in which we have

$$z_{A,A} = q^{-c/24} \operatorname{tr}_A q^{L_0} = q^{-1/48} \operatorname{tr}_A q^{L_0}, \tag{12.3.58}$$

where $q = e^{2\pi i\tau}$, $c = 1/2$, and the trace is taken in the antiperiodic sector of the theory, i.e. Neveu–Schwarz. Similarly, we have

$$z_{A,P} = \frac{1}{\sqrt{2}} q^{-1/48} \operatorname{tr}_P q^{L_0}, \tag{12.3.59}$$

where this time the trace is taken in the periodic sector, i.e. the Ramond sector (the factor $1/\sqrt{2}$ that enters the definition of this quantity is introduced to simplify later its modular properties).

It is necessary to discuss separately the case of periodic boundary conditions along the vertical axis. Since the natural boundary conditions for a fermionic field are the antiperiodic ones, to change them and make them periodic it is necessary to introduce in the trace an operator that anticommutes with the fermion. Such an operator is just $(-1)^F$, that was previously introduced, and for the remaining partition functions we then have

$$z_{P,A} = q^{-1/48} \operatorname{tr}_A (-1)^F q^{L_0} \tag{12.3.60}$$

$$z_{P,P} = q^{-1/48} \operatorname{tr}_P (-1)^F q^{L_0}. \tag{12.3.61}$$

The computation of these two traces is elementary. For each fermionic mode there are only two states and therefore, for each term of type $q^{n\psi_{-n}\psi_n}$, we have

$$\begin{aligned} \operatorname{tr} q^{n\psi_{-n}\psi_n} &= 1 + q^n \\ \operatorname{tr} (-1)^F q^{n\psi_{-n}\psi_n} &= 1 - q^n \end{aligned} \tag{12.3.62}$$

and therefore

$$\begin{aligned} \mathrm{tr}\, q^{\sum_{n>0} n\psi_{-n}\psi_n} &= \mathrm{tr} \prod_{n>0} q^{n\psi_{-n}\psi_n} = \prod_{n>0}(1+q^n) \\ \mathrm{tr}\,(-1)^F q^{\sum_{n>0} n\psi_{-n}\psi_n} &= \mathrm{tr} \prod_{n>0}(-1)^F q^{n\psi_{-n}\psi_n} = \prod_{n>0}(1-q^n). \end{aligned} \tag{12.3.63}$$

Using the expression for L_0 given in (12.3.46), in the two antiperiodic (NS) cases (with half-integer mode expansion) and in the periodic (R) case (with integer mode expansion), we get

$$\begin{aligned} z_{AA}(\tau) &= q^{-1/48}\mathrm{tr}_A\, q^{L_0} = q^{-1/48} \prod_{n=0}^{\infty}(1+q^{n+1/2}) = \sqrt{\frac{\theta_3(\tau)}{\eta}} \\ z_{PA}(\tau) &= q^{-1/48}\mathrm{tr}_A(-1)^F\, q^{L_0} = q^{-1/48} \prod_{n=0}^{\infty}(1-q^{n+1/2}) = \sqrt{\frac{\theta_4(\tau)}{\eta}} \\ z_{AP}(\tau) &= \frac{1}{\sqrt{2}}\, q^{-1/48}\mathrm{tr}_P\, q^{L_0} = q^{1/24} \prod_{n=0}^{\infty}(1+q^n) = \sqrt{\frac{\theta_2(\tau)}{\eta}} \\ z_{PP}(\tau) &= \frac{1}{\sqrt{2}}\, q^{-1/48}\mathrm{tr}_P(-1)^F\, q^{L_0} = q^{1/24} \prod_{n=0}^{\infty}(1-q^n) = 0 \end{aligned}$$

where $\theta_i(\tau)$ are the Jacobi functions defined in Problem 2. Note that the partition function z_{PP} vanishes, for the zero mode present with these boundary conditions and the integration rules (12.3.49). Under the modular transformation $\tau \to \tau + 1$, the partition functions change as

$$\begin{aligned} z_{AA}(\tau+1) &= e^{-i\pi/24}\, z_{PA}(\tau) \\ z_{PA}(\tau+1) &= e^{-i\pi/24}\, z_{AA}(\tau) \\ z_{AP}(\tau+1) &= e^{i\pi/12}\, z_{AP}(\tau) \end{aligned} \tag{12.3.64}$$

while under $\tau \to -1/\tau$

$$\begin{aligned} z_{AA}(-1/\tau) &= z_{PA}(\tau) \\ z_{PA}(-1/\tau) &= z_{AP}(\tau) \\ z_{AP}(-1\tau) &= z_{PA}(\tau). \end{aligned} \tag{12.3.65}$$

In light of these transformations, the modular invariant partition function is obtained by including all the three boundary conditions, namely

$$Z = |\, z_{AA}\,|^2 + |\, z_{AP}\,|^2 + |\, z_{PA}\,|^2 = \left|\frac{\theta_2}{\eta}\right| + \left|\frac{\theta_3}{\eta}\right| + \left|\frac{\theta_4}{\eta}\right|. \tag{12.3.66}$$

In Chapter 14 we will show that this partition function corresponds to the square of the partition function of the Ising model.

12.4 Bosonization

As we have seen in Appendix B of Chapter 1, in a system with one-dimensional space there is no distinction between the statistical and the interaction properties of the particles. The term "bosonization" refers to the possibility of describing a relativistic theory of Dirac fermions in $(1+1)$ dimensions in terms of a bosonic theory. Such a possibility permits in many cases a drastic simplification of the original fermionic theory.

The original idea of this transformation is due to D.C. Mattis and E.H. Lieb, who were able to exactly solve in this way the Thirring model. An important step forward in condensed matter physics was achieved by A. Luther and I. Peschel. In quantum field theory, the most famous work is due to Sidney Coleman, who proved the equivalence of the Sine–Gordon and the massive Thirring model.

In this section we present the main formulas of the dictionary that links the fermionic and bosonic fields. Note that the equivalence of these two theories is also suggested by the common value of their central charge, $c = 1$.

12.4.1 Bosonization Rules

The two-point correlation functions of the two components of the complex fermion field $\Psi(z, \bar{z})$ defined in (12.3.8) are given by eqn (12.3.7). Given the free nature of the theory, the multipoint correlators are computed in terms of Wick's theorem. Focusing attention on the analytic part of Ψ we have

$$\langle \chi^\dagger(z_1) \dots \chi^\dagger(z_n)\, \chi(w_1) \dots \chi(w_n) \rangle \;=\; \det\left(\frac{1}{z_i - w_j}\right). \tag{12.4.1}$$

With the choice $g = 1/4\pi$, the propagators of the bosonic field are $\langle \phi(z_1)\phi(z_2)\rangle = -\ln z_{12}$ and $\langle \bar{\phi}(\bar{z}_1)\bar{\phi}(\bar{z}_2)\rangle = -\ln \bar{z}_{12}$. Let's consider the purely analytic vertex operators $V_{+1} =: e^{i\phi(z)} :$ and $V_{-1} =: e^{-i\phi(z)}$, together with those purely anti-analytic $\bar{V}_{+1} =: e^{i\bar{\phi}(\bar{z})} :$ and $\bar{V}_{-1}(\bar{z}) =: e^{-i\bar{\phi}(\bar{z})} :$. Using these expressions and Wick's theorem, it is easy to prove that

$$\langle e^{i\phi(z_1)} \dots e^{i\phi(z_n)}\, e^{-i\phi(w_1)} \dots e^{-i\phi(w_n)} \rangle \;=\; \frac{\prod_{i<j}(z_i - z_j)(w_i - w_j)}{\prod_{i,j}(z_i - w_j)}. \tag{12.4.2}$$

At first sight, this expression seems different from the correlation functions of the fermion fields. However, there is a mathematical identity, due to Cauchy, that states the identity of the two expressions! Namely,

$$\frac{\prod_{i<j}(z_i - z_j)(w_i - w_j)}{\prod_{i,j}(z_i - w_j)} \;=\; \det\left(\frac{1}{z_i - w_j}\right). \tag{12.4.3}$$

To prove this identity, it is sufficient to check that both expressions have the same poles and zeros. Based on this equality between the correlation functions of the fermionic

and bosonic fields, it is natural to establish the following correspondence between the two operators[3]

$$\begin{aligned}
\chi(z) &=: e^{i\phi(z)} : \\
\chi^\dagger(z) &=: e^{-i\phi(z)} : \\
\bar\chi(\bar z) &=: e^{-i\bar\phi(\bar z)} : \\
\bar\chi^\dagger(z) &=: e^{i\bar\phi(\bar z)} : .
\end{aligned} \tag{12.4.4}$$

With these expressions, one can establish the operator identities

$$\begin{aligned}
\bar\chi^\dagger(\bar z)\,\chi(z) &= \bar\Psi\frac{(1+\sigma_3)}{2}\Psi =: e^{i\varphi(z,\bar z)} : \\
\chi^\dagger(z)\,\bar\chi(\bar z) &= \bar\Psi\frac{(1-\sigma_3)}{2}\Psi =: e^{-i\varphi(z,\bar z)} : .
\end{aligned} \tag{12.4.5}$$

(σ_3 is the Pauli matrix) or, equivalently

$$\bar\Psi\,\Psi =: \cos\varphi :; \qquad \bar\Psi\,\sigma_3\,\Psi = i : \sin\varphi : . \tag{12.4.6}$$

It is necessary, however, to pay particular attention to establish the bosonic espression of the fermionic current. In fact, a naive application of the bosonization rules would lead to

$$\bar\Psi\gamma^0\Psi = \bar\Psi\sigma_1\Psi = \chi^\dagger\chi + \bar\chi^\dagger\bar\chi$$

and therefore

$$\chi^\dagger(z)\chi(z) \to e^{-i\phi(z)+i\phi(z)} = 1,$$

an identity that is clearly false. To understand the origin of the discrepancy, we must recall that the normal order exponential operators do not obey the usual additive rule of the exponentials. In fact,

$$: e^{i\alpha\phi(z_1)} :: e^{i\beta\phi(z_2)} := (z_1-z_2)^{\alpha\beta} : e^{i\alpha\phi(z_1)+i\beta\phi(z_2)} : . \tag{12.4.7}$$

Taking this relation into account, it is convenient to compute the fermionic current corresponding to two points slightly separated, so that

$$\begin{aligned}
\chi^\dagger(z_1)\chi(z_2) &= (z_1-z_2)^{-1} : e^{-i(\phi(z_1)-\phi(z_2))} : \\
&= \frac{1}{z_1-z_2} - i\partial_z\phi + \cdots
\end{aligned} \tag{12.4.8}$$

If we omit the first term of this equation (which corresponds to the identity operator present in the OPE of the two fermionic fields), this is equivalent to defining the normal order of the operators as

$$: \chi^\dagger(z)\,\chi(z) := \lim_{\eta\to 0}\left[\chi^\dagger(z+\eta)\,\chi(z) - \langle\,\chi^\dagger(z+\eta)\,\chi(z)\,\rangle\right].$$

[3]For $g \neq 1/4\pi$ the formulas change as follows: $\chi \to e^{i\sqrt{4\pi g}\phi}$ and similarly for the others. Strictly speaking, when there are several fermions, it is necessary to update the formulas above by introducing the so-called *Klein factors* that implement the correct anticommutation relations of the fermion fields. To simplify the discussion we neglect here this aspect of the problem.

We arrive then at the operator identity

$$: \chi^\dagger(z)\,\chi(z) :=\, -i\,\partial_z\phi(z). \tag{12.4.9}$$

Repeating the same considerations for the anti-analytic part, we get an equivalent formula (with a change of sign)

$$: \bar\chi^\dagger(\bar z)\,\bar\chi(\bar z) :=\, i\,\partial_{\bar z}\bar\phi(\bar z). \tag{12.4.10}$$

Going back to euclidean coordinates (x_0, x_1), we have

$$\begin{aligned} j^0 =:\, \bar\Psi\sigma_1\Psi :=\, & -i\,(\partial_z\phi(z) - \partial_{\bar z}\bar\phi(\bar z)) \\ = & -i\,[(\partial_z - \partial_{\bar z})\phi + (\partial_z - \partial_{\bar z})\bar\phi] \\ = & -\partial_{x_1}\varphi(x_0, x_1) \end{aligned} \tag{12.4.11}$$

and similarly

$$j^1 =:\, \bar\Psi\sigma_2\Psi :=\, \partial_{x_0}\varphi \tag{12.4.12}$$

where σ_2 is the Pauli matrix. It is easy to prove that this current is conserved $\partial_\mu j^\mu = 0$, and its bosonization expression is

$$:\, \bar\Psi\gamma^\mu\Psi :=\, -\epsilon^{\mu\nu}\,\partial_\nu\varphi. \tag{12.4.13}$$

Other useful bosonization formulas are discussed in Problem 6 and 7 at the end of the chapter. A summary of the bosonization rules is given in Table 12.1.

Table 12.1: *Bosonization formulas.*

Bosonic theory	Fermionic theory
$\mathcal{A}_B = \frac{g}{2}\int d^2(\partial\varphi)^2$	$\mathcal{A}_F = \int d^2x[\bar\chi^\dagger\partial_z\chi + \chi^\dagger\partial_{\bar z}\chi]$
$\varphi(z,\bar z) = \phi(z) + \bar\phi(\bar z)$	$\Psi(z,\bar z) = \begin{pmatrix}\chi(z)\\ \bar\chi(\bar z)\end{pmatrix}$
$\theta(z,\bar z) = \phi(z) - \bar\phi(\bar z)$	
$: e^{i\sqrt{4\pi g}\phi} :$	χ
$: e^{-i\sqrt{4\pi g}\phi} :$	$\chi^\dagger$
$: e^{-i\sqrt{4\pi g}\bar\phi} :$	$\bar\chi$
$: e^{i\sqrt{4\pi g}\bar\phi} :$	$\bar\chi^\dagger$
$-i\sqrt{4\pi g}\,\partial_z\phi$	$: \chi^\dagger\chi(z) :$
$i\sqrt{4\pi g}\,\partial_{\bar z}\bar\phi$	$: \bar\chi^\dagger\bar\chi(z) :$
$: \cos\sqrt{4\pi g}\varphi :$	$: \bar\Psi\,\Psi :$
$i : \sin\sqrt{4\pi g}\varphi :$	$: \bar\Psi\sigma_3\,\Psi :$

References and Further Reading

The partition functions of free conformal theories on a torus are studied in the papers:

C. Itzykson, J.B. Zuber, *Two-dimensional conformal invariant theories on a torus*, Nucl. Phys. B 275 (1986), 580.

P. Ginsparg, *Curiosity at c =1* , Nucl. Phys. B 295 (1988), 153.

The papers where the bosonization approach has been developed are:

D. Mattis, E.H. Lieb, *Exact solution of a many fermion system and its associated boson field*, J. Math. Phys. 6 (1965), 304.

A. Luther, I. Peschel, *Calculation of critical exponents in two-dimensions from quantum field theory in onedimension*, Phys. Rev. B 12 (1975), 3908.

S. Coleman, *The quantum Sine–Gordon equation as the massive Thirring model*, Phys. Rev. D 11 (1975), 2088.

The bosonization approach and its applications are discussed in detail in the books:

A. Gogolin, A. Nersesyan, A. Tsvelik, *Bosonization and Strongly Correlated Systems*, Cambridge University Press, Cambridge, 1998.

T. Giamarchi, *Quantum Physics in One Dimension*, Oxford University Press, Oxford, 2004.

Problems

1. Vertex operators

In order to prove that the ground states $\mid \alpha\rangle$ of the free bosonic theory are obtained by applying the vertex operators $V_\alpha(z,\bar{0})$ to the conformal vacuum $\mid 0\rangle$ it is necessary to show that $V_\alpha(0,0)$ is an eigenstate of π_0 with eigenvalue α and, furthermore, that $a_n\, V_\alpha(0,0) \mid 0\rangle \,=\, 0$, with $n > 0$. To this end:

1. Prove the validity of the formula

$$[B, e^A] \,=\, e^A\, [B, A]$$

assuming that the commutator $[B, A]$ is a constant.

2. Impose $B = \pi_0$, $A = i\alpha\varphi(z,\bar{z})$, and show that

$$[\pi_0, V_\alpha] \,=\, \alpha\, V_\alpha.$$

Consequently

$$\pi_0\, V_\alpha(0,0) \mid 0\rangle \,=\, \alpha\, V_\alpha(0,0) \mid 0\rangle.$$

3. Show that

$$[a_n, V_\alpha(z,\bar{z})] \,=\, \alpha\, z^n\, V_\alpha(z,\bar{z})$$

and conclude that a_n annihilates $V_\alpha(0,0)\ |\,0\rangle$.

2. Modular transformations of the $\theta_i(\tau)$ functions

Let's denote by $\theta_i(z,\tau)$ the Jacobi theta functions, with $\theta_1(\tau) = 0$, while the others have the infinite series and product representations

$$\begin{aligned}
\theta_2(\tau) &= \sum_{n\in\mathbf{Z}} q^{(n+1/2)^2/2} \,=\, 2\,q^{1/8}\,\prod_{n=1}^{\infty}(1-q^n)\,(1+q^n)^2 \\
\theta_3(\tau) &= \sum_{n\in\mathbf{Z}} q^{n^2/2} \qquad\quad =\, \prod_{n=1}^{\infty}(1-q^n)\,(1+q^{n+1/2})^2 \\
\theta_4(\tau) &= \sum_{n\in\mathbf{Z}} (-1)^n\, q^{n^2/2} \,=\, \prod_{n=1}^{\infty}(1-q^n)\,(1-q^{n-1/2})^2.
\end{aligned}$$

Use these expressions and the Poisson resummation formula (12.2.55) to prove

$$\begin{aligned}
\theta_2(\tau+1) &= e^{i\pi/4}\,\theta_2(\tau), \quad & \theta_2(-1/\tau) &= \sqrt{-i\tau}\,\theta_4(\tau) \\
\theta_3(\tau+1) &= \theta_4(\tau), & \theta_3(-1/\tau) &= \sqrt{-i\tau}\,\theta_3(\tau) \\
\theta_4(\tau+1) &= \theta_3(\tau), & \theta_4(-1/\tau) &= \sqrt{-i\tau}\,\theta_2(\tau).
\end{aligned}$$

3. Dedekind function

The Dedekind function is defined by the infinite product ($q = e^{2i\pi\tau}$)

$$\eta(\tau) \,=\, q^{1/24}\,\prod_{n=1}^{\infty}(1-q^n).$$

1. Use the definition of the $\theta_i(\tau)$ functions in terms of the infinite product to prove the identity

$$\eta^3(\tau) \,=\, \frac{1}{2}\theta_2(\tau)\,\cdot\,\theta_3(\tau)\,\theta_4(\tau).$$

2. Use the modular transformations of $\theta_i(\tau)$ to prove

$$\begin{aligned}
\eta(\tau+1) &= e^{i\pi/12}\,\eta(\tau) \\
\eta(-1/\tau) &= \sqrt{-i\tau}\,\eta(\tau).
\end{aligned}$$

4. Bosonic partition function at the self-dual point

Consider the expression (12.2.56) of the partition function of a bosonic field with a compactification radius R equal to the self-dual value $R = \sqrt{2}$:

$$Z(\sqrt{2}) \;=\; \frac{1}{|\eta(\tau)|^2} \sum_{n,m\in\mathbf{Z}} q^{\frac{1}{4}(n+m)^2}\, \bar{q}^{\frac{1}{4}(n-m)^2}.$$

Prove that this expression can be written as

$$Z(\sqrt{2}) \;=\; |C_0|^2 + |C_1|^2$$

where

$$C_0(\tau) \;=\; \frac{1}{\eta} \sum_{k\in\mathbf{Z}} q^{m^2} \;=\; \frac{\theta_3(2\tau)}{\eta(\tau)}$$

$$C_1(\tau) \;=\; \frac{1}{\eta} \sum_{k\in\mathbf{Z}} q^{(m+1/2)^2} \;=\; \frac{\theta_2(2\tau)}{\eta(\tau)}.$$

5. Jacobi identity

The aim of this exercise is to prove, by physical arguments, the Jacobi identity

$$\prod_{n=1}^{\infty} (1-q^n)(1+q^{n-1/2}w)(1+q^{n-1/2}w^{-1}) \;=\; \sum_{n=-\infty}^{\infty} q^{n^2/2}\, w^n$$

which holds for $|q| < 1$ and $w \neq 0$. Consider then the partition function of a free system of f fermions and $\bar{f}$ antifermions, with energy levels $E = E_0(n - \frac{1}{2})$, $n \in \mathbf{Z}$, and total fermion number $N = N_f - N_{\bar{f}}$. Let $q = e^{-E_0/T}$ and $w = e^{\mu/T}$.

1. Show that the grand canonical partition function is given by

$$\begin{aligned} Z(w,q) = \sum_{f,\bar{f}} e^{-E/T+\mu N/T} \;&=\; \sum_{N=-\infty}^{\infty} w^N\, Z_N(q) \qquad (12.4.14)\\ &= \prod_{n=1}^{\infty} (1+q^{n-1/2}w)(1+q^{n-1/2}\,w^{-1}) \end{aligned}$$

   where $Z_N(q)$ is the partition function at a given number N of the fermions.

2. Consider now Z_0. The lowest energy states that contribute to this quantity have all negative energy levels occupied (they form the Dirac sea, with a total energy normalized to the value $E = 0$) whereas the excited states are described by the integers $k_1 \geq k_2 \geq k_3 \geq \cdots \geq k_l > 0$ with $\sum_i^l k_i = M$. The energy of these states is $E = ME_0$. Prove that Z_0 is given by

$$Z_0 \;=\; \sum_{M=0}^{\infty} P(M)\, q^M \;=\; \prod_{n=1}^{\infty} \frac{1}{1-q^n}$$

   where $P(M)$ is the combinatoric function that expresses in how many ways an integer M is expressed as a sum of numbers smaller than it.

3. Consider now the sector with fermionic number N, where the first positive levels are occupied. Argue that this sector contributes with the factor

$$q^{1/2} \cdots q^{N-3/2}\, q^{N-1/2} \;=\; q^{\sum_{n=1}^{N}(j-1/2)} \;=\; q^{N^2/2}$$

in their partition function, while the remaining excitation gives rise to the same partition function Z_0, so that

$$Z_N \;=\; q^{N^2/2}\, Z_0.$$

4. Now use eqn (12.4.14) to prove the Jacobi identity.

6. Quantum Pythagoras's theorem

A regularization of the normal order $:A(x)B(x):$ of two operators can be obtained by the limit $\lim_{\eta\to 0} :A(x-\eta/2)B(x+\eta/2):$ and an average on all directions of η, so that the final expression is invariant under the rotations. The average is equivalent to the substitution $\eta^\mu\eta^\nu/|\eta|^2 \to \frac{1}{2}\delta^{\mu\nu}$.

Use this regularization and the bosonization formulas of the text to prove the quantum version of the Pythagoras's theorem

$$(:\cos\varphi:)^2 + (:\sin\varphi:)^2 \;=\; -\frac{1}{4}\,(\partial\varphi)^2$$

(Observe that $(:\cos\varphi:)^2 \neq :\cos^2\varphi:$.)

7. Equivalence of the Sine–Gordon and Thirring models

Consider the Sine–Gordon model of a scalar bosonic field φ, whose lagrangian is

$$\mathcal{L} \;=\; \frac{1}{2}(\partial\varphi)^2 + \frac{m^2}{\beta^2}\,(\cos\beta\varphi - 1).$$

Use the bosonization formulas to prove that this lagrangian can be transformed into the lagrangian of the Thirring model

$$\mathcal{L} \;=\; i\bar\Psi\gamma^\mu\,\partial_\mu\Psi - M\,\bar\Psi\,\Psi - \frac{1}{2}g\,(\bar\Psi\gamma^\mu\Psi)\,(\bar\Psi\gamma_\mu\Psi)$$

where Ψ is a complex fermionic field, with the coupling constants related as

$$\frac{\beta^2}{4\pi} \;=\; \frac{1}{1+\frac{g}{\pi}}.$$

Note that $\beta^2 = 4\pi$ is equivalent to $g = 0$, i.e. a free fermionic model!

13
Conformal Field Theories with Extended Symmetries

Ideas are incredibly similar when you have a chance to know them.

Samuel Beckett

13.1 Introduction

This chapter deals with those field theories that present, in addition to conformal invariance, a symmetry under a larger group of transformations. These models can have interesting applications in a wide range of topics, such as the study of fundamental interactions, statistical mechanics, and condensed matter.

Our first example will be the superconformal models that have, in addition to the Virasoro generators, also their fermionic partners. The minimal models of these theories have a finite number of conformal families and rational values of the central charge and conformal weights. As in the pure bosonic case, the fusion rules of the unitary superconformal minimal models admit a remarkable interpretation in terms of Landau–Ginzburg theories. We will also study the conformal models that are invariant under the discrete $\mathbf{Z}_N$ symmetry, the so-called parafermion models. Finally, our study will focus on the conformal theories invariant under a current algebra based on a Lie group G and their lagrangian realization provided by the Wess–Zumino–Witten model.

A conformal theory is usually formulated in terms of an associative algebra that involves mutually local fields. However it is also useful to consider theories that have non-local fields. This is the case for both the superconformal and parafermion models. It is therefore convenient to define here the concept of non-local fields and refer to it later on: a field $\mathcal{O}_1(x)$ is γ-local with respect to another field $\mathcal{O}_2(x_2)$ if their product $\mathcal{O}(x_1)\,\mathcal{O}(x_2)$ acquires a phase $\exp(2\pi i\gamma)$ when the variable x_1 is analytically continued clockwise along a closed contour that encloses the point x_2, see Fig. 13.1.

13.2 Superconformal Models

In this section we present the main properties of conformal theories in which there is also a supersymmetry, i.e. a symmetry that links the bosonic and fermionic fields. They are a generalization of the conformal theories previously encountered. Since any supersymmetric theory is also superconformal on short scales, the classification

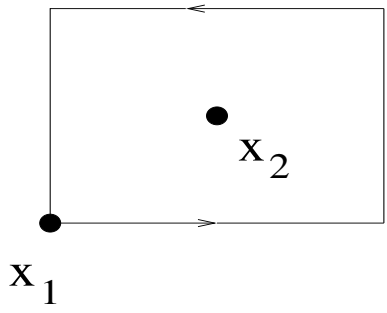


Fig. 13.1 *A closed loop of the variable x_1 around the point x_2.*

of the superconformal fixed points gives us useful information on the realization of all possible supersymmetric theories. Here we focus our attention only on the two-dimensional supersymmetric theories, referring the reader to the texts suggested at the end of the chapter for a broader discussion of the supersymmetric theories and their application in various fields of physics.

In two dimensions, superconformal invariance is associated to two supercurrents, $G(z)$ and $\bar{G}(\bar{z})$, the former a purely analytic field while the latter is a purely anti-analytic one. They are both fermionic fields, with conformal weights $(\frac{3}{2}, 0)$ and $(0, \frac{3}{2})$, respectively. The algebra of these generators is defined by the singular terms of their OPE: for $G(z)$ we have

$$G(z_1)\, G(z_2) \;=\; \frac{2c}{3(z_1 - z_2)^3} + \frac{2}{z_1 - z_2} T(z_2) + \cdots \tag{13.2.1}$$

with an analogous expression for $\bar{G}$. The parameter c is the central charge, the same quantity that enters the operator expansion of $T(z)$

$$T(z_1)T(z_2) \;=\; \frac{c}{2(z_1 - z_2)^4} + \frac{2}{(z_1 - z_2)^2} T(z_2) + \frac{1}{z_1 - z_2} \partial T(z_2) + \cdots \tag{13.2.2}$$

The field $G(z)$ (and $\bar{G}$) is itself a primary field, with operator product expansion

$$T(z_1)G(z_2) \;=\; \frac{3}{2(z_1 - z_2)^2} G(z_2) + \frac{1}{z_1 - z_2} \partial G(z_2) + \cdots \tag{13.2.3}$$

Let's define the generators L_n and G_n through the expansions

$$T(z) \;=\; \sum_{n=-\infty}^{\infty} \frac{L_n}{z^{2+n}}; \qquad G(z) \;=\; \sum_{m=-\infty}^{\infty} \frac{G_m}{z^{3/2+m}} \tag{13.2.4}$$

namely

$$L_n \;=\; \oint_C \frac{dz}{2\pi i} z^{n+1} T(z); \qquad G_m(z) \;=\; \oint_C \frac{dz}{2\pi i} z^{m+1/2} G(z).$$

Note that, in the expansion of the field $G(z)$, the indices can assume either integer or half-integer value. In fact, $G(z)$ is a fermionic field and, as we have seen in the

previous chapter for the free fermionic field ψ, is defined on the double covering of the plane, with a branch cut starting from the origin: making the analytic continuation $z \to e^{2\pi i}\, z$, we can have two possible boundary conditions

$$G(e^{2\pi i}\, z) \;=\; \pm G(z). \tag{13.2.5}$$

In the periodic case (relative to $+$), called the Neveu–Schwarz (NS) sector, the indices m are half-integers, $m \in \mathbf{Z} + \frac{1}{2}$. In the anti-periodic case (relative to $-$), called the Ramond (R) sector, the indices m are instead integer numbers, $m \in \mathbf{Z}$.

The OPE that involve $T(z)$ and $G(z)$ can be equivalently expressed as algebraic relations of their modes. Exchanging the order of the integration contours and taking into account the singular terms of their expansion (see Section 10.7), we arrive at the infinite-dimensional algebra

$$\begin{aligned}
[L_n, L_m] &= (n-m)L_{n+m} + \frac{c}{12}\, n(n^2-1)\delta_{n+m,0} \\
[L_n, G_m] &= \frac{1}{2}\,(n-2m)\, G_{n+m} \\
\{G_n, G_m\} &= 2L_{n+m} + \frac{c}{3}\left(n^2 - \frac{1}{4}\right)\, \delta_{n+m,0}.
\end{aligned} \tag{13.2.6}$$

The peculiar aspect of this algebra is the simultaneous presence of commutation and anticommutation relations.

As in the pure conformal case, the classification of superconformal theories reduces to finding all irreducible representations of the algebra (13.2.6) with the central charge c as a free parameter. The space $\mathcal{A}$ of these representations is given by the direct sum of the Neveu–Schwarz and Ramond subspaces:

$$\mathcal{A} = \mathcal{A}_{NS} \,\oplus\, \mathcal{A}_R.$$

Furthermore, each of the subspaces is decomposed into the direct sum of the *superconformal families*

$$\mathcal{A}_{NS} \;=\; \oplus_l [\Phi_l]_{NS}; \qquad \mathcal{A}_R \;=\; \oplus_\lambda [\Phi_\lambda]_R, \tag{13.2.7}$$

where the *primary fields* Φ_l and Φ_λ of this algebra satisfy

$$\begin{aligned}
L_n \Phi_a \;&=\; 0 \qquad\quad n > 0 \\
L_0 \Phi_a \;&=\; \Delta_a\, \Phi_a \\
G_m \Phi_a \;&=\; 0 \qquad\quad m > 0.
\end{aligned} \tag{13.2.8}$$

As for the Virasoro algebra, the representations are built starting from the primary fields and applying to them the *creation operators* L_n and G_m, with $n, m < 0$. So, the representations are uniquely identified by the conformal weights Δ_a of the primary fields. The same considerations hold for the anti-analytic sector of the theory.

Super-space. It is interesting to note that the operators

$$\delta_\epsilon \;=\; \oint_C \frac{dz}{2\pi i} \epsilon(z)\, T(z); \qquad \delta_\omega \;=\; \oint_C \frac{dz}{2\pi i} \omega(z)\, G(z) \tag{13.2.9}$$

can be interpreted as the (holomorphic) generators of the infinitesimal change of the coordinates $(Z, \bar{Z}) \;=\; (z, \theta; \bar{z}, \bar{\theta})$ of a $2+2$ dimensional *super-space*, where z and $\bar{z}$

are the usual complex coordinates, whereas θ and $\bar{\theta}$ are fermionic coordinates. For the analytic part of this super-space, we have the following superconformal transformation

$$z \to z + \epsilon(z) - \omega(z)\theta; \qquad \theta \to \theta + \frac{1}{2}\epsilon'(z) + \omega(z). \tag{13.2.10}$$

Hence, $\epsilon(z)$ and $\omega(z)$ are the bosonic and fermionic infinitesimal transformatiosn respectively. The peculiar nature of (13.2.10) consists of being the conformal transformation of the 1-form $dz + \theta\, d\theta$. It is therefore convenient to consider $G(z)$ and $T(z)$ as the components of a *super-stress–energy tensor*

$$W(z,\theta) \;=\; G(z) + \theta\, T(z). \tag{13.2.11}$$

Neveu–Schwarz sector. In the NS sector the representations are given in terms of the superfields

$$\Phi_l(Z,\bar{Z}) \;=\; \Phi_l(z,\bar{z}) + \theta\,\psi_l(z,\bar{z}) + \bar{\theta}\,\bar{\psi}_l(z,\bar{z}) + i\theta\,\bar{\theta}\,\tilde{\Phi}_l(z,\bar{z}) \tag{13.2.12}$$

where the primary field Φ_l is the first component while $\psi_l = G_{-1/2}\Phi_l$, $\bar{\psi}_l = \bar{G}_{-1/2}\Phi_l$ and $\tilde{\Phi}_l = -iG_{-1/2}\bar{G}_{-1/2}\Phi_l$.

Ramond sector. In the Ramond sector the field $G_m A_r$ cannot be local with respect to the fields of $\mathcal{A}_R$ and consequently the space $\mathcal{A}_R$ naturally decomposes into two locality classes: $\mathcal{A}_R = \mathcal{A}_R^{(+)} \oplus \mathcal{A}_R^{(-)}$, where all fields are mutually local in each class while any field $A_R^{(+)}$ is semilocal (with a semilocal index equal to $1/2$) with respect to $A_R^{(-)}$. The operators G_m act in $\mathcal{A}_R$ as $G_m : \mathcal{A}_R^{(\epsilon)} \to \mathcal{A}_R^{(-\epsilon)}$, with $\epsilon = \pm$. This implies, in particular, that the primary fields in the Ramond sector are organized in a doublet of fields $\Phi_\lambda \in \mathcal{A}_R^{(\epsilon)}$, with the operators G_0 and $\bar{G}_0$ that act on them as 2×2 matrices. From the algebraic relations of the modes, we also have

$$G_0^2 \;=\; L_0 - \frac{c}{24}. \tag{13.2.13}$$

Hence, for a scalar field Φ_λ with conformal weights $(\Delta_\lambda, \bar{\Delta}_\lambda)$ we get

$$G_0\,\Phi_\lambda^{(\epsilon)} \;=\; 2^{-3/2}\,(1 + i\epsilon)\,\beta_\lambda\,\Phi_\lambda^{(-\epsilon)}; \qquad \bar{G}_0\,\Phi_\lambda^{(\epsilon)} \;=\; 2^{-3/2}\,(1 - i\epsilon)\,\beta_\lambda\,\Phi_\lambda^{(-\epsilon)}$$

where $\tilde{c} = 2/3c$ and the parameter β subjected to the condition

$$\Delta_\lambda - \frac{\tilde{c}}{16} \;=\; \frac{1}{4}\beta_\lambda^2.$$

The only exception to these transformation laws is given by the Ramond field $\Phi_{(0)}$ of conformal weight $\Delta_{(0)} = \tilde{c}/16 = c/24$, if such a field actually exists in the theory: in this case, in fact, $G_0\Phi_{(0)} = \bar{G}_0\Phi_{(0)} = 0$ and therefore the second component is not necessarily present.

Irreducible representations and minimal models. The irreducible representations of the superconformal algebra are determined in the same way as those of

the Virasoro algebra previously discussed. In this case, the conformal weights can be expressed similarly to (11.2.7), namely

$$\Delta_{r,s} = \Delta_0 + \frac{1}{4}(r\beta_+ + s\beta_-)^2 + \frac{1}{32}[1-(-1)^{r+s}], \tag{13.2.14}$$

where

$$\begin{aligned} &\Delta_0 = (\tilde{c}-1)/16 \\ &\beta_\pm = \frac{1}{4}\left(\sqrt{1-\tilde{c}} \pm \sqrt{9-\tilde{c}}\right); \qquad \beta_+\,\beta_- = -\frac{1}{2}. \end{aligned} \tag{13.2.15}$$

In this formula r and s are two natural numbers: for the NS fields, $r+s \in 2\mathbf{Z}$, whereas for the Ramond fields $r+s \in 2\mathbf{Z}+1$. These degenerate fields have similar properties to the usual degenerate conformal fields, namely their operator product expansion enters only degenerate fields. Similarly, their correlation functions satisfy linear diffential equations. When the parameter $\rho = -\beta_-/\beta_+$ becomes a rational number, the operator algebra closes within a *finite number* of superconformal families. Particularly interesting are the unitary superconformal series, here denoted by $\mathcal{SM}_p$ ($p = 3, 4, 5, \ldots$), with

$$\rho = \frac{p}{p+2}.$$

In this case there are $[p^2/2]$ primary fields $\Phi_{r,s}$, where the indices r and s assume the values $r = 1, 2, \ldots, (p-1)$; $s = 1, 2, \ldots (p+1)$ (where $[x]$ is the integer part of x). The central charge and the conformal weights take the discrete values

$$\begin{aligned} &c = \frac{3}{2}\left[1 - \frac{8}{p(p+2)}\right], \quad p = 3, 4, \ldots \\ &\Delta_{r,s} = \frac{[(p+2)r - ps]^2 - 4}{8p(p+2)} + \frac{1}{32}[1-(-1)^{r+s}]. \end{aligned} \tag{13.2.16}$$

The modified Coulomb gas method can be generalized to the superconformal model, both in the Neveu–Schwarz and Ramond sectors, and permits us to determine the exact values of the structure constants of the operator algebra. It is interesting to note that, in the Ramond sector, the representation of the conformal fields can be implemented in terms of the magnetization operator σ of the Ising model, as will be discussed in detail in the next chapter.

Additional symmetry. The operator algebra of the minimal models $\mathcal{SM}_p$ may present additional symmetry, according to the value of p. In fact, if $p \in 2\,\mathbf{Z}+1$, the spaces $\mathcal{A}_R^{(+)}$ and $\mathcal{A}_R^{(-)}$ are isomorphic and therefore, for these values of p, the models are invariant under the duality transformation $\mathcal{A}_R^{(+)} \to \mathcal{A}_R^{(-)}$, similarly to the Kramers–Wannier duality of the Ising model. If p is instead an even number, the model $\mathcal{SM}_p$ contains the vacuum field $\Phi_{\frac{p}{2},\frac{p}{2}+1}$ of the Ramond sector and therefore it is not invariant under duality. However, it has a symmetry $Z_2 \times Z_2$, espressed in the form

$$\Phi_{r,s} \to (\epsilon_1)^{r+1}\,(\epsilon_2)^{s+1}\,\Phi_{r,s}$$

where the parameters $\epsilon_{1,2}$ can be either ± 1.

Landau–Ginzburg theory. Using arguments that are similar to those presented in the previous chapter, it can be shown that the unitary superconformal models $\mathcal{S}M_p$ are associated to a supersymmetric Landau–Ginzburg theory. The superpotential relative to the minimal models is given by $W(\Phi) = g\Phi^p$ and the action reads

$$\mathcal{A} \,=\, \int d^2x\, d^2\theta \,\left[-\frac{1}{2} D\Phi\, \bar{D}\Phi + W(\Phi)\right] \tag{13.2.17}$$

where

$$D \,=\, \partial_\theta - \theta\, \partial_z \qquad \bar{D} \,=\, \partial_{\bar{z}} - \bar{\theta}\, \partial_{\bar{z}}$$

are the covariant derivatives, θ and $\bar{\theta}$ are fermionic variables, while $\Phi(z, \bar{z}, \theta, \bar{\theta})$ is a superfield

$$\Phi(z, \bar{z}, \theta, \bar{\theta}) \,=\, \varphi + \theta\, \psi + \bar{\theta}\, \bar{\psi} + i\, \theta\bar{\theta}\, \chi.$$

The integration over the fermionic variables θ and $\bar{\theta}$ is done according to the rules of fermionic calculus presented in Section 12.3 of the previous chapter.

Identifying also in this case the NS superconformal primary field that sits in the position $(2, 2)$ of the Kac table with Φ, i.e. $\Phi_{2,2} \equiv \Phi$, and using the fusion rules of the superconformal minimal model, one can recursively define the composite operators $:\Phi^k:$ and show that their fusion rules lead to the operator identity

$$D\, \bar{D}\, \Phi \,\simeq\, \Phi^{p-1}. \tag{13.2.18}$$

This formula coincides with the equation of motion that can be derived by the supersymmetric action (13.2.17).

As for the minimal models of the Virasoro algebra, also for the superconformal minimal models we can determine the exact expression of the modular invariant partition functions on a torus. On this topic, we refer the reader to the original work by A. Cappelli, quoted at the end of the chapter.

The series of superconformal minimal models has an intersection with the Virasoro minimal models: in the next chapter we will see that the model $\mathcal{S}\mathcal{M}_3$ describes the tricritical Ising model, which coincides with the second minimal model of the Virasoro unitary series. The supersymmetry of this model provides a different interpretation of the primary fields and gives a reason for the particular relationships that exist among the structure constants of the conformal model. Furthermore, notice that the second minimal superconformal model has central charge $c = 1$ and can be regarded as a particular realization of the gaussian free bosonic theory analyzed in Chapter 12. It is worth stressing that supersymmetry, so long searched for in particle accelerators, has found its first physical realization in statistical mechanics!

13.3 Parafermion Models

Non-local operators naturally appear in field theories associated to the continuum limit of lattice statistical models with a $\mathbf{Z}_N$ symmetry. These theories have been investigated in detail by V. Fateev and A. Zamolodchikov. $\mathbf{Z}_N$ is an abelian group, generated by the powers of the generator Ω, and its elements are given by $\Omega, \Omega^2, \ldots, \Omega^{N-1}$, with $\Omega^N = 1$. In these models, the order parameter has $(N-1)$ components, here denoted

by σ_k, $k = 1, 2, \ldots, (N-1)$: they are scalar fields, with $\sigma_k^\dagger = \sigma_{N-k}$, and conformal weights $d_k = d_{N-k}$. These fields form a representation of $\mathbf{Z}_N$ and satisfy

$$\Omega\, \sigma_k \,=\, \omega^k\, \sigma_k, \qquad \omega = \exp(2\pi i/N). \tag{13.3.1}$$

Statistical models that are invariant under a $\mathbf{Z}_N$ symmetry can also be invariant under duality. For the self-dual theories, in addition to the $(N-1)$ order parameters, there are other $(N-1)$ operators μ_l ($l = 1, 2, \ldots, N-1$), with $\mu_l^\dagger = \mu_{N-l}$. These are the disorder operators, with the same conformal weights as the order parameters, $d_l = d_{N-l}$. The fields μ_l and σ_k are mutually local among themselves, but are non-local with respect to each other: the semilocal parameter of the fields σ_k and μ_l is equal to $\gamma_{kl} = kl/N$. The disorder fields form a representation of the dual group $\tilde{\mathbf{Z}}_N$, generated by $\tilde{\Omega}$ and satisfy

$$\tilde{\Omega}\, \mu_l \,=\, \omega^l\, \mu_l. \tag{13.3.2}$$

In light of this operator content, the self-dual models possess an enlarged symmetry $\mathbf{Z}_N \times \tilde{\mathbf{Z}}_N$. This allows us to introduce the concept of *charge*: we say that a field $\mathcal{O}_{(k,l)}$ has a charge (k, l) with respect to the group $\mathbf{Z}_N \times \tilde{\mathbf{Z}}_N$ if

$$\Omega \mathcal{O}_{(k,l)} \,=\, \omega^k\, \mathcal{O}_{(k,l)}, \qquad \tilde{\Omega} \mathcal{O}_{(k,l)} \,=\, \omega^l\, \mathcal{O}_{(k,l)}$$

with the integers k and l that are defined modulo N. Under an OPE, there is an abelian composition law for these fields, given (up to the actual value of the structure constants) by

$$\mathcal{O}^{(i)}_{(k,l)}\, \mathcal{O}^{(j)}_{(k',l')} \,=\, \sum_k \mathcal{O}^{(k)}_{(k+k',l+l')}, \tag{13.3.3}$$

where the sums over the indices are modulo N. With the definition given above, the fields σ_k have charge $(k, 0)$ while μ_l have charge $(0, l)$. In general, the semilocal index of two fields $\mathcal{O}_{(k,l)}$ and $\mathcal{O}_{(k',l')}$ is equal to $\gamma = (kl' + k'l)/N$. In addition to the symmetry $\mathbf{Z}_N \times \tilde{\mathbf{Z}}_N$, we also assume that these theories are invariant under the charge conjugation $\mathcal{C}$ and parity $\mathcal{P}$ transformations, with

$$\begin{array}{l} \mathcal{C} : \sigma_k \to \sigma_k^\dagger;\ \mu_l \to \mu_l^\dagger; \\ \mathcal{P} : \sigma_k \to \sigma_k;\ \mu_l \to \mu_l. \end{array} \tag{13.3.4}$$

In the next chapter we will see that the simplest representative of these theories is provided by the Ising model, invariant under the group $\mathbf{Z}_2 \times \tilde{\mathbf{Z}}_2$. In the operator content of this theory there is a Majorana fermion, whose analytic and anti-analytic components are $\psi(z)$ and $\bar{\psi}(\bar{z})$, respectively, which appear in the short-distance expansion of the order and disorder parameters

$$\sigma(z, \bar{z})\, \mu(0, 0) \,=\, \frac{1}{\sqrt{2}}\, (z\bar{z})^{-1/2}\, \left[z^{1/2}\, \psi(0) + \bar{z}^{1/2} \bar{\psi} + \cdots \right]. \tag{13.3.5}$$

These fields satisfy the analyticity and anti-analyticity conditions $\partial_{\bar{z}} \psi = \partial_z \bar{\psi} = 0$. We can now generalize these formulas to the case $\mathbf{Z}_N$: for the operator product expansion

of $\sigma_k(x)\,\mu_k(0)$ and $\sigma_k(x)\,\mu_k^\dagger(0)$ we impose

$$\begin{aligned}\sigma_k(z,\bar z)\,\mu_k(0,0) &= z^{\Delta_k-2d_k}\,\bar z^{\bar\Delta_k-2d_k}\,\psi_k(0,0)+\cdots\\ \sigma_k(z,\bar z)\,\mu_k^\dagger(0,0) &= z^{\bar\Delta_k-2d_k}\,\bar z^{\Delta_k-2d_k}\,\bar\psi_k(0,0)+\cdots\end{aligned}\tag{13.3.6}$$

where we have also used the symmetry (13.3.4). The fields ψ_k and $\bar\psi_k$ are operators with conformal weights Δ_k and $\bar\Delta_k$. From the semilocality of the operators σ_k and μ_k we can easily derive the condition

$$\Delta_k-\bar\Delta_k \;=\; -\frac{k^2}{N}\quad(\mathrm{mod}\,\mathbf{Z}).\tag{13.3.7}$$

Let's assume that in the self-dual critical theory the condition $\bar\Delta_k = 0$ holds. The fields ψ_k and $\bar\psi_k$ satisfy

$$\partial_{\bar z}\psi_k \;=\; 0;\quad \partial_z\bar\psi_k \;=\; 0,\tag{13.3.8}$$

so that $\psi_k = \psi_k(z)$ and $\bar\psi_k = \bar\psi_k(\bar z)$. In this case the conformal weights Δ_k coincide with the spins of the fields and their general expression is then

$$\Delta_k \;=\; m_k-\frac{k^2}{N},\tag{13.3.9}$$

where m_k are integer numbers. The operators ψ_k and $\bar\psi_k$ have charge equal to (k,k) and $(k,-k)$ respectively, and they are semilocal to each other. In contrast with the scalar order and disorder fields previously introduced, these fields have spins and therefore it is natural to call them *parafermions*. The simplest expression for (13.3.9) that also satisfies the condition $\Delta_k = \Delta_{N-k}$ is provided by

$$\Delta_k \;=\; \frac{k(N-k)}{N}.\tag{13.3.10}$$

In the following we assume that these are the conformal weights of the parafermions. The fields ψ_k generate a closed operatorial algebra

$$\begin{aligned}\psi_k(z_1)\,\psi_l(z_2) &= \mathbf{C}_{k,l}\,(z_{12})^{-2kl/N}\,\psi_{k+l}(z_2)+\cdots\\ \psi_k(z_1)\,\psi_k^\dagger(z_2) &= (z_{12})^{-2\Delta_k}\left[\mathbf{1}+\frac{2\Delta_k}{c}z_{12}^2\,T(z_2)+\cdots\right]\end{aligned}\tag{13.3.11}$$

where $T(z)$ is the analytic component of the stress–energy tensor, $\mathbf{C}_{k,l}$ are the structure constants of this algebra, whereas c is the central charge. These parameters can be fixed by imposing the associativity of this algebra. This condition leads to the values of the structure constants

$$\mathbf{C}_{k,l} \;=\; \frac{\Gamma(k+l+1)\,\Gamma(N-k+1)\,\Gamma(N-l+1)}{\Gamma(k+1)\,\Gamma(l+1)\,\Gamma(N-k-l+1)\,\Gamma(N+1)},\tag{13.3.12}$$

and the central charge

$$c \;=\; \frac{2(N-1)}{N+2}.\tag{13.3.13}$$

As for the Virasoro and the superconformal agebras, the fields of the self-dual $\mathbf{Z}_N \times \tilde{\mathbf{Z}}_N$ can be classified by the irreducible representation of the parafermionic algebra (13.3.11). Their Hilbert space is decomposed into *parafermionic conformal families*

$$\mathcal{A} \,=\, \oplus_{k=0}^{N-1}\, [\sigma_k]_\psi \,, \tag{13.3.14}$$

whose primary operators are the order parameters σ_k. Their conformal weights can be obtained by expressing $T(z)$ in terms of the normal order of the fields ψ_k and $\psi_k^\dagger$, and then using the operator product expansion (13.3.6). As a result we have

$$d_k \,=\, \frac{k(N-k)}{N(N+2)}. \tag{13.3.15}$$

In the next section we will show that the parafermionic theories also naturally appear in the Kac–Moody algebra $SU(2)_N$. In particular, using the results relative to this theory we can easily derive all the other conformal data of the parafermionic models. For instance, for the structure constants that enter the operator product expansion

$$\sigma_{k_1}(z,\bar z)\sigma_{k_2}(0,0) \,=\, \mathbf{C}_{k_1,k_2}(z\bar z)^{2d_{k_1+k_2}-d_{k_1}-d_{k_2}}\,\sigma_{k_1+k_2} + \dots$$

with the operators normalized as

$$\langle \sigma_k(z,\bar z)\sigma_{k'}^\dagger\rangle \,=\, \delta_{k,k'}\,(z\bar z)^{-2d_k},$$

one has

$$\mathbf{C}_{k_1,k_2} \,=\, \frac{\Gamma\left(\frac{1}{N+2}\right)\Gamma\left(\frac{1+k_1+k_2}{N+2}\right)\Gamma\left(\frac{N-k_1+1}{N+2}\right)\Gamma\left(\frac{N-k_2+1}{N+2}\right)}{\Gamma\left(\frac{N+1}{N+2}\right)\Gamma\left(\frac{N-k_1-k_2+1}{N+2}\right)\Gamma\left(\frac{k_1+1}{N+2}\right)\Gamma\left(\frac{k_2+1}{N+2}\right)}. \tag{13.3.16}$$

These quantities can be extracted by the four-point correlation functions of the σ_k operators. The simplest of them is given by

$$\langle \sigma_1(z_1,\bar z_1)\sigma_1^\dagger(z_2,\bar z_2)\sigma_k(z_3,\bar z_3)\sigma_k^\dagger(z_4,\bar z_4)\rangle \,=\, (z_{12}\bar z_{12})^{-2d_1}\,(z_{34}\bar z_{34})^{-2d_k}\,\mathcal{G}_{1,k}(x,\bar x),$$

where x and $\bar x$ are the harmonic ratios

$$x \,=\, \frac{z_{12}z_{24}}{z_{14}z_{23}}, \qquad \bar x \,=\, \frac{\bar z_{12}\bar z_{24}}{\bar z_{14}\bar z_{23}},$$

and the function $\mathcal{G}_(1,k)(x,\bar x)$ is expressed by

$$\begin{aligned}\mathcal{G}_{1,k}(x,\bar x) = (x\bar x)^{-k/N(N+2)}\,&\frac{\Gamma\left(\frac{1}{N+2}\right)\Gamma\left(\frac{N}{N+2}\right)}{\Gamma\left(\frac{N+1}{N+2}\right)\Gamma\left(\frac{2}{N+2}\right)}\\ \times\Bigg[&\frac{\Gamma\left(\frac{k+2}{N+2}\right)\Gamma\left(\frac{N-k+1}{N+2}\right)}{\Gamma\left(\frac{N-k}{N+2}\right)\Gamma\left(\frac{k+1}{N+2}\right)}\,F^{(1)}(k,x)F^{(1)}(k,\bar x)\\ +\,&\frac{\Gamma\left(1-\frac{k}{N+2}\right)\Gamma\left(\frac{k+1}{N+2}\right)}{\Gamma\left(\frac{k}{N+2}\right)\Gamma\left(1-\frac{k+1}{N+2}\right)}\,\frac{(x\bar x)^{(N+1-k)/(N+2)}}{(N+1-k)^2}\,F^{(2)}(k,x)F^{(2)}(k,\bar x)\Bigg].\end{aligned} \tag{13.3.17}$$

In this formula $F^{(i)}$ are the hypergeometric functions

$$F^{(1)}(k,x) = F\left(\frac{k}{N+2}, -\frac{1}{N+2}, \frac{k+1}{N+2}\,;\, x\right);$$
$$F^{(2)}(k,x) = F\left(\frac{N+1}{N+2}, \frac{N-k}{N+2}, \frac{2N-k+3}{N+2}\,;\, x\right).$$

Similarly one can also obtain the exact expression of the correlators that involves the order and disorder operators, the simplest example being

$$\langle \mu_1(z_1,\bar z_1)\mu_1^\dagger(z_2,\bar z_2)\sigma_k(z_3,\bar z_3)\sigma_k^\dagger(z_4,\bar z_4)\rangle \;=\; (z_{12}\bar z_{12})^{-2d_1}\,(z_{34}\bar z_{34})^{-2d_k}\,\mathcal{H}_{1,k}(x,\bar x),$$

where $\mathcal{H}_(1,k)(x,\bar x)$ is given by

$$\begin{aligned}\mathcal{H}_{1,k}(x,\bar x) = \bar x^{k/N}\,(x\bar x)^{-k/N(N+2)}\,&\frac{\Gamma\left(1+\frac{1}{N+2}\right)\Gamma\left(\frac{N}{N+2}\right)}{\Gamma\left(\frac{N+1}{N+2}\right)\Gamma\left(\frac{2}{N+2}\right)}\\ \times\Bigg[&\frac{\Gamma\left(\frac{k+2}{N+2}\right)\Gamma\left(\frac{N-k+1}{N+2}\right)}{\Gamma\left(\frac{N-k}{N+2}\right)\Gamma\left(1+\frac{k+1}{N+2}\right)}\,F^{(1)}(k,x)F^{(2)}(N-k,\bar x)\\ +\;&\frac{\Gamma\left(1-\frac{k}{N+2}\right)\Gamma\left(\frac{k+1}{N+2}\right)}{\Gamma\left(\frac{k}{N+2}\right)\Gamma\left(1+\frac{N+k+1}{N+2}\right)}\,x\,(x\bar x)^{-(k+1)/(N+2)}\,F^{(1)}(k,x)F^{(2)}(N-k,\bar x)\Bigg].\end{aligned} \tag{13.3.18}$$

This expression clearly shows that moving the point $(z_2,\bar z_2)$ along a closed contour that encloses the point $(z_3,\bar z_3)$, the correlation function acquires a phase factor, related to the non-locality of the two operators.

13.3.1 Relation with Lattice Models

The formulas of the previous section provide the exact solution of the quantum field theories of the critical points with a $\mathbf{Z}_N$ symmetry. It is useful to investigate their relation with the exactly solvable theories defined on a lattice that share the same symmetry. These theories are defined in terms of the variables σ_r, defined on any site r of the lattice, that take values ω^q, $q = 0, 1, \ldots (N-1)$. Assuming that their interaction is restricted to nearest neighbors, the partition function can be written as

$$Z \;=\; \sum_{\{\sigma_r\}} e^{-\sum_{r,a=1,2} H(\sigma_r,\sigma_{r+e_a})} \;=\; \sum_{\{\sigma_r\}}\prod_{r,a} W(\sigma_r,\sigma_{r+e_a}), \tag{13.3.19}$$

where e_a are the basis vectors of the lattice. The hamiltonian must be invariant under the $\mathbf{Z}_N$ transformations and the charge conjugation $\mathcal{C}$

$$H(\omega\sigma,\omega\sigma') \;=\; H(\sigma,\sigma') \;=\; H(\sigma^\dagger,\sigma^{\dagger,\prime}). \tag{13.3.20}$$

Consequently, the Boltzmann weights $W(\sigma, \sigma')$ can be written as

$$W(\sigma, \sigma') \;=\; e^{-H(\sigma,\sigma')} \;=\; \sum_{k=0}^{N-1} w_k \, (\sigma^\dagger \sigma')^k, \tag{13.3.21}$$

where the real and positive parameters w_k satisfy the condition $w_k = w_{N-k}$. As normalization we will choose $w_0 = 1$. Hence, such models are parameterized by the parameters w_k, with $k = 1, 2, \ldots \leq [N/2]$, where $[x]$ is the integer part of the number x.

The duality transformation of these lattice models can be performed as discussed in Chapter 4: the spins σ_r are replaced by the *dual spins* μ_l, associated to the sites of the dual lattice, with their interaction described by the same type of formulas shown in (13.3.19) and (13.3.21), where the dual parameters $\tilde{w}_k$ are expressed in terms of the original parameters w_i as

$$\tilde{w}_k \;=\; \left(1 + \sum_{q=1}^{N-1} w_q \, \omega^{kq}\right) \left(1 + \sum_{q=1}^{N-1} w_q\right)^{-1}. \tag{13.3.22}$$

The system is then self-dual if it satisfies the conditions

$$\tilde{w}_k \;=\; w_k, \quad k \;=\; 1, 2, \ldots (N-1). \tag{13.3.23}$$

For $N = 2, 3$, these lattice models coincide with the Ising and the three-state Potts models, respectively. Equation (13.3.23) identifies in both cases their critical temperature. For $N = 4$, the corresponding model is a special case of the Ashkin–Teller model. The self-dual line is described by

$$w_2 + 2w_1 \;=\; 1, \tag{13.3.24}$$

and the exact solution of this model can be found in the book by Baxter.[1] Its phase diagram is shown in Fig. 13.2. There are three phases, according to the values of the parameters: phase I, where $\langle\sigma\rangle \neq 0$ and $\langle\mu\rangle = 0$; phase II, where $\langle\sigma\rangle = 0$ and $\langle\mu\rangle \neq 0$; finally phase III, where $\langle\sigma\rangle = \langle\mu\rangle = 0$. The points of the segment AB of the diagram, that belong to the line (13.3.24), are all critical points of the system and therefore the critical exponents vary continuously along AB. There is however a peculiar point C, identified by the equations

$$w_1 \;=\; \frac{\sin(\pi/16)}{\sin(3\pi/16)}, \qquad w_2 = 1 - 2w_1, \tag{13.3.25}$$

where it is possible to show that the corresponding critical theory is precisely given by the parafermionic conformal theory $\mathbf{Z}_4 \times \tilde{\mathbf{Z}}_4$ previously analyzed.

[1] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, New York, 1982.

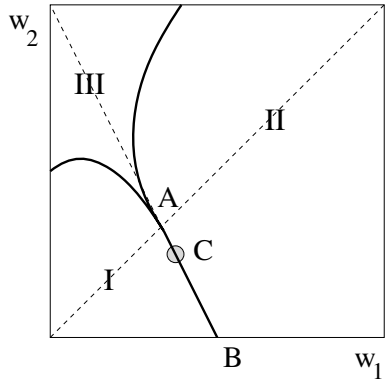


Fig. 13.2 *Phase diagram of the* $\mathbf{Z}_4$ *lattice model.*

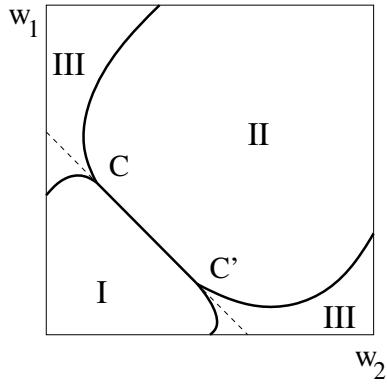


Fig. 13.3 *Phase diagram of the lattice* $\mathbf{Z}_5$ *model.*

Similarly, for a lattice model with $\mathbf{Z}_5$ symmetry, one has the phase diagram shown in Fig. 13.3. Also in this case there are three distinct phases, with the same characterization used for the previous $\mathbf{Z}_4$ model. The critical line is given by

$$w_1 + w_2 \,=\, \frac{1}{2}(\sqrt{5}-1). \tag{13.3.26}$$

This line contains, in particular, two symmetric bifurcation points C and C', whose corresponding theory in the continuum can be shown to coincide with the parafermionic conformal theory $\mathbf{Z}_5 \times \tilde{\mathbf{Z}}_5$.

In general, the points of the critical lines of the self-dual models described by the parafermionic theory have been identified by V. Fateev and A. Zamolodchikov. They correspond to the values

$$w_k \,=\, \prod_{l=0}^{k-1} \frac{\sin \pi l/N + \pi/4N)}{\sin(\pi(l+1)/N - \pi/4N)}. \tag{13.3.27}$$

13.4 Kac–Moody Algebra

In this section we consider the conformal field theories characterized by a set of analytic currents $J^a(z)$ of conformal weights $(\Delta, \bar{\Delta}) = (1, 0)$ and an analogous set of anti-analytic currents $\bar{J}^a(\bar{z})$ of conformal weights $(0, 1)$. As usual, we focus our attention on the analytic sector, with similar results for the anti-analytic one. Conformal theories based on a current algebra prove to be an important tool in the development of both string theory and condensed matter physics. Moreover, they give rise to one of the most general realizations of conformal field theory: the minimal models previously discussed are in fact a particular cases of them.

Let's start our discussion with the OPE of the currents. For dimensional reasons, this can be written as

$$J^a(z_1)\, J^b(z_2) \;=\; \frac{\tilde{k}^{ab}}{(z_1 - z_2)^2} + \frac{i f^{abc}}{z_1 - z_2}\, J^c(z_2) + \cdots \tag{13.4.1}$$

where, in the last term, it is meant to be a sum over the index c. The structure constants f^{abc} are obviously antisymmetric in the indices a and b. For the associativity of this operator expansion, they satisfy the Jacobi identity

$$\sum_d \left(f^{ade}\, f^{bcd} + f^{cde}\, f^{abd} + f^{bde}\, f^{cad} \right) \;=\; 0. \tag{13.4.2}$$

Therefore these quantities also play the role of the structure constants of a Lie algebra[2] $\mathcal{G}$. In the following we assume that this algebra is associated to a compact Lie group, characterized by a positive definite Cartan matrix. In this case the indices a, b, etc., run over the values $1, \ldots, |G| = \dim G$. In the algebra $\mathcal{G}$ it is always possibile to choose a basis such that

$$\tilde{k}^{ab} \;=\; \tilde{k}\, \delta^{ab}. \tag{13.4.3}$$

The algebra (13.4.1), defined by the operator expansion of the currents, is called the *affine algebra* or *Kac–Moody algebra*. Expanding the currents in modes, for instance at the origin

$$J^a(z) \;=\; \sum_{n=-\infty}^{\infty} \frac{J^a_n}{z^{n+1}}, \tag{13.4.4}$$

we can translate the operator expansion (13.4.1) into the commutation relations of the modes

$$\left[J^a_m, J^b_n \right] \;=\; i\, f^{abc}\, J^c_{m+n} + \tilde{k}\, m\, \delta^{ab}\, \delta_{m+n,0}. \tag{13.4.5}$$

Note that the zero modes of the currents, J^a_0, give rise to the usual commutation relations of the generators of the Lie algebras.

[2]The basic properties of the Lie algebras are summarized in the appendix at the end of the chapter.

The representation theory of the affine algebras can be developed along the lines of the Virasoro algebra. Also in this case, it is possible to define a vacuum state $|\,0\,\rangle$, annihilated by all positive modes of the currents

$$J^a_n\,|\,0\,\rangle \;=\; 0 \qquad n \;\geq\; 0. \tag{13.4.6}$$

There is also the notion of *primary field* $\varphi^l_{(r)}$, in this case made of a field multiplet, that satisfies the operator expansion

$$J^a(z_1)\,\varphi^l_{(r)}(z_2) \;=\; \frac{(R^a_{(r)})^{lk}}{z_1 - z_2}\,\varphi^k_{(r)}(z_2) + \cdots \tag{13.4.7}$$

where $(R^a_{(r)})^{lk}$ are the matries of the generators J^a, in the representation labeled by (r). The highest weight vectors of the Kac–Moody algebra are obtained acting with the primary fields on the vacuum state

$$|\,(r)\,\rangle \;=\; \varphi_{(r)}(0)\,|\,0\,\rangle. \tag{13.4.8}$$

In particular, this multiplet of states gives rise to a representation of the zero modes of the algebra, i.e. of the group G

$$J^a_0\,|(r)\,\rangle \;=\; R^a_{(r)}\,|(r)\,\rangle, \qquad J^a_n\,|(r)\,\rangle \;=\; 0 \qquad n > 0. \tag{13.4.9}$$

As for the stress–energy tensor and the primary fields of the Virasoro algebra, also in this case it is possible to prove that a Ward identity is satisfied by the currents:

$$\langle J^a(z)\varphi_{(r_1)}(z_1)\ldots\varphi_{(r_n)}(z_n)\rangle \;=\; \sum_{j=1}^{n}\frac{R^a_{(r_j)}}{z - z_j}\,\langle\varphi_{(r_1)}(z_1)\ldots\varphi_{(r_n)}(z_n)\rangle. \tag{13.4.10}$$

13.4.1 Virasoro Operators and Sugawara Formula

For the conformal field theories ruled by a set of currents it is natural to assume that the stress–energy tensor can be expressed as their composite operator. Since the conformal weight of $T(z)$ is equal to 2, while that of the currents J^a is 1, it should be possible to express $T(z)$ as a quadratic expression of J^a, invariant under the group. This reasoning leads to the *ansatz*

$$T(z) \;=\; \frac{1}{\gamma}\sum_{a=1}^{|G|} :\,J^a(z)\,J^a(z)\,:= \;\frac{1}{\gamma}\left(\lim_{w\to z}\sum_{a=1}^{|G|}J^a(w)\,J^a(z) - \frac{k\,|G|}{(w-z)^2}\right). \tag{13.4.11}$$

Expressing $T(z) = \sum L_n/z^{n+2}$, for the generators of the Virasoro algebra we have

$$L_n \;=\; \frac{1}{\gamma}\sum_{m=-\infty}^{\infty} :\,J^a_{m+n}\,J^a_{-m}\,:. \tag{13.4.12}$$

The constant γ in these formulas can be fixed demanding that the currents $J^a(z)$ are themselves primary fields with conformal weights$(1,0)$, fulfilling the OPE

$$T(z_1)J^a(z_2) \;=\; \frac{J^a(z_2)}{(z_1-z_2)^2} + \frac{\partial J^a(z_2)}{z_1-z_2} + \cdots \tag{13.4.13}$$

Note that this relation is equivalent to the commutation relations

$$[L_m, J^a_n] \;=\; -n\, J^a_{m+n}. \tag{13.4.14}$$

To determine γ, consider the expression for L_{-1} and apply this operator to a highest weight state $|\,(r)\,\rangle$. Using eqn (13.4.9), it is easy to check that in this procedure only the first term is different from zero, with the result

$$L_{-1}\,|\,(r)\,\rangle \;=\; \frac{2}{\gamma}\, J^a_{-1}\, R^a_{(r)}\,|\,(r)\,\rangle. \tag{13.4.15}$$

Applying to both terms of this expression J^b_1 and using eqn (13.4.14), we obtain

$$\begin{aligned} R^b_{(r)}\,|\,(r)\,\rangle &= \frac{2}{\gamma}\,(if^{abc}J^c_0 + \tilde{k}\delta^{ab})\, R^a_{(r)}\,|\,(r)\,\rangle \\ &= \frac{2}{\gamma}\,\left(if^{abc}\,\frac{1}{2}if^{dca}R^d_{(r)} + \tilde{k}R^b_{(r)}\right)\,|\,(r)\,\rangle \\ &= \frac{2}{\gamma}\left(\frac{1}{2}C_A + \tilde{k}\right)\, R^b_{(r)}\,|\,(r)\,\rangle, \end{aligned} \tag{13.4.16}$$

where we have defined the Casimir invariant C_A in the adjoint representation of the algebra through the formula

$$C_A\,\delta^{ab} \;=\; \sum_{c,d} f^{acd}f^{bcd}. \tag{13.4.17}$$

From (13.4.16) we arrive at the value of the constant γ

$$\gamma \;=\; 2\tilde{k} + C_A. \tag{13.4.18}$$

Once we have determined γ, we can compute the central charge of these theories by the two-point correlator of $T(z)$

$$\langle T(z_1)T(z_2)\rangle \;=\; \frac{c_G}{2}\,\frac{1}{(z_1-z_2)^4}. \tag{13.4.19}$$

Since

$$T(z) \;=\; \frac{1/2}{\tilde{k}+C_A/2}\sum_{a=1}^{|G|} :\, J^a(z)J^a(z)\,: \tag{13.4.20}$$

and

$$\langle J^a(z_1)J^b(z_2)\rangle \;=\; \frac{\tilde{k}\,\delta^{ab}}{(z_1-z_2)^2}, \tag{13.4.21}$$

this yields

$$c_G = \frac{\tilde{k}\,|G|}{\tilde{k} + C_A/2}. \tag{13.4.22}$$

In the literature the relation that links $T(z)$ to the currents J_a is known as the *Sugawara formula*.

13.4.2 Maximal Weights

In this section we discuss the representations of the Kac–Moody algebra associated to the irreducible and unitary maximal weights. These are also the representations that are irreducible for the ordinary Lie algebras and since they have the lowest eigenvalue of L_0, are called *vacuum representations*. The unitary conditions are expressed by $J^{a\dagger}(z) = J^a(z)$ and this implies $J_n^{a\dagger} = J_{-n}^a$. In the Cartan basis, the generators are given by $H^i(z)$ and $E^{\pm\alpha}(z)$, where $i = 1, \ldots, r_G$ are the indices that identify the generators that commute with each other, while the positive roots α denote the creation and annihilation operators. In this basis the highest weight states that form a vacuum representation satisfy

$$\begin{aligned} H_n^i \,|\, \lambda \,\rangle &= E_n^{\pm\alpha} \,|\, \lambda \,\rangle = 0, \quad n > 0 \\ H_0^i \,|\, \lambda \,\rangle &= \lambda^i \,|\, \lambda \,\rangle, \quad E_0^\alpha \,|\, \lambda \,\rangle = 0, \quad \alpha > 0. \end{aligned} \tag{13.4.23}$$

The remaining states are obtained acting on the state $|\, \lambda \,\rangle$ either by $E_0^{-\alpha}$ or by any mode J_{-n}^a, with $n > 0$.

The constant $\tilde{k}$ of the Kac–Moody algebra depends on the chosen normalization of the structure constants. Hence it is convenient to consider the following quantity $k = 2\tilde{k}/\psi^2$, called the *level of the affine algebra*, that is independent of the normalization of the structure constants. For the unitary conformal theories, the constant k is quantized and takes only integer value. To show this, it is convenient to consider firstly the case $G = SU(2)$. With the normalization $f^{abc} = \sqrt{2}\,\epsilon^{abc}$ and $\psi^2 = 2$, the generators are given by

$$I^\pm = \frac{1}{\sqrt{2}}(J_0^1 \pm i J_0^2) \qquad I^3 = \frac{1}{\sqrt{2}} J_0^3. \tag{13.4.24}$$

They satisfy

$$\left[I^+, I^-\right] = 2I^3, \qquad \left[I^3, I^\pm\right] = \pm I^\pm. \tag{13.4.25}$$

With the chosen normalization, the operator $2I^3$ has integer eigenvalues on any finite-dimensional representation of the group. There is, however, another set of operators that fulfill the same algebra $SU(2)$ given by

$$\begin{aligned} \tilde{I}^+ &= \frac{1}{\sqrt{2}}(J_{+1}^1 - i J_{+1}^2) \\ \tilde{I}^- &= \frac{1}{\sqrt{2}}(J_{-1}^1 + i J_{-1}^2) \\ \tilde{I}^3 &= \frac{1}{2}k - \frac{1}{\sqrt{2}} J_0^3. \end{aligned} \tag{13.4.26}$$

It is easy to show that they satisfy the relations $[\tilde{I}^+, \tilde{I}^-] = 2\tilde{I}^3$, $[\tilde{I}^3, \tilde{I}^\pm] = \pm\tilde{I}^\pm$. These commutation relations imply that also the operator $2\tilde{I}^3 = k - 2I^3$ possesses integer eigenvalues and, consequently, k is an integer number, $k \in \mathbf{Z}$.

The argument presented above is not only valid for $SU(2)$ but also for any other algebra $\mathcal{G}$. In fact, the highest root ψ always gives rise to a $SU(2)$ subalgebra, generated by

$$I^\pm = E_0^{\pm\psi}, \quad I^3 = \psi \cdot H_0/\psi^2. \tag{13.4.27}$$

This subalgebra is accompanied by another $SU(2)$ subalgebra given by

$$\tilde{I}^\pm = E_{\pm 1}^{\mp\psi} \quad \tilde{I}^3 = (\tilde{k} - \psi \cdot H_0)/\psi^2 \tag{13.4.28}$$

so that, repeating the steps of the previous argument, we arrive at the conclusion that also in this case the level $k = 2/\psi^2 = 2\tilde{I}^3 + 2I^3$ can take only integer values.

Conformal weights and constraint thereof. Let's now compute the conformal weights of the vacuum representations. Equation (13.4.12) yields

$$\begin{aligned} L_0 \,|\, (r) \,\rangle &= \frac{1/2}{\tilde{k} + C_A/2} \sum_{a,m} :\, J^a_m J^a_{-m} \,:\, |\, (r) \,\rangle \\ &= \frac{1/2}{\tilde{k} + C_A/2} \sum_a R^a_{(r)}\, R^a_{(r)} \,|\, (r) \,\rangle \;=\; \frac{C_r/2}{\tilde{k} + C_A/2} \,|\, (r) \,\rangle, \end{aligned} \tag{13.4.29}$$

where C_r is the Casimir invariant in the (r) representation. Thus, the conformal weight of the multiplet made of the primary fields $\varphi_{(r)}(z)$ is

$$\Delta_r \;=\; \frac{C_r/2}{\tilde{k} + C_A/2} \;=\; \frac{C_r/\psi^2}{k + \tilde{h}_G}, \tag{13.4.30}$$

where $\tilde{h}_G$ is the dual Coxeter number. However not all the representations can be accepted. To understand the constraint to which they are subjected, let's consider once again the $SU(2)$ case. For this algebra, the vacuum states transform in the spin j representation and therefore

$$L_0 \,|\, (j) \,\rangle \;=\; \frac{j(j+1)}{k+2} \,|\, (j) \,\rangle. \tag{13.4.31}$$

Fixing k, the only values of the spin j that can appear in this formula are those that satisfy the condition

$$2j \le k. \tag{13.4.32}$$

To this end, let's analyse in more detail the (j) representation. The $(2j+1)$ states of this representation are identified by their eigenvalue with respect to I^3, namely $I^3 |\, (j), m \,\rangle = m \,|\, (j), m \,\rangle$. Consider then the state with the maximum value of m, i.e. $m = j$, and the matrix element

$$\langle\, j \,|\, \tilde{I}^+ \tilde{I}^- \,|\, j \,\rangle \;=\; \langle\, j \,|\, \left[\tilde{I}^+, \tilde{I}^-\right] \,|\, j \,\rangle \;=\; \langle\, j \,|\, (k - 2I^3) \,|\, j \,\rangle \;=\; k - 2j \ge 0. \tag{13.4.33}$$

Hence, eqn (13.4.32) implies that for a fixed value of k, there are only $k+1$ possible values of j, given by $j = 0, \frac{1}{2}, \ldots, \frac{k}{2}$.

It is immediate to generalize the condition (13.4.32) to all other groups. Instead of $|\, j \,\rangle$, one needs to consider in the general case the state $|\, \lambda \,\rangle$, where λ is the highest weight of the vacuum representation. Using the previous argument, one arrives at the constraint

$$2\psi \cdot \lambda / \psi^2 \leq k. \tag{13.4.34}$$

This is the condition that determines the representations that can appear in the algebra at a fixed value of k.

With the identification of the primary fields of the Kac–Moody algebras, the remaining states that form the Verma modules of these theories are obtained by acting on the primary fields by the operators J^a_{-n}. As for the conformal theories with $c < 1$, these representations contain certain null-vectors, which are necessary to mode out in order to define the irreducible representations. For the affine algebras, one can show that all the null-vectors are descendents of only one primitive null-vector. For a generic affine algebra, this state is constructed using the generators (13.4.28) of the subalgebra $SU(2)$. Note, in fact, that the eigenvalues of $2\tilde{I}^3$ on the state with highest weight $|\, (r), \lambda \,\rangle$ are given by $M = k - 2\psi \cdot \lambda / \psi^2$. The set of states generated acting by subsequent powers of $\tilde{I}^-$ on $|\, (r)\, \lambda \,\rangle$ form then an irreducible and finite dimensional representation of the algebra (13.4.28). Hence M is an integer number and we have

$$(\tilde{I}^-)^{M+1} \,|\, (r)\, \lambda \,\rangle \;=\; 0. \tag{13.4.35}$$

This is precisely the primitive null-vector of the Verma module. For the group $SU(2)$ and its j representation, this condition translates into the equation

$$(J^+_{-1})^{k-2j+1} \,|\, (j), j \,\rangle \;=\; 0. \tag{13.4.36}$$

Correlation functions. Let's now address the correlation functions of the primary fields. As shown below, they satisfy a linear first-order differential equation. Consider, in fact, the Sugawara formula for the generator L_{-1} of the Virasoro algebra[3]

$$L_{-1} \;=\; \frac{1}{\tilde{k} + C_A/2} \, (J^a_{-1} J^a_0 + J^a_{-2} J^a_1 + \cdots). \tag{13.4.37}$$

Acting on a primary field, it yields

$$\left(L_{-1} - \frac{\sum_a J^a_{-1}\, R^a_{(r)}}{\tilde{k} + C_A/2} \right) \varphi_{(r)} \;=\; 0. \tag{13.4.38}$$

Consider now the Ward identity (13.4.10) and multiply both terms of this expression for $R^a_{(r_k)}$. Taking the limit $z \to z_k$ and using the operator product expansion of the currents, we arrive at the linear differential equation, called the Kniznik–Zamolodchikov

[3]Each term in the normal order product appears twice and this cancels the factor 1/2 in the formula (13.4.12).

equation

$$\left[\left(\tilde{k}+C_A/2\right)\frac{\partial}{\partial z_k}+\sum_{a,j\neq k}\frac{R^a_{(r_j)}\,R^a_{(r_k)}}{z_j-z_k}\right]\,\langle\varphi_{r_1}(z_1)\cdots\varphi_{r_n}(z_n)\rangle\;=\;0. \tag{13.4.39}$$

To obtain the final expression of the correlator it is necessary to implement the usual steps. Namely: solve this equation with the correct asymptotic expansion, together with the one relative to the anti-analytic part, and impose the monodromy invariant condition on the solutions.

13.4.3 Wess–Zumino–Witten Models

The conformal models that satisfy a Kac–Moody algebra differ from the other conformal models by an important property: they can be consistently defined by a lagrangian formalism based on a nonlinear sigma model with a topological term. The aim of this section is to present the main steps of this derivation. Consider initially the action

$$\mathcal{S}_0\;=\;\frac{1}{4\lambda^2}\int d^2x\,\mathrm{Tr}\,(\partial^\mu g^{-1}\partial_\mu g), \tag{13.4.40}$$

where λ^2 is a dimensionless positive constant. The bosonic field $g(x)$ is a matrix with values in a semisimple Lie group g. To have a real action, $g(x)$ must belong to a unitary representation of such a group. For the trace, we adopt the normalization

$$\mathrm{Tr}\,(t^a t^b)\;=\;2\,\delta^{ab}, \tag{13.4.41}$$

where t_a are the generators of the Lie algebra in the representation under consideration. Note that if g is a unitary matrix, $g^{-1}\partial_\mu g$ is an antihermitian matrix since

$$(g^{-1}\partial_\mu g)^\dagger\;=\;\partial_\mu g^{-1}\,g\;=\;-g^{-1}\partial_\mu g, \tag{13.4.42}$$

and $\partial_\mu g^{-1}\;=\;-g^{-1}\partial_\mu g g^{-1}$, where the last relation comes from the identity $\partial_\mu(gg^{-1})\;=\;0$.

Although the theory above is conformally invariant at the classical level, it is well known that this invariance is broken at the quantum level by the renormalization procedure. For the ultraviolet divergences one is forced to introduce a length-scale and therefore the $\beta(\lambda)$ function is different from zero. At the quantum level, the theory becomes asymptotically free and its spectrum is purely massive.

The breaking of conformal invariance of the action (13.4.40) at the quantum level can be directly checked by the absence of conserved currents that are purely analytic and anti-analytic. Under the variation $g\to g+\delta g$, we have

$$\delta\mathcal{S}_0\;=\;\frac{1}{2\lambda^2}\int d^2x\,\mathrm{Tr}\,\left[g^{-1}\,\delta g\,\partial^\mu(g^{-1}\partial_\mu g)\right], \tag{13.4.43}$$

and therefore we get the equations of motion

$$\partial^\mu (g^{-1}\partial_\mu g) \,=\, 0. \tag{13.4.44}$$

They can be interpreted as the conservation law of the currents

$$J_\mu \,=\, g^{-1}\partial_\mu g. \tag{13.4.45}$$

Switching to complex coordinates and introducing the notation

$$\bar{J}_z \,=\, g^{-1}\partial_z g, \qquad \bar{J}_{\bar{z}} \,=\, g^{-1}\partial_{\bar{z}} g, \tag{13.4.46}$$

we have

$$\partial_z \bar{J}_{\bar{z}} + \partial_{\bar{z}} \bar{J}_z \,=\, 0. \tag{13.4.47}$$

In order to have a separate conservation law of the two components of the currents, it is necessary that each of the two terms of this equation vanishes separately. However, this is impossible, because this would lead to some inconsistencies. In fact, assuming that $\partial_z(g^{-1}\partial_{\bar{z}} g) \,=\, 0$, one would also have

$$\partial_z \partial_{\bar{z}} g \,=\, \partial_{\bar{z}} g\, g^{-1}\, \partial_z g. \tag{13.4.48}$$

The left-hand side is clearly symmetric under the exchange $z \leftrightarrow \bar{z}$ and this would imply the identity

$$\partial_{\bar{z}} g\, g^{-1}\, \partial_z g \,=\, \partial_z g\, g^{-1}\, \partial_{\bar{z}}\, g. \tag{13.4.49}$$

However, this identity is generically false for the elements of a non-commutative group, since it would correspond to the equality $ABC = CBA$, with $A = \partial_{\bar{z}} g$, $B = g^{-1}$ and $C = \partial_z g$.

In order to have separate conservation of the analytic and anti-analytic components, the correct choice is

$$J_z \,=\, \partial_z g\, g^{-1}, \qquad J_{\bar{z}} \,=\, g^{-1}\, \partial_{\bar{z}} g. \tag{13.4.50}$$

In this case, the conservation of one quantity implies the conservation of the other

$$\partial_z (g^{-1}\partial_{\bar{z}} g) \,=\, g^{-1}\, \partial_{\bar{z}} (\partial_z g\, g^{-1})\, g. \tag{13.4.51}$$

Hence, the question is whether it is possible to modify the action (13.4.40) in such a way that the conserved currents become those defined by eqn (13.4.50) instead of those given in eqn (13.4.46). There is indeed a positive answer and the way to implement it is to use the Wess–Zumino term

$$\Gamma \,=\, -\frac{i}{24\pi} \int_B d^3y\, \epsilon^{ijk}\, \mathrm{Tr} \left(\tilde{g}^{-1} \frac{\partial \tilde{g}}{\partial y_i}\, \tilde{g}^{-1} \frac{\partial \tilde{g}}{\partial y_j}\, \tilde{g}^{-1} \frac{\partial \tilde{g}}{\partial y_k} \right). \tag{13.4.52}$$

This expression needs an explanation. Imagine the original complex plane, with the point at infinity, compactified into the Riemann sphere S. The matrix g is then a map of the surface S onto the group G. However, this map can be extended to a new map $\tilde{g}(y)$, from all the *internal* points of the three-dimensional sphere B, with boundary

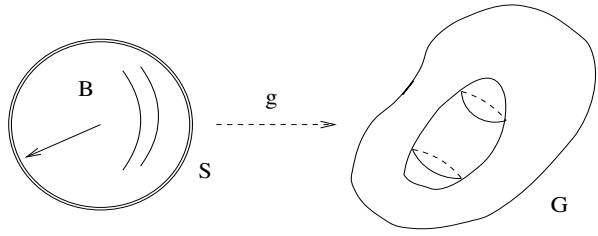


Fig. 13.4 *The map $\tilde{g}$ of the three-dimensional sphere B (with boundary given by the two-dimensional surface S) onto the group G.*

given by the surface S, onto the group G, as shown in Fig. 13.4. The new matrix $\tilde{g}$ is the one that appears in eqn (13.4.52), where the coordinates of the three-dimensional sphere are denoted by y_1, y_2 and y_3.

The Wess–Zumino term (13.4.52) has the important property of being defined up to an additive quantity that is an integer multiple of 2π. This ambiguity comes from the existence of topologically distinct ways of extending the original map g to the map $\tilde{g}$ that involve the internal points of the three-dimensional sphere. Although the expression (13.4.52) is a three-dimensional integral, its integrand is a total derivative and therefore its final value depends only on the values of $\tilde{g}$ at the boundary, i.e. on the original function g. To understand the origin of the ambiguity of Γ, let's consider the case $G = SU(2)$. The parameter space of this group is a three-dimensional sphere, whose parameterization is given by the angles ψ, θ and φ, with a line element

$$ds^2 \;=\; d\psi^2 + \sin^2\theta\,(d\theta^2 + \sin^2\theta\, d\varphi^2). \tag{13.4.53}$$

Using the parameterization of the matrix g in terms of ψ, θ, φ and the Pauli matrices

$$\begin{aligned} g \;&=\; \exp\left(\frac{i}{2}\varphi\,\sigma_3\right)\exp\left(\frac{i}{2}\theta\,\sigma_1\right)\exp\left(\frac{i}{2}\psi\,\sigma_3\right) \\ &=\; \begin{pmatrix} \cos(\theta/2)\,\exp[i(\varphi+\psi)/2] & i\sin(\theta/2)\,\exp[i(\varphi-\psi)/2] \\ i\sin(\theta/2)\,\exp[i(\psi-\varphi)/2] & \cos(\theta/2)\,\exp[-i(\varphi+\psi)/2] \end{pmatrix}, \end{aligned} \tag{13.4.54}$$

it is easy to see that the integrand in eqn (13.4.52) corresponds to the Jacobian of the transformation from the coordinates (ψ, θ, φ) to (y_1, y_2, y_3)

$$\Gamma_{SU(2)} \;=\; \frac{i}{4\pi}\int d^3y\,\frac{\partial(\psi,\theta,\varphi)}{\partial(y_1,y_2,y_3)} \;=\; \frac{i}{4\pi}\int d^2x\epsilon_{\mu\nu}\varphi\,\sin\theta\,\partial_\mu\theta\,\partial_\nu\psi \tag{13.4.55}$$

and this explicitly shows that Γ depends only on the boundary values of $\tilde{g}$. However, the result of the integration cannot be expressed in a local form in terms of g. This

matrix is in fact periodic in φ, whereas Γ is not: when φ changes of $2\pi n$, Γ changes in

$$\Delta\Gamma \;=\; i\,\frac{n}{2}\int d^2x\,\epsilon_{\mu\nu}\sin\theta\,\partial_\mu\theta\,\partial_\nu\psi. \tag{13.4.56}$$

The last integral is however an integer, since it expresses the number of times the vector field $\vec{n} = (\cos\theta, \sin\theta\,\cos\psi, \sin\theta\,\sin\psi)$ wraps round the three-dimensional sphere.

It is important to stress that the explicit result shown for $SU(2)$ also applies to all other semisimple Lie groups, by a topological theorem due to Bott. For this ambiguity of the Wess–Zumino term, the coupling constant that multiplies Γ must necessarily be an integer, here denoted by k. Hence, let's consider the new action

$$\mathcal{S} \;=\; \mathcal{S}_0 + k\,\Gamma, \tag{13.4.57}$$

and its variation under $g \to g + \delta g$. For $\delta\mathcal{S}_0$ we have the previous result (13.4.43), whereas for $\delta\Gamma$ we have

$$\delta\Gamma \;=\; \frac{i}{8\pi}\int d^2x\,\epsilon_{\mu\nu}\,\mathrm{Tr}\,(g^{-1}\delta g\,\partial^\mu(g^{-1}\partial^\nu g)). \tag{13.4.58}$$

Putting together the two terms, the equation of motion becomes

$$\partial^\mu(g^{-1}\partial_\mu g) + i\,\frac{\lambda^2 k}{4\pi}\,\epsilon_{\mu\nu}\partial^\mu(g^{-1}\partial^\nu g) \;=\; 0 \tag{13.4.59}$$

which, in complex coordinates, can be written as

$$\left(1+\frac{\lambda^2 k}{4\pi}\right)\,\partial_z(g^{-1}\partial_{\bar z}g) + \left(1-\frac{\lambda^2 k}{4\pi}\right)\,\partial_{\bar z}(g^{-1}\partial_z g) \;=\; 0. \tag{13.4.60}$$

This equation shows that, choosing

$$\lambda^2 \;=\; \frac{4\pi}{k}, \tag{13.4.61}$$

we have the desired conservation law

$$\partial_z(g^{-1}\partial_{\bar z}g) \;=\; 0. \tag{13.4.62}$$

Since λ^2 is a positive quantity, the integer k is positive as well. Choosing the other solution, $\lambda^2 = -4\pi/k$ with $k < 0$, we obtain instead the conservation of the dual current, $\partial_{\bar z}(g^{-1}\partial_z g) = 0$. With this choice of the coupling constant, the solution of the equation of motion assumes the factorized form

$$g(z,\bar z) \;=\; h(z)\bar h(\bar z), \tag{13.4.63}$$

where $h(z)$ and $\bar h(\bar z)$ are two arbitrary functions. The separated conservation law of the analytic and anti-analytic components of the currents implies furthermore the

invariance of the action under the transformation

$$g(z,\bar{z}) \to \mathcal{G}(z)\, g(z,\bar{z})\, \bar{\mathcal{G}}^{-1}(\bar{z}), \tag{13.4.64}$$

where $\mathcal{G}$ and $\bar{\mathcal{G}}$ are two arbitrary matrices of the group G, in the same representation of g. For infinitesimal values, we have

$$\mathcal{G}(z) \simeq 1 + \omega(z), \qquad \bar{\mathcal{G}}(\bar{z}) \simeq 1 + \bar{\omega}(\bar{z}),$$

and

$$\delta_\omega g = \omega\, g, \qquad \delta_{\bar{\omega}} g = -\bar{\omega} g.$$

With the choice (13.4.61), the variation of the action under $g \to g + \delta_\omega g + \delta_{\bar{\omega}} g$ is given by

$$\begin{aligned} \delta\mathcal{S} &= \frac{k}{4\pi} \int d^2x\, \mathrm{Tr}\left(g^{-1}\, \delta g\, \left[\partial_z (g^{-1} \partial_{\bar{z}} g)\right]\right) \\ &= \frac{k}{2\pi} \int d^2x\, \mathrm{Tr}\left[\omega(z) \partial_{\bar{z}} (\partial_z g g^{-1}) - \bar{\omega}(\bar{z}) \partial_z (g^{-1} \partial_{\bar{z}} g)\right], \end{aligned} \tag{13.4.65}$$

which clearly vanishes after an integration by parts. Therefore, the original global symmetry $G \times G$ of the sigma model, in the presence of the Wess–Zumino term, is enhanced with the choice (13.4.61) to a local symmetry $G(z) \times G(\bar{z})$. The analytic currents

$$\begin{aligned} J(z) &\equiv -k\, J_z(z) = -k\, \partial_z g\, g^{-1}, \\ \bar{J}(\bar{z}) &\equiv k\, J_{\bar{z}}(\bar{z}) = k g^{-1}\, \partial_{\bar{z}} g, \end{aligned} \tag{13.4.66}$$

give rise to the Kac–Moody current algebra of the previous section, where k is the same integer that enters the operator product expansion (13.4.1).

This scenario can be explicitly confirmed by a perturbative computation of the β-function. For instance, for the group $SO(N)$ one gets

$$\beta(\lambda) = -\frac{\lambda^2 (N-2)}{4\pi} \left[1 - \left(\frac{\lambda^2 k}{4\pi}\right)^2\right], \tag{13.4.67}$$

and this function has a fixed point at $\lambda^2 = \left|\frac{4\pi}{k}\right|$, as shown in Fig. 13.5. At these values of the coupling constant the correlation length of the model diverges and the theory acquires a conformal symmetry described by the Kac–Moody algebra.

13.5 Conformal Models as Cosets

The conformal theories associated to the Kac–Moody algebra are useful to construct a vast class of models. The method that we are going to present here, known as the *coset approach*, is based on a simple observation. Consider a group G and one of its subgroup H. The currents associated to the original group will be generically denoted by J^a_G, while those of H by J^i_H, where the index i assumes values on the adjoint

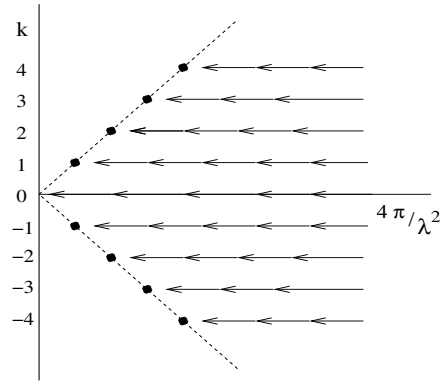


Fig. 13.5 *Renormalization group flows of the coupling constant λ^2. The strong coupling region is on the right of the graph. The coupling constant stops its growth at the fixed points of the β function, i.e. $\lambda^2 = \left|\frac{4\pi}{k}\right|$.*

representation of H, namely $i = 1, \ldots, |H|$, where $|H| = \dim H$. Using the Sugawara formula, we can construct the two stress–energy tensors associated to these groups[4]

$$T_G \;=\; \frac{1/2}{k_G + \tilde{h}_G} \sum_{a=1}^{|G|} :\, J_G^a(z) J_G^a(z) \,: \tag{13.5.1}$$

and

$$T_H \;=\; \frac{1/2}{k_H + \tilde{h}_H} \sum_{i=1}^{|H|} :\, J_H^i(z) J_H^i(z) \,:\,. \tag{13.5.2}$$

For the OPE of the currentsJ_H^i with both stress–energy tensors we have

$$\begin{aligned} T_G(z_1) J_H^i(z_2) &= \frac{J_H^i(z_2)}{(z_1 - z_2)^2} + \frac{\partial J_H^i(z_2)}{z_1 - z_2} + \cdots \\ T_H(z_1) J_H^i(z_2) &= \frac{J_H^i(z_2)}{(z_1 - z_2)^2} + \frac{\partial J_H^i(z_2)}{z_1 - z_2} + \cdots \end{aligned} \tag{13.5.3}$$

As a consequence, the operator product expansion of $(T_G - T_H)$ with J_H^i does not have singular terms. Since T_H is entirely constructed in terms of the currents J_H^i, $T_{G/H} \equiv T_G - T_H$ also has an operator expansion without singular terms with T_H. Imposing

$$T_G \;=\; (T_G - T_H) + T_H \;\equiv\; T_{G/H} + T_H, \tag{13.5.4}$$

we have an orthogonal decomposition of the original Virasoro algebra – associated to T_G – in two Virasoro algebras that commute with each other – associated to $T_{G/H}$,

[4] In the following we assume the normalization $\psi^2 = 1$.

and T_H, respectively. The central charge of the Virasoro algebra associated to $T_{G/H}$ is thus given by

$$c_{G/H} \,=\, c_G - c_H \,=\, \frac{k_G\,|G|}{k_G + \tilde{h}_G} - \frac{k_H\,|H|}{k_K + \tilde{h}_H}. \tag{13.5.5}$$

A significant class of conformal field theories is obtained by the coset $G \times G/G$, where the group G in the denominator corresponds to the diagonal subgroup of the two groups in the numerators. Denoting by $J^a_{(1)}$ and $J^a_{(2)}$ the currents in the two groups of the numerators, for those of the denominator we have $J^a = J^a_{(1)} + J^a_{(2)}$. The most singular part of their operator expansion is given by

$$J^a(z_1)J^b(z_2) \simeq J^a_{(1)}(z_1)J^b_{(1)}(z_2) + J^a_{(2)}(z_1)J^b_{(2)}(z_2) \simeq \frac{(k_1 + k_2)\,\delta^{ab}}{(z_1 - z_2)^2} + \cdots \tag{13.5.6}$$

and therefore the level of G at the denominator is $k = k_1 + k_2$. An important example of this construction is

$$G/H \,=\, SU(2)_{k-1} \times SU(2)_1/SU(2)_k. \tag{13.5.7}$$

The central charge of these theories is

$$c_{G/H} \,=\, \frac{3(k-1)}{k+1} + 1 - \frac{3k}{k+2} \,=\, 1 - \frac{6}{(k+1)(k+2)}. \tag{13.5.8}$$

Note that, with the position $q = k + 1 = 3, 4, \ldots$, these values coincide with those of eqn (11.3.2), i.e. the same central charge of the unitary minimal models of the Virasoro algebra!

Another significant example is obtained by considering

$$G/H = SU(2)_{k-1} \times SU(2)_2/SU(2)_{k+1}$$

whose central charge is

$$c_{G/H} \,=\, \frac{3(k-1)}{k+1} + \frac{3}{2} - \frac{3(k+1)}{k+3} \,=\, \frac{3}{2}\left(1 - \frac{8}{(k+1)(k+3)}\right). \tag{13.5.9}$$

These are the values of the central charge of the minimal unitary superconformal models, given in eqn (13.2.16).

Finally, let's analyze how to obtain the states of the model associated to the coset G/H. To this end, it is necessary to study the decomposition of the representations of G in the splitting (13.5.4) of the stress–energy tensors. Let $|c_G, \lambda_G\rangle$ be the representations of the affine algebra associated to G, where c_G is the central charge relative to the level k_G and λ_G is the highest weight of the vacuum representation. Since $T_G = T_{G/H} + T_H$, these representations decompose into a direct sum of the irreducible representations

$$|c_G, \lambda_G\rangle \,=\, \oplus_j \left[\, |c_{G/H}, \Delta^j_{G/H}\rangle \,\otimes\, |c_H, \lambda^j_H\rangle \,\right], \tag{13.5.10}$$

where $|c_{G/H}, \Delta^J_{G/H}\rangle$ denotes the irreducible representation of $T_{G/H}$ with the lowest eigenvalue of L_0 given by $\Delta^j_{G/H}$. Some significant examples of this formula will be discussed in the next chapter.

13.5.1 Relation with Parafermions

There is an important relationship between the Kac–Moody theories based on the group $SU(2)$ and the parafermionic models. This relationship can be established as follows. Let's initially introduce a free massless boson satisfying the equation $\partial\,\bar{\partial}\varphi = 0$

$$\varphi(z,\bar{z}) = \phi(z) + \bar{\phi}(\bar{z}),$$

with correlators

$$\begin{aligned}\langle\phi(z)\phi(0)\rangle &= 2\log z\\ \langle\bar{\phi}(z)\bar{\phi}(0)\rangle &= 2\log\bar{z}\\ \langle\phi(z)\phi(0)\rangle &= 0.\end{aligned}$$

Its stress–energy tensor $T_b(z) = (\partial\phi)^2$ generates a Virasoro algebra with central charge $c = 1$. Suppose that, in addition to this bosonic field, there are also the parafermionic fields associated to $\mathbf{Z}_N \times \tilde{\mathbf{Z}}_N$, that are decoupled by φ. In terms of the operators of both theories let's construct the currents

$$\begin{aligned}J^3(z) &= N\,\partial\phi(z),\\ J^+(z) &= N\,\psi_1(z)\,:e^{i/N^{1/2}\,\phi(z)},\\ J^-(z) &= N\,\psi^\dagger\,:e^{-i/N^{1/2}\,\phi(z)}.\end{aligned} \tag{13.5.11}$$

It is easy to check that the conformal weights of these currents are $(1,0)$: this is obvious for J^3, for the other two currents their conformal weight is given by the sum of the conformal weights of the two fields

$$\Delta_\pm = \frac{N-1}{N} + \frac{1}{N} = 1.$$

Using the operator expansion of the fields ψ_1, $\psi_1^\dagger$ and the vertex operator of the bosonic field ϕ, one can check that these currents satisfy

$$J^a(z_1)J^b(z_2) = \frac{N\,q^{ab}}{(z_1-z_2)^2} + \frac{f_c^{ab}}{z_1-z_2}J^c(z_2) + \cdots \tag{13.5.12}$$

where $q^{00} = 1/2q^{+-} = 1/2q^{-+} = 1$, whereas f_c^{ab} are the structure constants of $SU(2)$

$$\begin{aligned}f_+^{0+} &= -f_+^{+0} = -f_-^{0-} = f_-^{-0} = 1,\\ f_0^{+-} &= -f_-^{-+} = 2,\end{aligned}$$

Hence these currents give rise to a Kac–Moody algebra $SU(2)$ of level $k = N$. The stress–energy tensor of such a theory is the sum of the stress-energy tensor of the free bosonic theory and that of the parafermionic model

$$T_t(z) = T_b(z) + T_{pf}(z), \tag{13.5.13}$$

and the central charge is the sum of the central charges of the two theories

$$c_t = 1 + \frac{2(N-1)}{N+2} = \frac{3N}{N+2}. \tag{13.5.14}$$

This indeed coincides with the central charge of the Kac–Moody algebra $SU(2)$ of level $k = N$. In the light of this result, the parafermionic models $\mathbf{Z}_N$ can be considered as the

coset theory $SU(2)_N/U(1)$. This permits us to identify the fields of the parafermionic theory in terms of the decomposition of the representations of $SU(2)_N$ with respect to the subgroup $U(1)$, an observation that greatly simplifies the computation of the correlation functions of the parafermionic models.

Appendix 13A. Lie Algebra

In this appendix we recall the main results of the Lie algebra, inviting the reader to consult the literature at the end of the chapter for further analysis on the subject. First of all, for any compact Lie group with n parameters there is a Lie algebra of dimension n and vice versa. For the compact groups there are the following properties: (a) there is always a unitary representation; (b) any irreducible representation is finite-dimensional; (c) in order to find a representation of the group it is sufficient to find a representation of the algebra.

A Lie algebra $\mathcal{G}$ of dimension n is a vector space with an internal composition law given by

$$(\lambda_i, \lambda_j) \to [\lambda_i, \lambda_j] \,=\, \sum_k f_{ij}^k \, \lambda_k, \tag{13.A.1}$$

where f_{ij}^k are the structure constants of the algebra and $[\,,\,]$ is the commutator. This composition law satisfies the Jacobi identity

$$[x,[y,z]] + [z,[x,y]] + [y,[z,x]] \,=\, 0. \tag{13.A.2}$$

A representation of the Lie algebra is obtained by associating each of its elements x to a matrix $M(x)$, with the condition $M([x,y]) = [M(x), M(y)]$. Particularly important is the *adjoint representation* given by $x \to ad(x)$, where $ad(x)$ is a linear application of $\mathcal{G}$ in itself, defined by

$$ad(x)y \,=\, [x,y]. \tag{13.A.3}$$

For the Jacobi identity, we have $[ad(x), ad(y)] = ad([x,y])$. In terms of this representation we can define a bilinear form, i.e. a scalar product among the elements of the algebra, by the formula

$$\langle x|y\rangle \,=\, \mathrm{Tr}\,(ad(x)\, ad(y)). \tag{13.A.4}$$

An invariant subspace under the adjoint representation is called an *ideal* $\mathcal{I}$ of $\mathcal{G}$, namely $y \in \mathcal{I}$ if $ad(x)y = [x,y] \in \mathcal{I}$, for every $x \in \mathcal{G}$. The ideals are crucial for the further analysis of the Lie algebras. In fact, there are three classes of algebras:

1. The simple Lie algebras that have no ideals at all.
2. The semisimple Lie algebras that do not have abelian ideals.
3. All other algebras.

Presently there is a complete mathematical theory only for the first two classes. Let's now introduce another useful concept: a subalgebra $\mathcal{C}$ is a Cartan subalgebra if it has the properties: (a) $\mathcal{C}$ is a maximal abelian subalgebra, i.e. there is no other subalgebra that contains $\mathcal{C}$; (b) if $h \in \mathcal{C}$, then in any representation of $\mathcal{C}$ on a complex vector

space $A(h)$ is a diagonalizable operator. The dimension r of $\mathcal{C}$ is the *rank* of $\mathcal{G}$. Let's now recall, without giving proofs, the theory of semisimple Lie algebras.

Let $\mathcal{G}$ be an n-dimensional Lie algebra (with complex coefficients) and $\mathcal{C}$ its Cartan subalgebra of dimension r.

- Any operator $ad(h_i)$ with $h_i \in \mathcal{C}$ is diagonalizable in $\mathcal{G}$. Since $[h_i, h_j] = 0$, there exists a set of common eigenvectors $e_{\alpha_1,\ldots,\alpha_r}$, with

$$ad(h_i)e_{\alpha_1,\ldots,\alpha_r} \;=\; \alpha_i\, e_{\alpha_1,\ldots,\alpha_r}.$$

- The h_i can always be chosen (by an appropriate choice of the basis) in such a way that the eigenvalues α_i are all real. The r-dimensional vector $\alpha = (\alpha_1, \ldots, \alpha_r)$ is called a *root*. The algebra $\mathcal{G}$ can be written as a direct sum $\mathcal{G} = \mathcal{C} \oplus_a \mathcal{G}_a$, where $\mathcal{C}$ corresponds to the null root $(0, \ldots, 0)$ while $\mathcal{G}_a$ corresponds to the vector subspace associated to the non-vanishing root a. It is possible to prove that this is a one-dimensional space. Hence there are $n - r$ non-vanishing roots.
- Consider the restriction of the scalar product (13.A.4) in $\mathcal{C}$, namely

$$g_{ij} = \mathrm{Tr}\,(ad(h_i)\, ad(h_j)).$$

  In the basis $\{h_i, e_\alpha\}$, the operators $ad(h_i)$ are diagoonal and therefore $g_{ij} = \sum_\alpha \alpha_i\, \alpha_j$. Since $g_{ij} = g_{ji}$ and g_{ij} is a real matrix, it can be diagonalized. Moreover, one can show that g_{ij} is a non-singular positive definite matrix. Hence, introducing its inverse by the definition $g_{ij}g^{jk} = \delta_i^k$, we can define a scalar product among the roots

$$\langle\, \alpha|\beta\,\rangle \;=\; \sum_i \alpha^i\, \beta_i \;=\; \sum_{i,j} g^{ij}\, \alpha_i\, \beta_j. \tag{13.A.5}$$

  One can always choose a basis in which $g_{ij} = \delta_{ij}$, so that $\langle\alpha|\beta\rangle = \sum_i \alpha_i\beta_i$. As we shall see soon, in the basis $\{h_i, e_\alpha\}$ all the commutation relations of the Lie algebra are fixed by the roots.

Roots. The roots are the building blocks of the Lie algebras. They satisfy a series of properties enumerated below:

1. If α is a root, then $k\alpha$ is a root only if $k = 0, \pm 1$. Hence the $n - r$ roots come in pairs and we have $n - r = 2m$.
2. If α and β are two roots, they uniquely identify two non-negative integers p and q such that $\beta - p\alpha, \beta - (p-1)\alpha, \ldots, \beta + q\alpha$ are the only roots of the form $\beta + k\alpha$. This series of roots is called the *string* α *containing* β. Exchanging α with β, we can identify two other non-negative integers p' and q' that characterize the string β containing α. These numbers satisfy

$$p - q \;=\; 2\,\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}, \qquad p' - q' \;=\; 2\,\frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle}. \tag{13.A.6}$$

   Since

$$-p \;\leq (q - p) \;\leq q, \qquad -p' \;\leq (q' - p') \;\leq q'$$

if α and β are two non-vanishing roots we have that

$$\beta - 2\,\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle}\,\alpha, \qquad \alpha - 2\,\frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle}\,\beta,$$

are also non-vanishing roots. Note that the first is obtained by reflecting β with respect to the orthogonal plane to α, while the second is reflecting α with respect to the orthogonal plane to β.

3. Since
$$ad(h_i)[e_\alpha, e_\beta] \;=\; [h_i, [e_\alpha, e_\beta]] \;=\; (\alpha_i + \beta_i)[e_\alpha, e_\beta]$$
there are the following cases:
   (a) $\alpha + \beta \neq 0$, with $\alpha + \beta$ not a root. In this case $[e_\alpha, e_\beta] = 0$, otherwise $[e_\alpha, e_\beta]$ would be an eigenvector of $ad(h_i)$ and $\alpha + \beta$ a root.
   (b) $\alpha + \beta = 0$, in this case $[e_\alpha, e_{-\alpha}] \in \mathcal{C}$ and then it can be written as
   $$[e_\alpha, e_{-\alpha}] \;=\; \sum_i \lambda^i\, h_i. \tag{13.A.7}$$
   Choosing the normalization $\langle e_\alpha | e_{-\alpha}\rangle = 1$ (which determines the roots up to a factor d_α such that $d_\alpha d_{-\alpha} = 1$), one has $\lambda^i \;=\; \alpha^i$.
   (c) $\alpha + \beta \neq 0$, but with $\alpha + \beta$ a root. Since the space of eigenvectors is one-dimensional, one has $[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}$ and the coefficient $N_{\alpha,\beta}$ satisfies the conditions
   $$N_{\alpha,\beta} = N_{\beta,-\alpha-\beta} = N_{-\alpha-\beta,\alpha} = -N_{\beta,\alpha}. \tag{13.A.8}$$
   From the normalization condition $\langle e_\alpha | e_{-\alpha}\rangle = 1$, we can always choose d_α in such a way that $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$ and, in this case, we arrive at the condition
   $$N^2_{\alpha,\beta} \;=\; \frac{q(p+1)}{2}\langle\beta|\beta\rangle. \tag{13.A.9}$$
   This relation determines $N_{\alpha,\beta}$ up to a sign, which can be chosen to satisfy the relations (13.A.8).

In summary, all the commutation relations of the Lie algebra are encoded in the following formulas

$$\begin{aligned}
&[h_i, h_j] \;=\; 0,\\
&[h_i, e_{\pm\alpha}] \;=\; \pm\alpha_i\, e_{\pm\alpha},\\
&[e_\alpha, e_{-\alpha}] \;=\; \sum_i \alpha^i\, h_i,\\
&[e_\alpha, e_\beta] \;=\; \begin{cases} 0 & \text{if } \alpha+\beta \neq 0 \text{ and } \alpha+\beta \text{ is not a root}\\ N_{\alpha,\beta} e_{\alpha+\beta} & \text{if } \alpha+\beta \neq 0 \text{ and } \alpha+\beta \text{ is a root.}\end{cases}
\end{aligned} \tag{13.A.10}$$

As we anticipated earlier, the roots of a Lie algebra uniquely fix its structure. Hence the classification of the Lie algebras reduces to studing the vector space of dimension r that satisfies the properties discussed above.

Simple roots. A root is called *positive* if its first non-vanishing component is positive. A root is simple if: (a) it is a positive root; and (b) it cannot be written as a sum of positive roots. The simple roots have two important properties that are easy to prove: (i) if α and β are simple roots, then $\alpha - \beta$ is not a root; (ii) $\langle\alpha|\beta\rangle \leq 0$ and moreover

$$2\,\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle} = p - q \;=\; -q, \tag{13.A.11}$$

since, for the point (a), $p = 0$.

The utility of the simple roots is stated by the following theorem: there are exactly r simple roots, all linearly independent, and any other positive root can be written as their linear combination. In addition, if α is a positive root but not simple, there always exists a simple root $\alpha^{(k)}$ so that $\alpha - \alpha^{(k)}$ is a positive root. These two properties ensure that all the roots of the algebra can be determined in terms of the simple roots. From eqn (13.A.6) we infer that there are severe constraints on the angle between two roots and the ratio of their lengths. In fact, since

$$2\frac{\langle\alpha|\beta\rangle}{\langle\alpha|\alpha\rangle} \;=\; m, \quad 2\frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle} \;=\; n \tag{13.A.12}$$

one has

$$\frac{(\langle\alpha|\beta\rangle)^2}{\langle\alpha|\alpha\rangle\,\langle\beta|\beta\rangle} \;=\; \frac{mn}{4} \;=\; \cos^2\varphi_{\alpha,\beta} \leq 1 \tag{13.A.13}$$

and, if $m, n \neq 0$,

$$\frac{\langle\alpha|\alpha\rangle}{\langle\beta|\beta\rangle} \;=\; \frac{n}{m}. \tag{13.A.14}$$

If we now specialize these equations to the case in which α and β are simple roots, we have both $m, n < 0$ and there are only the following cases

m	n	φ	$\langle\alpha\|\alpha\rangle/\langle\beta\|\beta\rangle$
-1	-1	$120°$	1
-1	-2	$135°$	2
-1	-3	$150°$	3
0	0	$90°$	arbitrary

The scalar product of the simple roots defines the *Cartan matrix*

$$A_{ij} \;=\; \frac{2\langle\alpha_i|\alpha_j\rangle}{\langle\alpha_j|\alpha_j\rangle}. \tag{13.A.15}$$

The matrix elements of A_{ij} are necessarily integers and its diagonal elements are equal to 2. If the roots do not have the same length, A_{ij} is not a symmetric matrix. It is convenient to introduce a special notation for the quantity $2\alpha_i/|\alpha_i|^2$, with $|\alpha_i|^2 = \langle\alpha_i|\alpha_i\rangle$

$$\alpha_i^\vee \;=\; \frac{2\alpha_i}{|\alpha_i|^2}. \tag{13.A.16}$$

Hence the Cartan matrix can be elegantly written as $A_{ij} = \langle \alpha_i | \alpha_j^\vee \rangle$. Let's also define the *dual Coxeter number*, given by

$$\tilde{h}_G = \sum_{i=1}^{r} \alpha_i^\vee + 1. \tag{13.A.17}$$

Since any semisimple Lie algebra is the direct sum of simple algebras, it is sufficient to discuss the classification of the latter ones.

Classification of the simple Lie algebras. This problem consists of finding all sets of r simple roots that satisfy the condition discussed above, with none of them orthogonal to the others. The fundamental result of the theory can be expressed in a graphical way in terms of the Dynkin diagrams. In fact, since the length of the simple roots can take at most two values, let's associate a circle to each root. Two circles are linked by one, two, or three lines according to whether their angle is equal to 120°, 135° or 150°, respectively. If the two roots are orthogonal the relative circles are not connected. The black circles are associated to the shorter roots. The final classification of the simple Lie algebras is given in Fig. 13.6.

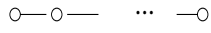
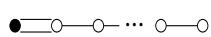
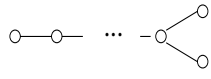
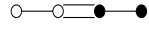
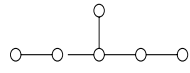
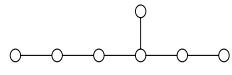
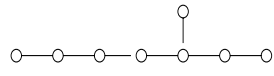
Group	Algebra	Dynkin diagram	Dimension	
SU(r+1)	A_r		r (r+2)	$r \geqslant 1$
O(2 r +1)	B_r		r (2 r +1)	$r \geqslant 2$
Sp(2 r)	C_r		r (2 r +1)	$r \geqslant 2$
O(2 r)	D_r		r (2 r −1)	$r \geqslant 3$
	G_2		14	
	F_4		52	
	E_6		78	
	E_7		133	
	E_8		248	

Fig. 13.6 *Simple Lie algebras and Dynkin diagrams.*

These algebras are all distinct when $r \geq 4$. Note that

1. When $r = 1$ there is only one Lie algebra, A_1.
2. When $r = 2$ the Dynkin diagrams of B_2 and C_2 are identical, therefore the two algebras coincide.
3. When $r = 3$ A_3 and D_3 have the same Dynkin diagram, so $A_3 = D_3$.
4. There are four families of algebras with an arbitrarily large number of simple roots: the series A_r that corresponds to the group of unitary matrices $SU(r+1)$; the series B_r, relative to group of orthogonal matrices $O(2r+1)$; the series C_r relative to the sympletic matrices $Sp(2r)$ (these are the linear transformations U that leave invariant an antisymmetric non-singular matrix I, namely $U^t\, I\, U = I$), and the series D_r that corresponds to the group of orthogonal matrices $O(2r)$. In additional to these families, there are five exceptional algebras, called G_2, F_4, E_6, E_7, and E_8.
5. Among the Lie algebras, only A_n, D_n, and the three exceptional algebras E_6, E_7, E_8 have roots all of the same length. These algebras are known as *simply laced algebras*.

Let's now discuss representation theory.

Representation theory. Let's recall that a representation to a Lie algebra on a complex vector field L is defined by a linear map $x \to T(x)$, where $x \in \mathcal{G}$ and T is an operator that acts in L, such that $T([x,y]) = [T(x), T(y)]$. In the following we only deal with the finite-dimensional representations, to which applies the Weyl theorem: any finite-dimensional representation of a semisimple Lie algebra is completely reducible. Hence we can restrict our attention only to the irreducile representations.

Choosing a basis $\{h_i, e_\alpha, e_{-\alpha}\}$ in $\mathcal{G}$, let $\{H_i, E_\alpha, E_{-\alpha}\}$ be the corresponding operators in a given representation. It is always possible to implement the conditions $H_i = H_i^\dagger$ and $E_\alpha^\dagger = E_{-\alpha}$. The H_i's are a set of hermitian operators that commute with each other. Hence they can be simultaneously diagonalized and their eigevalues are real. Let $M = (M_1, \ldots, M_r)$ be the set of eigenvalues on a common eigenvectors of the H_i

$$H_i\, |M\,\rangle \;=\; M_i\, |M\,\rangle. \tag{13.A.18}$$

M can be regarded as an r-dimensional real vector and it is called the *weight vector*. Denoting by L_M the space of eigenvectors associated to the weight M, the vector space L decomposes as

$$L \;=\; \oplus L_M. \tag{13.A.19}$$

In general, the spaces L_M are not one-dimensional and therefore the operators H_i do not form a complete set of commuting operators. Therefore some of the weights M can be degenerate. There is no general procedure to remove this degeneracy. However, it is possible to show that the number of operators that commute with all H_i that permits us to remove such a degeneracy is at most equal to $(n-3r)/2$.

Properties of the weight vectors. If $|M\,\rangle$ is a vector of L_M, from the commutation relations of H_i with E_α we get

$$H_i\, E_\alpha\, |M\,\rangle \;=\; (\alpha_i + M)\, E_\alpha\, |M\,\rangle. \tag{13.A.20}$$

Supposing that $E_\alpha|M\rangle \neq 0$, we have that also $M+\alpha = (M_1+\alpha_1, \ldots, M_r+\alpha_r)$ is a weight and $E_\alpha|M\rangle$ belongs to $L_{M+\alpha}$. If $E_\alpha E_\alpha|M\rangle \neq 0$, we can repeat the same reasoning to conclude that also $M+2\alpha$ is a weight, with $E_\alpha^2|M\rangle$ belonging to $L_{M+2\alpha}$. For recurrence, if $E_\alpha^k|M\rangle \neq 0$, then $M+k\alpha$ is a weight and $E_\alpha^k|M\rangle \in L_{M+k\alpha}$. However, since L is a finite dimensional space, such a procedure must stop, i.e. there should exist an integer q such that $E_\alpha^q|M\rangle \neq 0$ (therefore $M+q\alpha$ is a weight) but $E_\alpha^{q+1}|M\rangle = 0$. Repeating the same steps with $E_{-\alpha}$ we can determine an integer p such that $E_{-\alpha}^p|M\rangle \neq 0$ but $E_{-\alpha}^{p+1}|M\rangle = 0$. From this we derive that these vectors

$$M-p\alpha, \ldots, M+q\alpha \tag{13.A.21}$$

are all and only the weight vectors of the form $M+k\alpha$. The operators E_α and $E_{-\alpha}$ are the raising and lowering operators of the spectrum. In terms of the tensor g^{ij} we can introduce a scalar product among the weights and the roots

$$\langle M|\alpha\rangle = \sum_{ij} g^{ij}\alpha_i M_j,$$
$$\langle M|M'\rangle = \sum_{ij} g^{ij} M_i M'_j.$$

If p and q are the integers previously introduced, one has

$$\frac{2\langle M|\alpha\rangle}{|\alpha|^2} = p-q,$$

and therefore

$$M - \frac{2\langle M|\alpha\rangle}{|\alpha|^2}$$

is a weight. It is worth stressing that the considerations done above are very similar to those used for the roots – a circumstance not surprising since the roots are nothing else but the weight vectors of a particular representation, the adjoint.

Since the r simple roots $\alpha^{(i)}$ form a basis in the r-dimensional space of real vectors, any weight can be expressed in terms of them as

$$M = \sum_{i=1}^{r} M_i \alpha^{(i)}. \tag{13.A.22}$$

We can introduce an order in this space. We say that $M > M'$ if the first component of the vector $M-M'$ is positive. For a finite number of distinct weights, there exists then a *highest weight vector*, i.e. a weight that is greater than the other. As a consequence of this definition, if α is a positive root and $|\Lambda\rangle$ is an eigenvector belonging to the space of the highest weight, then $E_\alpha|\Lambda\rangle = 0$.

Let R be the representation of $\mathcal{G}$ in the linear space L, and $|\Lambda,1\rangle, |\Lambda,2\rangle, \ldots, |\Lambda,k\rangle$ a set of independent vectors belonging to the space of the highest weight Λ. Consider the subspace $L^{(1)}$ defined by the vectors

$$E_{-\alpha}E_{-\beta}\ldots|\Lambda,1\rangle, \tag{13.A.23}$$

obtained by applying a finite product of $E_{-\alpha}$ (including the repetition of the same operator) where $\alpha, \beta, \ldots$ are positive roots. It is easy to see that this is an invariant and

irreducible space. It is obviously invariant under the action of the operators H_i and $E_{-\alpha}$, while applying one of the operators E_α (with $\alpha > 0$), this can be moved, using its commutation relations, to the end of the product, where we get $E_\alpha|\Lambda, 1\rangle = 0$. Doing so, we generate a sequence of vectors having the form (13.A.23). If the representation R is irreducible we then have $R = L^{(1)}$. In $L^{(1)}$ there is only one independent vector with highest weight Λ, all other weights have the form

$$\Lambda - \sum_{\alpha>0} k_\alpha \alpha, \tag{13.A.24}$$

where k_α are integer numbers, equal to the number of times the operator $E_{-\alpha}$ appears in (13.A.23).

The importance of the concept of the highest weight is stressed by the following theorems due to Cartan. The first theorem states that two irreducible representations that have the same highest weight are equivalent. The second theorem states that an r-dimensional vector Λ is the highest weight vector of an irreducible representation if and only if

$$\Lambda_{\alpha_i} \;=\; \frac{2\langle\Lambda|\alpha^{(i)}\rangle}{|\alpha^{(i)}|^2}, \tag{13.A.25}$$

is a non-negative integer for any simple root $\alpha^{(i)}$. Hence, once we choose a set of simple roots $\alpha^{(i)}$, any set of non-negative integers $(\Lambda_{\alpha_1}, \Lambda_{\alpha_2}, \ldots, \Lambda_{\alpha_r})$ uniquely defines an irreducible representation of $\mathcal{G}$ and all representations are obtained in this way. The other weights have the form (13.A.24) and are obtained by applying the decreasing operators $E_{-\alpha}$.

Other useful formulas. In this last part of the appendix we discuss some formulas entering the formalism of the Kac–Moody algebras. The constant $C_A/2$ that appears in the expression of the stress–energy tensor and the central charge generally depends on the chosen normalization of the structure constants f^{abc}. Let $R^a_{(r)}$ the matrices of a representation (r) of G, with dimension d_r and normalization

$$\mathrm{Tr}\, R^a_{(r)}\, R^b_{(r)} \;=\; l_r\, \delta^{ab}. \tag{13.A.26}$$

Summing over the indices a and b, in the range $1, \ldots, |G|$, we get

$$C_r\, d_r \;=\; l_r\, |G| \tag{13.A.27}$$

where C_r is the quadratic Casimir operator in the representation r. Summing instead only over the indices of the Cartan subalgebra of G $(a, b = 1, \ldots, r_G)$, we get

$$\sum_{j=1}^{d_r} \mu^2_{(j)} \;=\; l_r\, r_G \tag{13.A.28}$$

where r_G is the rank of G and μ are the weights of the representation (r).

Table 13.1: *Dual Coxeter numbers of the Lie algebras.*

SU(n) $(n \geq 2)$	$\tilde{h}_{SU(n)} = n$	$l_{(n)} = \frac{1}{2}\psi^2$
SO(n) $(n \geq 4)$	$\tilde{h}_{SO(n)} = n-2$	$l_{(n)} = \psi^2$
E_6	$\tilde{h}_{E_6} = 12$	$l_{(27)} = 3\psi^2$
E_7	$\tilde{h}_{E_7} = 18$	$l_{(56)} = 6\psi^2$
E_8	$\tilde{h}_{E_8} = 30$	$l_{(248)} = 30\psi^2$
Sp(2n) $(n \geq 1)$	$\tilde{h}_{Sp(2n)} = n+1$	$l_{2n} = \frac{1}{2}\psi^2$
G_2	$\tilde{h}_{G_2} = 4$	$l_{(7)} = \psi^2$
F_4	$\tilde{h}_{F_4} = 9$	$l_{26} = 3\psi^2$

For the adjoint representation, we have $d_A = |G|$ and

$$C_A = l_{(A)} = r_G^{-1} \sum_{a=1}^{|G|} \alpha_{(a)}^2 \tag{13.A.29}$$

where α are the roots. Denoting by ψ the highest root, the quantity $\tilde{h}_G \equiv C_A/\psi^2$ is independent of the normalization and it is expressed by

$$\tilde{h}_G = \frac{C_A}{\psi^2} = \frac{1}{r_G}\left(n_L + \left(\frac{S}{L}\right)^2 n_S\right). \tag{13.A.30}$$

In this formula $n_{S,L}$ is the number of the short (long) roots of the algebra (the highest root ψ is always a long root) whereas S/L is the ratio of their lengths. As seen above, for the Lie algebras the roots can have at most two different lengths. The quantity $\tilde{h}_G$ is the *dual Coxeter number*, previously defined by the formula (13.A.17).

The simply laces algebras (A, D, E) have simple roots of the same length. The remaining algebras have roots of two different lengths and their ratio L/S is $\sqrt{2}$ for $SO(2n+1), Sp(2n)$ and F_4, while it is $\sqrt{3}$ for G_2.

We can now easily compute all dual Coxeter numbers for the compact Lie algebras; see Table 13.1.

References and Further Reading

For a general introduction to supersymmetry and its application the reader can consult:

M.F. Sonhius, *Introducing supersymmetry*, Phys. Reports 128 (1985), 39.

K. Efetov, *Supersymmetry in Disorder and Chaos*, Cambridge University Press, Cambridge, 1999.

F. Cooper, A. Khare, U. Sukhatme, *Supersimmetry in Quantum Mechanics*, World Scientific, Singapore, 2001.

The two-dimesional superconformal models are discussed in the articles:

D. Friedan, Z. Qiu, S. Shenker, *Superconformal invariance in two dimensions and the tricritical Ising model*, Phys. Lett. B 151 (1985), 37.

Z. Qiu, *Supersymmetry, two-dimensional critical phenomena and the tricritical Ising Model*, Nucl. Phys. B 270 [FS16] (1986), 205.

A. Cappelli, *Modular invariant partition functions of superconformal theories*, Phys. Lett. B 185 (1987), 82.

G. Mussardo, G. Sotkov, M. Stanishkov, *N=2 superconformal minimal models*, Int. J. Mod. Phys. A 4 (1989), 1135.

The parafermionic models and lattice statistical systems with $\mathbf{Z}_N$ are discussed in the articles:

V. Fateev, A.B. Zamolodchikov, *Parafermionic currents in the two-dimensional conformal quantum field theory and self-dual critical points in $Z(n)$ statistical systems*, Sov. Phys. JETP 62 (1985), 215; Phys. Lett. 92 A (1982), 37.

Quantum field theories with a Wess–Zumino term have been studied by many authors. A set of fundamental articles is given by:

A. Polyakov, P. Wiegman, *Goldstone fields in two dimensions with multivalued actions*, Phys. Lett. B 141 (1983), 223.

E. Witten, *Non-abelian bosonization in two dimensions*, Comm. Math. Phys. 92 (1984), 455.

Conformal models with a Kac–Moody symmetry were originally proposed in the article:

V.G. Knizhnik, A.B. Zamolodchikov, *Current algebra and Wess–Zumino model in two dimensions*, Nucl. Phys. B 247 (1984), 63.

The construction of conformal models using the coset approach is due to P. Goddard, A. Kent and D. Olive and is covered in the articles:

P. Goddard, A. Kent, D. Olive, *Virasoro algebra and coset space models*, Phys. Lett. B 152 (1985), 88.

P. Goddard, A. Kent, D. Olive, *Unitary representations of the Virasoro and super-Virasoro algebras*, Comm. Math. Phys. 103 (1986), 105.

Coset models have found remarkable applications in strongly correlated low-dimensional systems, see:

I. Affleck, *Field Theory Methods and Quantum Critical Phenomena*, Les Houches, Session XLIX, 1988, Fields, Strings and Critical Phenomena, Elsevier Science Publishers, B.V., 1989.

To deepen knowledge on group theory and Lie algebra the reader can consult the monograph:

B. Wybourne, *Classical Groups for Physicists*, John Wiley, New York, 1973.

Problems

1. Spontaneous supersymmetry breaking

Let Q be the generator of a $N = 1$ supersymmetric theory and $Q^\dagger$ its adjoint operator. With a proper normalization one has

$$\{Q, Q^\dagger\} = H,$$

where H is the hamiltonian of the system.

a Show that the hamiltonian of a supersymmetric theory contains no negative eigenvalues.

b Show that any state whose energy is not zero cannot be invariant under a supersymmetry transformation.

c Show that supersymmetry is spontaneously broken if and only if the energy of the lowest lying state (the vacuum) is not exactly zero.

d Consider the two-dimensional superconformal models on a cylinder, for which $Q = G_0$ and $H = Q^2 = L_0 - c/24$. Show that in the first model of the minimal unitary series, given by the tricritical Ising model, supersymmetry is broken while in the second minimal model, given by the gaussian field theory, it is exact.

2. Central charge of the parafermions

On a physical basis argue why in the limit $N \to \infty$ the central charge of the parafermionic systems is equal to $c = 2$.

3. Polyakov–Wiegman identity

Consider the action of the sigma model with a Wess–Zumino topological term

$$\mathcal{S}(g) \;=\; \frac{k}{16\pi} \int d^2x \,\mathrm{Tr}\,(\partial^\mu g^{-1} \partial_\mu g) + k\,\Gamma.$$

Prove the identity

$$\mathcal{S}(gh^{-1}) \;=\; \mathcal{S}(g) + \mathcal{S}(h) + \frac{k}{2\pi} \int d^2x \,\mathrm{Tr}\,(g^{-1}\partial_{\bar z} g\, h^{-1} \partial_z\, h).$$

Show that this identity gives rise to the invariance of the action under the transformation

$$g(z, \bar z) \to \mathcal{G}(z)\, g(z, \bar z)\, \mathcal{G}^{-1}(\bar z).$$

4. Correlation functions of the currents

For the conformal models with a Kac–Moody algebra, compute the four-point correlation functions of the analytic currents

$$\langle J^a(z_1)J^b(z_2)J^c(z_3)J^d(z_4)\rangle.$$

5. Bosonization of the $SU(2)_1$ theory

Verify that the central charge of the theory $SU(2)_1$ is $c = 1$. Compute the spectrum of the conformal weights of this theory and determine a representation of the corresponding conformal fields in terms of the vertex operators of a bosonic field φ.

14
The Arena of Conformal Models

Madamina il catalogo è questo.
Leporello, *Don Juan*

14.1 Introduction

In this chapter we will study some significant minimal conformal models. As shown in Chapter 11, these models provide explicit examples of exactly solved quantum field theories: of these theories we know the operator content, the fusion rules of their fields, the corresponding structure constants, the correlation functions of the order parameters and, finally, their modular invariant partition function on a torus. Despite this large amount of knowledge, there is still an important open problem, namely the identification of the classes of universality they are describing. Is there a way to associate these exactly solved critical theories to the continuum limit of lattice statistical models? Unfortunately there is no direct method to answer this question: the identification of the various classes of universality can be achieved only by comparison of the critical exponents predicted by conformal field theory with the values obtained by the exact solution of the models defined on a lattice, further supporting this identification on the basis of the symmetry of the order parameters. This has been the approach followed, for instance, by Huse who identified a particular critical regime of the lattice RSOS models solved by Andrew, Baxter, and Forrester with the unitarity minimal models of conformal field theory. In this chapter, rather than going into a technical analysis of this identification, we prefer to analyze in detail the first minimal models (in the following denoted, in general cases, by $\mathcal{M}_{p,q}$ and $\mathcal{M}_q$ for the unitary cases), in particular those corresponding to the Ising model, the tricritical Ising model and the Yang–Lee model. We will also discuss the three-state Potts model as an example of a statistical model associated to a partition function of the type (A, D), according to the notation introduced in Chapter 11. Finally, we will study the statistical models of geometric type (as, for instance, those that describe self-avoiding walks) and their formulation in terms of conformal minimal models.

14.2 The Ising Model

Consider the first minimal unitary conformal model, obtained by substituting $q = 3$ in eqn (11.3.2). Such a model has the central charge $c = \frac{1}{2}$ and the Kac table is reported in Table 14.1.

Table 14.1: *Kac table of the minimal unitary model $\mathcal{M}_3$.*

$\frac{1}{2}$	$\frac{1}{16}$	0
0	$\frac{1}{16}$	$\frac{1}{2}$

To denote the operator content of this theory let's introduce the notation[1]

$$\begin{aligned} \mathbf{1} &= (0,0) \\ \psi &= \left(\tfrac{1}{2},0\right) \\ \bar{\psi} &= \left(0,\tfrac{1}{2}\right) \\ \epsilon &= \left(\tfrac{1}{2},\tfrac{1}{2}\right) \\ \sigma &= \left(\tfrac{1}{16},\tfrac{1}{16}\right) \\ \mu &= \left(\tfrac{1}{16},\tfrac{1}{16}\right). \end{aligned} \tag{14.2.1}$$

Below we present a series of arguments to show that the conformal field theory described by this minimal model corresponds to the exact solution of the two-dimensional Ising model at its critical point.

The first indication comes from the numerical values of the Kac table. Assuming that the scalar field σ can be associated to the continuum limit of the magnetization field of the two-dimensional Ising model, for the corresponding critical index η the value is

$$\eta = \frac{1}{4}, \tag{14.2.2}$$

and coincides with the exact value known for this critical index from the exact lattice solution. Analogously, assuming that the scalar field ϵ describes the continuum limit of the energy operator of the two-dimensional Ising model (i.e. the conjugate operator to the temperature displacement $|T - T_c|$), we can derive the critical exponents ν and α:

$$\nu = 2 - 2\Delta_\epsilon = 1, \qquad \alpha = 2 - 1/(1 - \Delta_\epsilon) = 0. \tag{14.2.3}$$

Also in this case, these quantities coincide with their known exact values obtained by the lattice solution.

Further support for the hypothesis that the class of universality is that of the Ising model comes from the skeleton form of the fusion rules. Using the results of Chapter 11, the operator algebra that involves the fields σ, μ and the chiral field ψ (with analogous relations for the antichiral field $\bar{\psi}$) is given by

$$\begin{aligned} \psi\,\psi &= \mathbf{1} \\ \psi\,\sigma &= \mu \\ \psi\,\mu &= \sigma. \end{aligned} \tag{14.2.4}$$

[1] $(\Delta, \bar{\Delta})$ are the conformal weights provided by the Kac table.

These relations show that ψ is a fermionic field (here subject to antiperiodic boundary conditions) and that the operators σ and μ play the role of order and disorder fields. The fermionic structure present in the conformal model $\mathcal{M}_3$ perfectly matches the fermionic structure identified in the lattice version of the Ising model, discussed in Chapter 9, where we have showed that the continuum limit of the Ising model corresponds to a free fermionic theory for a Majorana field, with central charge $c = \frac{1}{2}$.

From the algebra of the scalar fields we have

$$\begin{aligned} \sigma\,\sigma &= \mathbf{1} + \epsilon \\ \mu\,\mu &= \mathbf{1} + \epsilon \\ \epsilon\,\sigma &= \sigma \\ \epsilon\,\mu &= \mu \\ \epsilon\,\epsilon &= \mathbf{1}. \end{aligned} \tag{14.2.5}$$

This algebra highlights the Z_2 spin symmetry of the Ising model, under which both σ and μ are odd fields ($\sigma \to -\sigma$, $\mu \to -\mu$) while ϵ is even, $\epsilon \to \epsilon$. Moreover, at its critical point the Ising model is also invariant under the Kramers–Wannier duality transformation, under which $\epsilon \leftarrow -\epsilon$ and $\sigma \leftrightarrow \mu$. The odd parity of ϵ under the duality transformation naturally explains the absence of ϵ in the operator product expansion of this field with itself.

Finally, note that the algebra (14.2.5) of the scalar fields can also be interpreted as the algebra of the composite operators of a φ^4 Landau–Ginzburg theory – a theory notoriously associated to the class of universality of the Ising model. In fact, following the general discussion presented in Section 11.6, let's impose $\sigma \equiv \varphi$. Using the operator expansion, we have $:\varphi^2: = \epsilon$ and $:\varphi^3: = \partial_z\,\partial_{\bar{z}}\varphi$. This shows that this conformal model provides the exact solution of the field theory associated to the lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\varphi)^2 + g\varphi^4.$$

Let's now discuss the correlation functions and the structure constants of this model.

14.2.1 Operator Product Expansion and Correlation Functions

If we identify the chiral field ψ with the analytic component of the Majorana fermion of the Ising model and $\bar{\psi}$ with its anti-analytic component, the continuum limit of the energy operator ϵ is given by[2] $\epsilon(z,\bar{z}) = i\,\bar{\psi}(\bar{z})\psi(z)$. The fermionic representation of this operator permits us to easily compute all its correlators using Wick's theorem. Since Wick's theorem always involves the contractions pairwise of different fields, it is easy to see that the only non-zero correlators are those with an even number of fields ϵ. The same conclusion can be reached based on the duality property of the model, since under this transformation $\epsilon \to -\epsilon$ and therefore only the correlation functions with an even number of ϵ can be different from zero. Using the factorization in the analytic and anti-analytic components, we have

[2]The i in this definition is necessary for the anticommutation rule of the fermionic field and the positivity of the correlation function $\langle\epsilon(z,\bar{z})\epsilon(w,\bar{w})\rangle = \frac{1}{|z-w|^2}$.

$$\begin{aligned} G_{2n} &= \langle \epsilon(z_1, \bar{z}_1) \dots \epsilon(z_n, \bar{z}_n) \rangle \\ &= (-1)^n \langle \psi(z_1) \bar{\psi}(\bar{z}_1) \dots \psi(z_n) \bar{\psi}(\bar{z}_n) \rangle \\ &= \langle \psi(z_1) \dots \psi(z_n) \rangle \, \langle \bar{\psi}(\bar{z}_1) \dots \bar{\psi}(\bar{z}_n) \rangle . \end{aligned} \tag{14.2.6}$$

For each of the two terms, Wick's theorem leads to the sum of all possible two-point correlation functions multiplied by the sign of the corresponding permutation. The final result can be expressed in terms of a Pfaffian of the $(2n) \times (2n)$ antisymmetric matrix A, with matrix elements $A_{ij} = -A_{ji} = \langle \psi(z_i) \psi(z_j) \rangle = 1/(z_i - z_j)$. We have then

$$\begin{aligned} \langle \epsilon(z_1, \bar{z}_1) \dots \epsilon(z_n, \bar{z}_n) \rangle \; &= \; \left| \mathrm{Pf} \left[\frac{1}{z_i - z_j} \right]_{1 \le i,j \le 2n} \right|^2 \\ &= \; \det \left[\frac{1}{z_i - z_j} \right] \end{aligned} \tag{14.2.7}$$

since the square of the Pfaffian of an antisymmetric matrix A is equal to its determinant.

For the computation of the correlation functions that involve the fields σ and μ, we can proceed in two different ways.

- The first method consists of applying the general strategy explained in Chapter 12: the operators σ and μ occupy the position $(1, 2)$ in the Kac table and therefore their correlators satisfy a second-order linear differential equation, whose explicit solution can be obtained by using the modified Coulomb gas approach. If we consider, for instance, the four-point correlation function[3]

$$F(\eta, \bar{\eta}) \; = \; \langle \sigma(\infty) \sigma(1,1) \sigma(\eta, \bar{\eta}) \sigma(0,0) \rangle ,$$

one gets

$$F(\eta, \bar{\eta}) \; = \; \left(\frac{1}{\eta \bar{\eta} (1-\eta)(1-\bar{\eta})} \right)^{1/8} \left[\mid Y_+(\eta) \mid^2 + \mid Y_-(\eta) \mid^2 \right] , \tag{14.2.8}$$

where

$$Y_\pm(\eta) \; = \; \sqrt{1 \pm \sqrt{1-\eta}} .$$

From the analysis of the singularity of this expression for $\eta \to 0$ and the operator expansion

$$\sigma(z_1, \bar{z}_1) \sigma(z_2, \bar{z}_2) \; = \; \frac{1}{|z_1 - z_2|^{1/4}} \, [\mathbf{1} + \cdots] + C^{\epsilon}_{\sigma\sigma} \, |z_1 - z_2|^{3/4} \, [\epsilon(z_2, \bar{z}_2) + \cdots]$$

one infers that the function $\mid Y_+(\eta) \mid^2$ corresponds to the channel of the identity operator $\mathbf{1}$ while $\mid Y_-(\eta) \mid^2$ corresponds to the channel of the operator ϵ. Using

[3] We use the Moebius invariance to fix three of the four points of this correlator at the positions $z_1 = \infty$, $z_2 = 1$ and $z_4 = 0$.

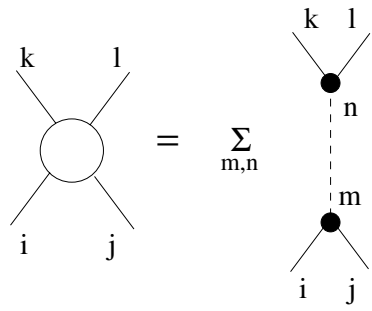


Fig. 14.1 *Expansion of the correlation functions in conformal blocks.*

the decomposition of the correlators in the conformal blocks showed in Fig. 14.1, we arrive at the quadratic equation for the structure constant $C^{\epsilon}_{\sigma\sigma}$

$$(C^{\epsilon}_{\sigma\sigma})^2 \,=\, \frac{1}{4}.$$

Note that this equation cannot fix the sign of the structure constant: hence our choice to take the positive sign

$$C^{\epsilon}_{\sigma\sigma} \,=\, \frac{1}{2} \tag{14.2.9}$$

is purely arbitrary. There is a point, though, with the choice of the sign of the structure constants. In order to appreciate this aspect, it is sufficient to observe that the four-point correlation function of the disorder operator

$$F(\eta,\bar{\eta}) \,=\, \langle \mu(\infty)\mu(1,1)\mu(\eta,\bar{\eta})\mu(0,0)\rangle,$$

is expressed in terms of the same function (14.2.8) and, from the singular term of its expression, we arrive at the same quadratic equation for the structure constant $C^{\epsilon}_{\mu\mu}$:

$$\left(C^{\epsilon}_{\mu\mu}\right)^2 \,=\, \frac{1}{4}.$$

However, in this case, we have to choose the negative solution

$$C^{\epsilon}_{\mu\mu} \,=\, -\frac{1}{2}. \tag{14.2.10}$$

To prove that this is the right choice, consider the four-point correlation function that involves both fields

$$G(\eta,\bar{\eta}) \,=\, \langle \mu(\infty)\sigma(1,1)\sigma(\eta,\bar{\eta})\mu(0,0)\rangle.$$

It satisfies the same second-order differential equation fulfilled by the previous correlators. However, its solution must take into account the semilocal property of these fields, i.e. the correlator should acquire a (-1) sign when the variable η

is analytically continued along the close contours that enclose either the origin or the point at infinity. Hence, in this case, the solution is given by

$$G(\eta,\bar\eta) \;=\; \frac{1}{2}\,\left(\frac{1}{\eta\bar\eta(1-\eta)(1-\bar\eta)}\right)^{1/8}\,\left[Y_+(\eta)\,Y_-(\bar\eta)+Y_-(\eta)\,Y_+(\bar\eta)\right]. \qquad (14.2.11)$$

Studying the singularities that are present in this expression when $\eta \to 1$ we get the equation

$$C^{\epsilon}_{\sigma\sigma}\,C^{\epsilon}_{\mu\mu} \;=\; -\frac{1}{4}, \qquad (14.2.12)$$

which clearly shows the equal and opposite value of the two structure constants.

Other correlation functions can be computed as well using straightforwardly the modified Coulomb gas. Instead of presenting these results, let's go on to illustrate another efficient method to compute the correlation functions of the Ising model.

- The second method for computing the correlators of the Ising model is based on the bosonization rules, exploiting the circumstance that the Ising model is a free fermionic theory. As a theory of real Majorana fermions, it cannot be directly bosonized but, if we consider two copies of the same theory, we can define a Dirac fermion theory that can be instead bosonized. Let $i = 1, 2$ be the index of each copy of the Ising model. In terms of the two Majorana fermions ψ_1 and ψ_2 (together with their anti-analytic components), we can define the Dirac field as

$$\Psi(z,\bar z) \;=\; \begin{pmatrix}\chi(z)\\ \bar\chi(\bar z)\end{pmatrix} \;=\; \frac{1}{\sqrt{2}}\,\begin{pmatrix}\psi_1+i\psi_2\\ \bar\psi_1+i\bar\psi_2\end{pmatrix} \qquad (14.2.13)$$

and apply the bosonization rule

$$\chi(z) \;=\; e^{i\phi(z)}, \qquad \bar\chi(\bar z) \;=\; e^{-i\bar\phi(\bar z)}. \qquad (14.2.14)$$

It is now essential to provide the bosonization representation of the various fields of the two copies of the Ising model. Let's start from the energy operator of the two-copy model, given by $\tilde\epsilon = \epsilon_1 \times \epsilon_2$. Using eqn (12.4.6) we have

$$\begin{aligned}(\bar\Psi\Psi)(z,\bar z) \;&=\; \psi_1\bar\psi_1+\psi_2\bar\psi_2 \\ &=\; i(\epsilon_1+\epsilon_2) \;=\; \cos\varphi(z,\bar z).\end{aligned} \qquad (14.2.15)$$

Since $\psi_1\,\psi_2 = i\partial_z\varphi$ we also have

$$\begin{aligned}\epsilon_1\epsilon_2 = (i\psi_1\bar\psi_1)\,(i\psi_2\bar\psi_2) \;&=\; \psi_1\psi_2\,\bar\psi_1\bar\psi_2 \\ &= -\partial_z\varphi\,\partial_{\bar z}\varphi.\end{aligned} \qquad (14.2.16)$$

Using these expressions and the correlators of the bosonic field φ, one easily recovers the previous expressions (14.2.7) of the correlators of the ϵ_i operators.

Let's consider now the correlators of the spin fields. For the two-copy model, the spin operator is expressed by the product of the spin operators of each copy, $\tilde\sigma = \sigma_1 \times \sigma_2$. Since the two copies do not interact with each other, the correlation

functions of $\tilde{\sigma}$ provide the *square* of the correlation functions of the original Ising model. Taking into account the conformal weight of the spin field, we can impose

$$\tilde{\sigma} \to \sqrt{2} \cos \frac{\varphi}{2} \tag{14.2.17}$$

and, using the two-point correlation function of this vertex operator, we find

$$\langle \tilde{\sigma}(z,\bar{z})\tilde{\sigma}(w,\bar{w})\rangle \;=\; \langle \sigma(z,\bar{z})\sigma(w,\bar{w})\rangle^2 \;=\; \frac{1}{|z-w|^{1/2}}. \tag{14.2.18}$$

Equation (14.2.17) enables us to compute all the (squares of the) correlators of the field σ of the Ising model. In fact,

$$\begin{aligned} \langle \sigma(z_1,\bar{z}_1)\cdots\sigma(z_n,\bar{z}_n)\rangle^2 \;&=\; 2^{n/2} \left\langle \prod_{i=1}^{n} \cos\frac{\varphi}{2}(z_i,\bar{z}_i) \right\rangle \\ &=\; 2^{-n/2} \sum_{\{\alpha_i=\pm 1\}} \prod_{i<j} |z_i - z_j|^{\alpha_i\alpha_j/2}. \end{aligned} \tag{14.2.19}$$

To characterize the disorder operator $\tilde{\mu} = \mu_1 \times \mu_2$ of the two-copy system, it is necessary to use duality. Under this transformation $\epsilon \to -\epsilon$, while $\sigma \leftrightarrow \mu$, and therefore in the bosonization formalism this symmetry is implemented by the substitution $\varphi \to \pi - \varphi$. In this way, we arrive at the identification

$$\tilde{\mu}(z,\bar{z}) \to \sqrt{2} \sin \frac{\varphi}{2}(z,\bar{z}). \tag{14.2.20}$$

We can now easily compute the mixed correlator

$$\begin{aligned} &\langle \sigma(z_1,\bar{z}_1)\mu(z_2,\bar{z}_2)\sigma(z_3,\bar{z}_3)\mu(z_4,\bar{z}_4)\rangle^2 \\ &\quad= \frac{1}{2} \frac{|z_{13}z_{24}|^{1/2}}{z_{14}z_{23}z_{12}z_{34}|^{1/2}} \left[-1 + \frac{|z_{12}z_{34}|}{|z_{13}z_{24}|} + \frac{|z_{14}z_{23}|}{|z_{13}z_{24}|} \right]. \end{aligned} \tag{14.2.21}$$

Thanks to this expression we can fix the operator product expansion of the order and disorder operators

$$\sigma(z_1,\bar{z}_1)\mu(z_2,\bar{z}_2) \;=\; \frac{C^{\psi}_{\sigma\mu}\,(z_1-z_2)^{1/2}\,[\psi(z_2)+\cdots] + C^{\bar{\psi}}_{\sigma\mu}\,(\bar{z}_1-\bar{z}_2)^{1/2}\,\left[\bar{\psi}(\bar{z}_2)+\cdots\right]}{|z_1-z_2|^{1/4}}.$$

Analyzing the limit $z_1 \to z_2$ and $z_3 \to z_4$, we have

$$C^{\psi}_{\sigma\mu} \;=\; \frac{e^{i\pi/4}}{\sqrt{2}} \qquad C^{\bar{\psi}}_{\sigma\mu} \;=\; \frac{e^{-i\pi/4}}{\sqrt{2}}.$$

In conclusion, from the bosonization procedure we get the following OPE

$$\begin{aligned}
\psi(z_1)\,\sigma(z_2,\bar{z}_2) &= \frac{e^{i\pi/4}}{\sqrt{2}(z_1-z_2)^{\frac{1}{2}}}\,\mu(z_2,\bar{z}_2)\\
\psi(z_1)\,\mu(z_2,\bar{z}_2) &= \frac{e^{-i\pi/4}}{\sqrt{2}(z_1-z_2)^{\frac{1}{2}}}\,\sigma(z_2,\bar{z}_2)\\
\bar{\psi}(\bar{z}_1)\,\sigma(z_2,\bar{z}_2) &= \frac{e^{-i\pi/4}}{\sqrt{2}(\bar{z}_1-\bar{z}_2)^{\frac{1}{2}}}\,\mu(z_2,\bar{z}_2)\\
\bar{\psi}(\bar{z}_1)\,\mu(z_2,\bar{z}_2) &= \frac{e^{i\pi/4}}{\sqrt{2}(\bar{z}_1-\bar{z}_2)^{\frac{1}{2}}}\,\sigma(z_2,\bar{z}_2).
\end{aligned} \tag{14.2.22}$$

Note that, in the basis of σ and μ chosen above, the 2×2 matrix representations of the zero-modes ψ_0 and $\bar{\psi}_0$ of the fermionic field are non-diagonal and given by

$$\psi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & e^{i\pi/4} \\ e^{-i\pi/4} & 0 \end{pmatrix}, \qquad \bar{\psi}_0 = \begin{pmatrix} 0 & e^{-i\pi/4} \\ e^{i\pi/4} & 0 \end{pmatrix}.$$

14.2.2 Coset Constructions and E_8 Algebra

At the critical point, remarkably enough, the Ising model can be described by two different coset conformal models. The first coset is based on the affine algebra of the group $SU(2)$, namely

$$\mathcal{M}_3 = \frac{SU(2)_1 \otimes SU(2)_1}{SU(2)_2}. \tag{14.2.23}$$

The representations of the affine algebra $SU(2)$ at the level $k=1$ are given by the multiplets of spin $(0)_1$ and $(\frac{1}{2})_1$, with conformal weights equal to 0 and $\frac{1}{4}$, whereas the representations at the level $k=2$ are given by the multiplets of spin $(0)_2$, $(\frac{1}{2})_2$ and $(1)_2$, with conformal weights 0, $\frac{3}{16}$, and $\frac{1}{2}$. The products of the two representations of $SU(2)_1$ decompose as

$$\begin{aligned}
(0)_1 \times (0)_1 &= [(0)_{Ising} \otimes (0)_2] \oplus \left[\left(\frac{1}{2}\right)_{Ising} \otimes (1)_2\right]\\
(0)_1 \times \left(\frac{1}{2}\right)_1 &= \left(\frac{1}{16}\right)_{Ising} \otimes \left(\frac{1}{2}\right)_2\\
\left(\frac{1}{2}\right)_1 \times \left(\frac{1}{2}\right)_1 &= [(0)_{Ising} \otimes (1)_2] \oplus \left[\left(\frac{1}{2}\right)_{Ising} \otimes (0)_2\right]
\end{aligned} \tag{14.2.24}$$

and therefore one recovers the Kac table of the model.

Quite surprisingly, the second coset construction uses the exceptional algebra E_8. Consider, in fact, the coset

$$\frac{(E_8)_1 \otimes (E_8)_1}{(E_8)_2}. \tag{14.2.25}$$

The dual Coxeter number of E_8 is $\tilde{h}=30$. Using formula (13.5.5) given in Chapter 13, we have $c=\frac{1}{2}$. At level $k=1$ there is only the representation of the identity field, with

conformal weight equal to 0. At level 2, there are instead three different representations, here denoted by the symbol Π_i, of the conformal weights

$$(E_8)_2 \to \quad \{1, \Pi_1, \Pi_7\} = \{0, \tfrac{15}{16}, \tfrac{3}{2}\}\,. \tag{14.2.26}$$

In particular, Π_1 is the adjoint representation of the group E_8. Their components, with respect to the basis of the simple roots of E_8 ($n_1, n_2, \ldots, n_8$, with n_i integers) are

$$\begin{array}{l} 1 \ \to (0,0,0,0,0,0,0,0) \\ \Pi_1 \to (1,0,0,0,0,0,0,0) \\ \Pi_7 \to (0,0,0,0,0,0,1,0). \end{array} \tag{14.2.27}$$

The Ising model is obtained by the decomposition

$$(0)_1 \times (0)_1 = [(0)_{Ising} \otimes (0)_2] \oplus \left[(\tfrac{1}{16})_{Ising} \otimes (\tfrac{15}{16})_2\right] \oplus \left[(\tfrac{1}{2})_{Ising} \otimes (\tfrac{3}{2})_2\right]. \tag{14.2.28}$$

The underlying E_8 structure of the Ising model will be decisive to understand its off-critical behavior when an external magnetic field is present.

14.2.3 Characters and Partition Function

The Kac table of the minimal model $\mathcal{M}_3$ has three fields and there are correspondingly three different characters of the Virasoro algebra, χ_0, $\chi_{\frac{1}{2}}$, and $\chi_{\frac{1}{16}}$. Their explicit expression can be computed by the Rocha–Caridi formula given in eqn (11.7.17) but, as we show below, they can also be computed by taking advantage of the fermionic formulation of the model and using the results of Section 12.3.2.

On a cylinder, the fermion has an expansion in half-integer or integer modes according if it satisfies antiperiodic or periodic boundary conditions along the space direction. Consider initially the antiperiodic case. If $|\,0\,\rangle$ is the lowest energy state in this sector, the excited states are given by $\psi_{-n_1} \ldots \psi_{-n_k}\,|\,0\rangle$, where $n_i \in \mathbf{Z} + \frac{1}{2}$. We can use the expression of $L_0 = \sum_{n>0} n\,\psi_{-n}\psi_n$ to order their sequence as the growth of their eigenvalues:

L_0 eigenvalue	state	
0	$\|0\rangle$	
$\frac{1}{2}$	$\psi_{-1/2}\,\|0\rangle$	
$\frac{3}{2}$	$\psi_{-3/2}\,\|0\rangle$	
2	$\psi_{-3/2}\,\psi_{-1/2}\,\|0\rangle$	
$\frac{5}{2}$	$\psi_{-5/2}\,\|0\rangle$	
3	$\psi_{-5/2}\,\psi_{-1/2}\,\|0\rangle$	
$\frac{7}{2}$	$\psi_{-7/2}\,\|0\rangle$	
4	$\psi_{-7/2}\,\psi_{-1/2}\,\|0\rangle$	$\psi_{-5/2}\,\psi_{-3/2}\,\|0\rangle$
$\cdots$	$\cdots$	

(14.2.29)

We have then

$$\mathrm{Tr}_A\, q^{L_0} \ = \ 1 + q^{1/2} + q^{3/2} + q^2 + q^{5/2} + q^3 + q^{7/2} + 2q^4 + \cdots \tag{14.2.30}$$

The states (14.2.29) form a representation of the Virasoro algebra with $c = \frac{1}{2}$ but such a representation is reducible for it can be decomposed into the direct sum of the two representations $[0] \oplus \left[\frac{1}{2}\right]$ of the minimal model $\mathcal{M}_3$. First of all, note that the states with conformal weights $\Delta = 0$ and $\Delta = \frac{1}{2}$ appear only once in the tower of these states. This means that these conformal families have a multiplicity equal to 1. Furthermore, note that the states that belong to the family $[0]$ are obtained by applying an even number of fermionic fields, while those of the family $\left[\frac{1}{2}\right]$ are obtained by acting on $|0\rangle$ by an odd number of operators ψ_{-n}. These two sets are therefore distinguished by their opposite eigenvalue with respect to the operator $(-1)^F$, and the irreducible representations are recovered by using the projectors $\frac{1}{2}(1 \pm (-1)^F)$

$$\begin{aligned} \chi_0(q) &\equiv q^{-1/48}\,\mathrm{Tr}_{\Delta=0} q^{L_0} \;=\; q^{-1/48}\,\mathrm{Tr}_A\,\frac{1}{2}(1+(-1)^F)\,q^{L_0} \\ \chi_{\frac{1}{2}}(q) &\equiv q^{-1/48}\,\mathrm{Tr}_{\Delta=\frac{1}{2}} q^{L_0} \;=\; q^{-1/48}\,\mathrm{Tr}_A\,\frac{1}{2}(1-(-1)^F)\,q^{L_0}. \end{aligned} \tag{14.2.31}$$

Let's now consider the periodic sector of the fermionic field, whose expression for L_0 on the cylinder is given by

$$L_0 = \sum_{n>0} n\psi_{-n}\psi_n + \frac{1}{16} \qquad n \in \mathbf{Z}.$$

The zero mode of the fermionic field has a two-dimensional representation space, spanned by $|\sigma\rangle = |\frac{1}{16}\rangle_+$ and $|\mu\rangle = |\frac{1}{16}\rangle_-$, which have eigenvalues ± 1 with respect to the operator $(-1)^F$. The tower of states in the periodic sector is expressed by

L_0 eigenvalue	state	
$\frac{1}{16}+0$	$\lvert\frac{1}{16}\rangle_\pm$	
$\frac{1}{16}+1$	$\psi_{-1}\lvert\frac{1}{16}\rangle_\pm$	
$\frac{1}{16}+2$	$\psi_{-2}\lvert\frac{1}{16}\rangle_\pm$	
$\frac{1}{16}+3$	$\psi_{-3}\lvert\frac{1}{16}\rangle_\pm$	$\psi_{-2}\,\psi_{-1}\lvert\frac{1}{16}\rangle_\pm$
$\cdots$	$\cdots$	

(14.2.32)

Hence, there are two irreducible representations associated to the two states $|\frac{1}{16}\rangle_\pm$. One may think of separating them using once more the projectors $\frac{1}{2}(1 \pm (-1)^F)$. However in this sector $\mathrm{Tr}_R\,(-1)^F\,q^{L_0} = 0$ identically because at each level there is always the same number of states with equal and opposite fermion number. In conclusion, there is the same expression for the character of the two families (another manifestation of the self-duality of the model) and this is given by

$$\chi_{\frac{1}{16}}(q) \;\equiv\; q^{-1/48}\,\mathrm{Tr}_P\,\frac{1}{2}(1\pm(-1)^F)q^{L_0} \;=\; q^{1/24}\,(1+q+q^2+2q^3+\cdots). \tag{14.2.33}$$

Partition functions. We can now use the characters χ_0, $\chi_{\frac{1}{2}}$, and $\chi_{\frac{1}{16}}$ to compute different partition functions on a torus and extract the relative operator content of the

model. Adopting the order of the characters given above, the modular matrix $\mathcal{S}$ that implements their transformation under $\tau \to -1/\tau$ is

$$\mathcal{S} = \frac{1}{2} \begin{pmatrix} 1 & 1 & \sqrt{2} \\ 1 & 1 & -\sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \end{pmatrix}. \tag{14.2.34}$$

If we consider the partition function with periodic boundary conditions along both horizontal and vertical axes of the torus, this quantity is given by the diagonal solution of the modular equation

$$Z_{PP}(q) = \mid \chi_0(q) \mid^2 + \mid \chi_{\frac{1}{2}} \mid^2 + \mid \chi_{\frac{1}{16}} \mid^2 . \tag{14.2.35}$$

In the presence of these boundary conditions, the operator content of the theory is expressed by the scalar conformal families $\{\mathbf{1}\}$, $\{\epsilon\}$, and $\{\sigma\}$.

We can also use the Z_2 symmetry of the model to implement other boundary conditions. Suppose we would like to compute the partition function with periodic boundary conditions along the space axis but with antiperiodic ones along the time axis for the spin field. This corresponds to computing the trace of an operator that implements a change of sign to the conformal family of the spin field $\sigma \to -\sigma$ but that leaves invariant both the identity and energy fields. The final expression is then

$$Z_{AP} = \mid \chi_0 \mid^2 + \mid \chi_{\frac{1}{2}} \mid^2 - \mid \chi_{\frac{1}{16}} \mid^2 . \tag{14.2.36}$$

Also in this case the operator content of the theory is expressed by the scalar conformal families $\{\mathbf{1}\}$, $\{\epsilon\}$, and $\{\sigma\}$, with a negative multiplicity of the last family for the given boundary conditions.

We can now use the modular transformation $\tau \to -1/\tau$ that induces a change of the horizontal and vertical axes to compute the partition function with antiperiodic boundary conditions along the horizontal axis and periodic along the vertical axis. Using eqn (14.2.34) to transform the characters, we have

$$Z_{PA} = \chi^* \chi_{\frac{1}{2}} + \chi^*_{\frac{1}{2}} \chi_0 + \mid \chi_{\frac{1}{16}} \mid^2 . \tag{14.2.37}$$

The operator content of the theory with these boundary conditions is expressed by the conformal scalar family $\{\sigma\}$ but, in this case, there are also the chiral and antichiral families $\{\psi, \mathbf{1}\}$ and $\{\mathbf{1}, \bar{\psi}\}$.

It is interesting to observe that the combination $Z_{AP} + Z_{PA}$ is invariant, by construction, under the modular transformation $\mathcal{S}$, and it is also invariant under $\mathcal{T}^2$, where $\mathcal{T}$ implements the transformation $\tau \to \tau + 1$. The partition function expressed by this combination

$$Z = Z_{AP} + Z_{PA} = \mid \chi_0 + \chi_{\frac{1}{2}} \mid^2 \tag{14.2.38}$$

corresponds to the operator content of the Ising model given by the fields $\mathbf{1}$, ψ, $\bar{\psi}$ and ϵ that are all mutually local. The spin field is not local with respect to both ψ and $\bar{\psi}$ and is therefore absent in this situation.

14.3 The Universality Class of the Tricritical Ising Model

Let's now discuss the universality class of the Tricritical Ising Model (TIM), associated to the second unitary minimal model $\mathcal{M}_4$. One of its microscopic realizations is provided by the Blume–Capel model that was discussed in Section 7.7.2. Equivalently, this class of universality can be associated to a Landau–Ginzburg lagrangian based on a scalar field φ, a formulation that has the advantage of easy bookkeeping of the Z_2 symmetry property of each order parameter. The euclidean action is

$$\mathcal{S} = \int d^D x \left[\frac{1}{2} (\partial_\mu \varphi)^2 + g_1 \varphi + g_2 \varphi^2 + g_3 \varphi^3 + g_4 \varphi^4 + \varphi^6 \right], \tag{14.3.1}$$

with the tricritical point identified by the condition $g_1 = g_2 = g_3 = g_4 = 0$. We recall that the statistical interpretation of the coupling constants reads as follows: g_1 plays the role of an external magnetic field h, g_2 measures the displacement of the temperature from its critical value, i.e. $g_2 \sim (T - T_c)$, g_3 may be regarded as a subleading magnetic field h' and, finally, g_4 may be interpreted as a chemical potential for the vacancies.

In two dimensions – the case of interest here – there are strong fluctuations of the order parameters and this implies that the critical exponents and the universal ratios are quite different from their estimates provided by a mean field theory. We can use the conformal theory to obtain an exact solution of this model at its critical point. In fact, as we show below, it is described by the second unitary minimal model $\mathcal{M}_4$: its central charge is $c = \frac{7}{10}$ and the exact values of its conformal weight are

$$\Delta_{l,k} = \frac{(5l - 4k)^2 - 1}{80}, \quad \begin{matrix} 1 \le l \le 3 \\ 1 \le k \le 4. \end{matrix} \tag{14.3.2}$$

They are organized in the Kac Table 14.2.

There are six scalar primary fields and, out of them, four are relevant operators: the operator product expansion algebra and the relative structure constants are reported in Table 14.3. The correlation functions of these fields can be computed straightforwardly using the modified Coulomb gas, as proposed in Problem 2, and will not be presented here.

Landau–Ginzburg. The six primary fields perfectly match the identification provided by the composite fields of the Landau–Ginzburg theory and by the symmetries

Table 14.2: *Kac table of the unitary minimal model* $\mathcal{M}_4$.

$\frac{3}{2}$	$\frac{6}{10}$	$\frac{1}{10}$	0
$\frac{7}{16}$	$\frac{3}{80}$	$\frac{3}{80}$	$\frac{7}{16}$
0	$\frac{1}{10}$	$\frac{6}{10}$	$\frac{3}{2}$

Table 14.3: *Fusion rules of the tricritical Ising model.*

even $*$ *even* $\epsilon * \epsilon = [1] + c_1\,[t]$ $t * t = [1] + c_2\,[t]$ $\epsilon * t = c_1\,[\epsilon] + c_3\,[\varepsilon'']$	$c_1 = \frac{2}{3}\sqrt{\frac{\Gamma(\frac{4}{5})\Gamma^3(\frac{2}{5})}{\Gamma(\frac{1}{5})\Gamma^3(\frac{3}{5})}}$ $c_2 = c_1$ $c_3 = \frac{3}{7}$
even $*$ *odd* $\epsilon * \sigma' = c_4\,[\sigma]$ $\epsilon * \sigma = c_4\,[\sigma'] + c_5\,[\sigma]$ $t * \sigma' = c_6\,[\sigma]$ $t * \sigma = c_6\,[\sigma'] + c_7\,[\sigma]$	$c_4 = \frac{1}{2}$ $c_5 = \frac{3}{2}c_1$ $c_6 = \frac{3}{4}$ $c_7 = \frac{1}{4}c_1$
odd $*$ *odd* $\sigma' * \sigma' = [1] + c_8\,[\varepsilon'']$ $\sigma' * \sigma = c_4\,[\epsilon] + c_6\,[t]$ $\sigma * \sigma = [1] + c_5\,[\epsilon] + c_7\,[t] + c_9\,[\varepsilon'']$	$c_8 = \frac{7}{8}$ $c_9 = \frac{1}{56}$

of the model. There are two different Z_2 symmetries, one associated to the spin transformation, the other to the duality.

With respect to the Z_2 spin symmetry $\varphi \to -\varphi$ we have

1. two odd fields: the magnetization operator $\sigma = \phi_{\frac{3}{80},\frac{3}{80}} \equiv \varphi$ and the sub-leading magnetic operator $\sigma' = \phi_{\frac{7}{16},\frac{7}{16}} \equiv: \varphi^3 :$;
2. four even fields: the identity operator $1 = \phi_{0,0}$, the energy operator $\varepsilon = \phi_{\frac{1}{10},\frac{1}{10}} \equiv: \varphi^2 :$, and the density operator $t = \phi_{\frac{6}{10},\frac{6}{10}} \equiv: \varphi^4 :$, associated to the vacancies. Finally, there is also the irrelevant field $\varepsilon'' = \phi_{\frac{3}{2},\frac{3}{2}}$. The operator product expansion of these fields gives rise to a subalgebra of the fusion rules.

As for the Ising model, also for the TIM there is another Z_2 associated to the duality transformation, under which the fields change as

- the magnetization order parameters change into the disorder operators

$$\mu = D^{-1}\sigma D = \tilde{\phi}_{\frac{3}{80},\frac{3}{80}}, \quad \mu' = D^{-1}\sigma' D = \tilde{\phi}_{\frac{7}{16},\frac{7}{16}}; \tag{14.3.3}$$

- the even fields transform instead in themselves

$$D^{-1}\varepsilon D = -\varepsilon, \quad D^{-1}tD = t, \quad D^{-1}\varepsilon'' D = -\varepsilon'', \tag{14.3.4}$$

  ε and ε'' are odd fields while t is an even field under this transformation.

Supersymmetry. It is interesting to note that this critical model provides an explicit realization of a supersymmetric field theory. In fact, the TIM is also the first model of the minimal unitary superconformal series: the Z_2 even fields enter the definition of a

superfield of the Neveu–Schwarz sector

$$\mathcal{N}(z,\bar{z},\theta,\bar{\theta}) = \varepsilon(z,\bar{z}) + \bar{\theta}\,\psi(z,\bar{z}) + \theta\,\bar{\psi}(z,\bar{z}) + i\,\theta\bar{\theta}\,t(z,\bar{z}), \tag{14.3.5}$$

while the Z_2 odd magnetization operators form two irreducible representations of the Ramond sector. The supersymmetric Landau–Ginzburg model can be written as

$$\mathcal{S} = \int d^2x\, d^2\theta \left[\frac{1}{2}\mathcal{D}\mathcal{N}\,\bar{\mathcal{D}}\mathcal{N} + \mathcal{N}^3\right], \tag{14.3.6}$$

where $\mathcal{D}$ and $\bar{\mathcal{D}}$ are the covariant derivatives

$$\mathcal{D} = \frac{\partial}{\partial\theta} - \theta\frac{\partial}{\partial z}, \quad \overline{\mathcal{D}} = \frac{\partial}{\partial\overline{\theta}} - \overline{\theta}\frac{\partial}{\partial\overline{z}}. \tag{14.3.7}$$

Note that the supersymmetry and the organization of its Z_2 even primary fields in a superfield are at the root of the relationships that link the various structure constants (see, for instance, the identity $c_2 = c_1$).

Exceptional algebra E_7. In addition to the conformal and superconformal invariance, the TIM holds another surprise. In fact, it can also be realized in terms of a coset on the exceptional algebra E_7

$$\mathcal{M}_4 \;=\; \frac{(E_7)_1 \otimes (E_7)_1}{(E_7)_2}. \tag{14.3.8}$$

For E_7, the dual Coxeter number is $\tilde{h} = 18$ and therefore the central charge of this coset theory is $c = \frac{7}{10}$. At the level $k = 1$, the possible representations are given by the identity 1 and the representation Π_6, with conformal weights equal to 0 and $\frac{3}{4}$, respectively:

$$(E_7)_1 \to \quad \{1, \Pi_6\} = \{0, \tfrac{3}{4}\}. \tag{14.3.9}$$

Their components with respect to the simple roots of E_7 ($n_1, n_2, \ldots, n_7$, with n_i integers) are

$$\begin{array}{l} 1 \;\to (0,0,0,0,0,0,0) \\ \Pi_6 \to (0,0,0,0,0,1,0). \end{array} \tag{14.3.10}$$

At the level $k = 2$, there are instead the representations

$$(E_7)_2 \to \quad \{1, \Pi_1, \Pi_2, \Pi_5, \Pi_6\} = \{0, \tfrac{9}{10}, \tfrac{21}{16}, \tfrac{7}{5}, \tfrac{57}{80}\}\,, \tag{14.3.11}$$

with the corresponding fundamental weights given by

$$\begin{array}{l} \Pi_1 \to (1,0,0,0,0,0,0) \\ \Pi_2 \to (0,1,0,0,0,0,0) \\ \Pi_5 \to (0,0,0,0,1,0,0). \end{array} \tag{14.3.12}$$

Π_1 is the adjoint representation E_7. We can recover the conformal weights of the TIM by the decomposition of the various representations

$$\begin{aligned}
(0)_1 \times (0)_1 &= [(0)_{TIM} \otimes (0)_2] \quad + \left[\left(\frac{1}{10}\right)_{TIM} \otimes (\Pi_1)_2\right] + \left[\left(\frac{6}{10}\right)_{TIM} \otimes (\Pi_5)_2\right] \\
(0)_1 \times \left(\frac{3}{4}\right)_1 &= \left[\left(\frac{7}{16}\right)_{TIM} \otimes (\Pi_2)_2\right] + \left[\left(\frac{3}{80}\right)_{TIM} \otimes (\Pi_6)_2\right] \\
\left(\frac{3}{4}\right)_1 \times \left(\frac{3}{4}\right)_1 &= \left(\frac{3}{2}\right)_{TIM} \otimes (0)_2 \,.
\end{aligned} \tag{14.3.13}$$

Note that the energy operator $\Phi_{\frac{1}{10},\frac{1}{10}}$ is associated to the adjoint representation of E_7, an observation that will be crucial in the analysis of the off-critical model, when the temperature is moved away from its critical value $T \neq T_c$.

14.4 Three-state Potts Model

On a square lattice, the hamiltonian of the three-state Potts model is given by

$$H \;=\; -\frac{J}{2} \sum_{x,\alpha} (\sigma_x\, \bar{\sigma}_{x+\alpha} + \bar{\sigma}_x\, \sigma_{x+\alpha}) \;=\; -J \sum_{x,\alpha} \cos(\eta_x - \eta_{x+\alpha}), \tag{14.4.1}$$

where the discrete spin variables are represented by $\sigma = \exp(i\eta)$, $\bar{\sigma} = \exp(-i\eta)$, with $\eta = 0, \pm\frac{2\pi}{3}$. It is known that this model has a duality symmetry and, at its self-dual point $J_c = \frac{2}{3}\,\log(\sqrt{3}+1)$, presents a second-order phase transition. The lattice theory is exactly solvable and consequently all critical exponents are known. In this section we plan to show that the conformal theory that emerges at the critical point coincides with the unitary minimal model $\mathcal{M}_5$. More precisely, the operator content of the three-state Potts model is given only by a subset of the Kac table of the conformal model $\mathcal{M}_5$. The subset of fields are those entering the modular invariant partition function of type (A, D).

To find the field theory description of the microscopic statistical model, let's assume that there exists the continuum limit of its spin and energy operators, here denoted by $\sigma(x)$, $\bar{\sigma}(x)$, and $\epsilon(x)$. Moreover, let's assume that

$$\begin{aligned}
\sigma(x_1)\bar{\sigma}(x_2) + \bar{\sigma}(x_1)\sigma(x_2) \;&=\; \frac{1}{|x_1 - x_2|^{2\Delta_\sigma}} + \mathbf{C}^{\epsilon}_{\sigma\bar{\sigma}}\, \frac{1}{|x_1 - x_2|^{2\Delta_\sigma - \Delta_\epsilon}}\, \epsilon(x_2) + \cdots \\
\epsilon(x_1)\sigma(x_2) \;&=\; \mathbf{C}^{\sigma}_{\epsilon\sigma}\, \frac{1}{|x_1 - x_2|^{\Delta_\epsilon}}\, \sigma(x_2) + \cdots \\
\epsilon(x_1)\epsilon(x_2) \;&=\; \frac{1}{|x_1 - x_2|^{2\Delta_\epsilon}}.
\end{aligned} \tag{14.4.2}$$

From the known expression of the critical exponents $\alpha = \frac{1}{3}$ and $\beta = \frac{1}{9}$ coming from the exact solution of the lattice model, we can infer the conformal weights of the scaling operators

$$\Delta_\sigma \;=\; \Delta_{\bar{\sigma}} \;=\; \frac{1}{15}, \qquad \Delta_\epsilon \;=\; \frac{2}{5}. \tag{14.4.3}$$

Let's now consider the Kac table of the minimal model $\mathcal{M}_5$, reported in Table 14.4.

Table 14.4: *Kac table of the unitary minimal model* $\mathcal{M}_5$.

3	$\frac{13}{8}$	$\frac{2}{3}$	$\frac{1}{8}$	0
$\frac{7}{5}$	$\frac{21}{40}$	$\frac{1}{15}$	$\frac{1}{40}$	$\frac{2}{5}$
$\frac{2}{5}$	$\frac{1}{40}$	$\frac{1}{15}$	$\frac{21}{40}$	$\frac{7}{5}$
0	$\frac{1}{8}$	$\frac{2}{3}$	$\frac{13}{8}$	3

In this table there is the field $\Phi_{3,3} = \Phi_{2,3}$, with conformal weight $\Delta_\sigma = \frac{1}{15}$ and the field $\Phi_{2,1} = \Phi_{3,5}$ with $\Delta_\epsilon = \frac{2}{5}$. It is therefore natural to identify these conformal fields with the scaling operators associated to the spin and energy operators of the lattice model. However the exact solution of the lattice model does not have operators with conformal weights $\frac{1}{8}$, $\frac{1}{40}$, $\frac{21}{40}$, and $\frac{13}{8}$. What is then the correct identification of the Z_3 Potts model?

To answer this question, one should recall that, for $p \geq 5$, the conformal minimal model $\mathcal{M}_p$ admits *two* different partition functions. The first of them is the purely diagonal partition function, i.e. the one in which all fields of the Kac table appear each with multiplicity equal to 1. This leads to the expression

$$Z_{diag} = \frac{1}{2} \sum_{r=1}^{4} \sum_{s=1}^{5} |\chi_{r,s}|^2. \tag{14.4.4}$$

The field theory associated to the operator content of this partition function does not correspond to the three-state Potts model but it rather defines the critical theory of a Landau–Ginzburg scalar field φ, that presents only a Z_2 invariance $\varphi \to -\varphi$. Its action is

$$\mathcal{S} = \int d^2x \left[\frac{1}{2} (\partial_\mu \varphi)^2 + \varphi^8 \right]. \tag{14.4.5}$$

There is, however, another modular invariant partition function associated to the minimal model $\mathcal{M}_5$ expressed by

$$Z_{Potts} = \sum_{r=1,2} \left\{ |\chi_{r,1} + \chi_{r,5}|^2 + 2|\chi_{r,3}|^2 \right\}. \tag{14.4.6}$$

The operator content identified by this partition function is different from the previous one: it involves only a subset of the fields of the Kac table of the minimal model $\mathcal{M}_5$. There are, in fact, only the fields $\Phi_{r,s}$ with $s = 1, 5$ and $r = 1, 2$. Combining the analytic and the anti-analytic parts, the critical theory described by this partition function has the scalar fields given in Table 14.5. These fields close an operator algebra, also in the absence of the other fields of the Kac table. Their skeleton fusion rules are reported in Table 14.6.

Table 14.5: *Scalar operators of the non-diagonal partition function of the model $\mathcal{M}_5$.*

(r,s)	Δ	Field	Interpretation
$(1,1)$ or $(4,5)$	0	$\mathbf{1}$	Identity
$(2,1)$ or $(3,5)$	$\frac{2}{5}$	ϵ	energy
$(3,3)$ or $(2,3)$	$\frac{1}{15}$	σ	spin
$(3,1)$ or $(2,5)$	$\frac{7}{5}$	$\mathbf{X}$	
$(4,1)$ or $(1,5)$	3	$\mathbf{Y}$	
$(4,3)$ or $(1,3)$	$\frac{2}{3}$	$\mathbf{Z}$	

Table 14.6: *Fusion rules of the scalar fields of the three-state Potts model.*

$\epsilon \times \epsilon = \mathbf{1} + X$	$\epsilon \times \sigma = \sigma + Z$
$\epsilon \times X = \epsilon + Y$	$\epsilon \times Y = X$
$\sigma \times \sigma = \mathbf{1} + \epsilon + \sigma + X + Y + Z$	$\sigma \times X = \sigma + Z$
$\sigma \times Y = \sigma$	$\sigma \times Z = \epsilon + \sigma + X$
$X \times X = \mathbf{1} + X$	$X \times Y = \epsilon$
$X \times Z = \sigma$	$Y \times Y = \mathbf{1}$
$Y \times Z = Z$	$Z \times Z = \mathbf{1} + Y + Z$

In addition to these scalar fields there are certain fields with spin, here denoted by their conformal weights $\Phi_{(\Delta,\bar{\Delta})}$. They are constructed by combining in a non-diagonal way the analytic and anti-analytic fields: $W = \Phi_{(3,0)}$, $\bar{W} = \Phi_{(0,3)}$, $J = \Phi_{(\frac{7}{5},\frac{2}{5})}$, and $\bar{J} = \Phi_{(\frac{2}{5},\frac{7}{5})}$.

It is interesting to observe that the three-state Potts models at criticality can also be obtained by the parafermionic theory $\mathbf{Z}_N$ with $N = 3$. It is easy to check the equality of the central charges of both theories, as well as the conformal weights of the spin operator $\Delta_\sigma = \frac{1}{15}$. The role of the parafermionic current is here played by the chiral operator $\Phi_{1,3}$, with conformal weight $\Delta_\psi = \frac{2}{3}$.

Generalization. Remarkably, the analysis presented for the three-state Potts model can be generalized to the $\mathcal{Q}$-state Potts model, where $\mathcal{Q}$ is regarded as a continuous variable (see Chapter 2). The range of values of $\mathcal{Q}$ for which the Potts model is critical is given by the interval $\mathcal{Q} \in (1, 4)$: for $\mathcal{Q} = 1$, the Potts model describes the critical phenomenon of percolation, for $\mathcal{Q} = 2$ we have the usual Ising model, while for $\mathcal{Q} > 4$, the Potts model presents a first-order phase transition that cannot be described by a conformal field theory. The relation that identifies the minimal models $\mathcal{M}_p$ with the $\mathcal{Q}$-state Potts model is

$$\mathcal{Q} = 4 \cos^2 \frac{\pi}{p+1}, \tag{14.4.7}$$

and it is easy to check that it correctly reproduces, for $p = 3$ the Ising model ($with \mathcal{Q} = 2$), for $p = 5$ the three-state Potts model and for $p \to \infty$ the four-state Potts model. The exact solution of the lattice models is known for generic values of $\mathcal{Q}$ and therefore all values of the thermal and magnetic critical exponents are known as well. This permits the identification of the anomalous dimension of several order parameters, as

$$\begin{aligned} X_{T_n} &= 2\Delta_{n+1,1} = \frac{n^2 + ny}{2 - y}, \\ X_{H_n} &= 2\Delta_{N-1-n,n} = \frac{(2n_1)^2 - y^2}{4(2 - y)}, \end{aligned} \tag{14.4.8}$$

where we have introduced the notation

$$N \equiv \frac{p+1}{2}, \qquad y \equiv \frac{1}{N}. \tag{14.4.9}$$

14.5 The Yang–Lee Model

Among the minimal non-unitary models, a simple but particularly significant example is given by the model $\mathcal{M}_{2,5}$. Its central charge is $c = -22/5$ and the Kac table consists of only one row, as shown in Table 14.7. In addition to the identity operator, there is only a field φ of conformal weight $\Delta = -1/5$. Hence the effective central charge is $c_{eff} = c - 24\Delta_{min} = 2/5$. As shown originally by J.L. Cardy, this model admits a statistical interpretation in terms of a field theory associated to the Yang–Lee zero singularities of the Ising model. Let's discuss the main steps that lead to this conclusion, by initially recalling the Yang–Lee theorem. The partition function of a statistical model defined on a lattice is an analytic function of its parameters as long as the number N of the fluctuating variables is finite. Its singularities only emerge in the thermodynamical limit $N \to \infty$. Consider the Ising model at a given value T of the temperature and in the presence of an external magnetic field B. As a function of B, at finite N, the zeros of the partition function cannot be on the real axis of B,

Table 14.7: *Kac table of the minimal non-unitary model* $\mathcal{M}_{2,5}$.

0	$-\frac{1}{5}$	$-\frac{1}{5}$	0

since Z is expressed by a sum of positive terms. Hence they are placed in complex conjugate points of the complex plane B and they tend to accumulate along certain curves in the limit $N \to \infty$. In particular, as shown by C.N. Yang and T.D. Lee, in the Ising model these zeros accumulate along the imaginary axis $B = ih$. Correspondingly the free energy of the system can be expressed in terms of the density of these zeros along the imaginary axis

$$F(b) \,=\, \log Z \,=\, \int_{-\infty}^{+\infty} dx\, \rho(x,T)\, \log(h - ix), \tag{14.5.1}$$

with the magnetization given by

$$M \,=\, \frac{\partial F}{\partial B} \,=\, \int_{-\infty}^{+\infty} dx\, \frac{\rho(x,T)}{h - ix}. \tag{14.5.2}$$

Below the critical temperature, i.e. for $T < T_c$, the distribution of the zeros extends to the real axis, so that $\rho(0,T) \neq 0$. Consequently the magnetization is a discontinuous function of the variable B when it crosses the real axis, and the system presents a first-order phase transition. Precisely at $T = T_c$ we have $\rho(0,T_c) = 0$ and there is a second-order phase transition. In the high-temperature phase $T > T_c$, the system is paramagnetic and the distribution of the zeros starts from two symmetric critical values $\pm h_c(T)$ and then extends along the magnetic axis (see Fig. 14.2).

In the vicinity of h_c, the density of the zeros has an anomalous behavior

$$\rho(h,T) \,=\, (h - h_c)^{\sigma}. \tag{14.5.3}$$

An analogous anomalous behavior is present in the magnetization, as a function of the (complex) magnetic field

$$M(ih) \simeq (h - h_c)^{\sigma}. \tag{14.5.4}$$

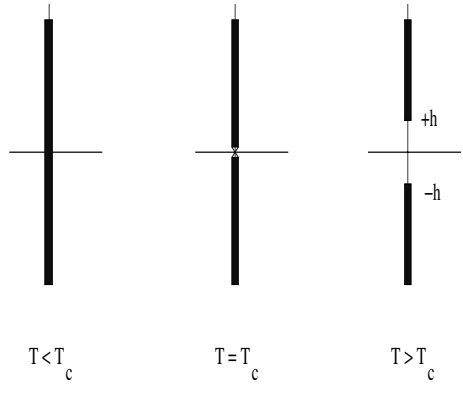


Fig. 14.2 *Density of the zeros of the partition function of the Ising model in the complex plane of the variable B by varying the temperature.*

Thanks to the thermodynamic relations discussed in Chapter 1, one can link the critical exponent σ to the critical exponent η of the operator corresponding to the fluctuations of the model in the presence of an imaginary magnetic field

$$\sigma \,=\, \frac{1}{\delta} \,=\, \frac{d-2+\eta}{d+2-\eta}. \tag{14.5.5}$$

Fisher has identified the effective action of the order paramter, given by the Landau–Ginzburg theory:

$$\mathcal{S} \,=\, \int d^d x \left[\frac{1}{2}(\partial\varphi)^2 + i(h-h_c)\,\varphi + ig\,\varphi^3\right]. \tag{14.5.6}$$

Note that the non-unitarity of the model manifests itself in the imaginary value of the coupling constant. In two dimensions, a model that could reproduce the dynamics of such a theory should satisfy two main requests: (i) it must be a non-unitary model; (ii) it must have only one relevant field φ satisfying the fusion rule

$$\varphi \times \varphi \,=\, \mathbf{1} + C^{\varphi}_{\varphi,\varphi}\,\varphi, \tag{14.5.7}$$

with a purely imaginary structure constant $C^{\varphi}_{\varphi,\varphi}$.

These are precisely the features of the minimal non-unitary model $\mathcal{M}_{2,5}$ whose structure constant is given by

$$C^{\varphi}_{\varphi,\varphi} \,=\, i\left[\frac{\Gamma^2\left(\frac{6}{5}\right)\Gamma\left(\frac{1}{5}\right)\Gamma\left(\frac{2}{5}\right)}{\Gamma\left(\frac{3}{5}\right)\,\Gamma^3\left(\frac{4}{5}\right)}\right]^{1/2}. \tag{14.5.8}$$

This quantity can be computed by using the exact expression of the four-point correlation function of the field φ. Since this field occupies the position $(1,2)$ of the Kac table, its correlators are given either by solving the corresponding second-order differential equation or applying the modified Coulomb gas method. The result is

$$\begin{aligned}&\langle\varphi(z_1,\bar z_1)\varphi(z_2,\bar z_2)\varphi(z_3,\bar z_3)\varphi(z_4,\bar z_4)\rangle \\ &= \left|\frac{z_{13}\,z_{24}}{z_{12}\,z_{23}\,z_{34}\,z_{14}}\right|^{-4/5}\left\{|F_1(\eta)|^2 + C^2\,|F_2(\eta)|^2\right\}\end{aligned} \tag{14.5.9}$$

where η is the harmonic ratio $\eta = z_{12}z_{34}/z_{13}z_{24}$ and $F_i(\eta)$ are the hypergeometric functions

$$\begin{aligned}F_1(\eta) \,&=\, F\left(\frac{3}{5},\frac{4}{5},\frac{6}{5},\eta\right) \\ F_2(\eta) \,&=\, \eta^{-1/5}\,F\left(\frac{3}{5},\frac{2}{5},\frac{4}{5},\eta\right).\end{aligned} \tag{14.5.10}$$

The value of the critical exponent σ predicted by this conformal model is $\sigma = -1/6$, in reasonable agreement with its numerical determination $\sigma = -0.163$.

14.6 Conformal Models with $O(n)$ Symmetry

As we have seen in Chapter 2, the spin models with a continuous symmetry $O(n)$ provide a generalization of the Ising model and, in particular, their limit $n \to 0$ describes the universality class of self-avoiding random walks. In these theories, the spins are vectors $\vec{S}$ with n components and length $|\vec{S}|^2 = n$. Taking advantage of the universality of critical phenomena, we can choose any microscopic lattice to study their behavior. The most convenient one turns out to be a lattice with coordination number equal to 3, as for instance the hexagonal lattice shown in Fig. 14.3.

We assume that the partition function of the system is expressed by

$$Z = \int \prod_k d\vec{S}_k \prod_{\langle i,j \rangle} (1 + x\, \vec{S}_i \cdot \vec{S}_j), \tag{14.6.1}$$

where the product on i and j is on the nearest neighbor sites. The integration rules on the spins are

$$\int dS^a\, S^a = 0$$
$$\int dS^a\, (S^a)^2 = 1$$
$$\int d\vec{S}\, \mathbf{S}^2 = n.$$

Now expand the product $\prod_{\langle i,j \rangle}(1 + x\vec{S}_i \cdot \vec{S}_j)$ and integrate over the values of the spins: due to the coordination number of the lattice and the integration rules stated above, the only terms that are different from zero are those relative to the self-avoiding closed circuits. Since each of these circuits carries a factor n coming from the integration on the spins and a factor x for each of its segments, the partition function becomes

$$Z = \sum_{closed\ circuit} n^{N_C}\, x^{N_S}, \tag{14.6.2}$$

where N_C is the number of close a circuits, while N_S is the number of segments. We can use this expression to analytically continue the definition of the model to arbitrary

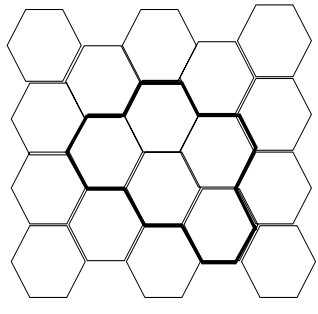


Fig. 14.3 *Hexagonal lattice and one of its closed spin circuits.*

values of n, not necessarily integers. The partition function presents a critical point x_c given by

$$x_c \,=\, (2+\sqrt{2-n})^{-1/2}, \tag{14.6.3}$$

at which there is a second-order phase transition. This is described by a conformal field theory with central charge $c(n) = 1 - 6/k(k+1)$, where the relation that links n and k is expressed by

$$n \,=\, 2\,\cos(\pi/k), \qquad k \geq 1. \tag{14.6.4}$$

Note, in particular, that $c = 0$ when $n = 0$ but its derivative $\partial c/\partial n$ at $n = 0$ is different from zero and equal to $5/3\pi$. For $n = 1$, $c = 1/2$ and we recover the Ising model. The anomalous dimension of the energy operator of these theories is

$$\eta_e \,=\, 2(k-1)/(k+1), \tag{14.6.5}$$

i.e. $\eta_e = 2/3$ for $n = 0$. This exponent is related to the exponent ν that characterizes the divergence of the correlation length by the relation $\nu = 1/(2 - x_e) = 3/4$. This value is in perfect agreement with the critical exponent of the exact lattice solution of the self-avoiding random walk found by B. Nienhuis.

References and Further Reading

For exactly solved lattice models and their identification with minimal models of CFT see:

G.E. Andrew, R.J. Baxter and P.J. Forrester, *Eight-vertex SOS model and generalized Rogers–Ramanujan-type identities*, J. Stat. Phys. 35 (1984), 193.

D. Huse, *Exact exponents for infinitely many new multicritical points*, Phys. Rev. B 30 (1984), 2908.

A review article on the universality class of the tricritical point is:

I. Lawrie, S. Serbach, *Theory of tricritical points*, in Phase Transitions, Vol. 9 (1984),

The superconformal invariance of the two-dimensional tricritical Ising model was pointed out by Friedan, Qiu, and Shenker and studied in the articles:

D. Friedan, Z.A. Qiu, S. Shenker, *Superconformal invariance in two-dimensions and the tricritical ising model*, Phys. Lett. B 151 (1985), 37.

Z.A. Qiu, *Supersymmetry, Two-dimensional critical phenomena and the tricritical Ising model*, Nucl. Phys. B 270 (1986), 205.

G. Mussardo, G. Sotkov, M. Stanishkov, *Ramond sector of the supersymmetric minimal models*, Phys. Lett. B 195 (1987), 397.

The identification of the universality class of the three-state Potts model with unitary minimal model $\mathcal{M}_5$ is discussed in the article:

V.S. Dotsenko, *Critical behavior and associated conformal algebra of the Z(3) Potts model*, Nucl. Phys. B 235 (1984), 54.

The identification of the minimal non-unitary model $\mathcal{M}_{2,5}$ with the Yang–Lee edge singularity of the zeros of the partition function has been proposed in:

J.L. Cardy, *Conformal invariance and the Yang–Lee edge singularity in two dimensions*, Phys. Rev. Lett. 54 (1985), 1354.

The Coulomb gas formalism and the relation with $O(n)$ statistical models are the subject of the review paper:

B. Nienhuis, *Critical behavior of two-dimensional spin models and charge asymmetry in the Coulomb gas*, J. Statist. Phys. 34 (1984), 731.

Problems

1. Correlator of the Ising model

Consider the following correlator of the two-dimensional Ising model

$$H(\eta, \bar{\eta}) \;=\; \langle \sigma(\infty)\epsilon(1,1)\epsilon(\eta,\bar{\eta})\sigma(0,0)\rangle.$$

Use the modified Coulomb gas to show that it is given by

$$H(\eta, \bar{\eta}) \;=\; \frac{1}{4}\left|\frac{\eta+1}{\eta^{1/2}(1-\eta)}\right|.$$

2. Structure constants

Use the modified Coulomb gas to compute the correlation functions of the tricritical Ising model. Determine the values of the structure constants given in the text.

3. Vacua of the multicritical Ising model

Consider the potential of the multicritical Ising model

$$V(\varphi) \;=\; g_1\varphi + g_2\varphi^2 + g_3\varphi^3 + g_4\varphi^4 + g_5\varphi^5 + g_6\varphi^6 + \varphi^8.$$

a Show that, by fine tuning the parameters, the model has a phase with four degenerate vacua.

b Argue that this is enough information to conclude that the universality class of this model does not coincide with that of the three-state Potts model, although the two models share the same value of the central charge.

Part IV

Away from Criticality

This page intentionally left blank

15
In the Vicinity of the Critical Points

Lume v'è dato a bene e a malizia.

Dante Alighieri

15.1 Introduction

In the previous chapters we have dedicated ample space to the study of two-dimensional statistical systems at criticality, providing their exact solutions in terms of conformal field theories. In this chapter we start investigating the deformations of conformal field theories that move the statistical systems away from criticality. As pointed out in Chapter 8, in the renormalization group approach the characterization of the universality classes must include, in addition to the conformal theory of the fixed points, also the description of the *scaling region* nearby.

The scaling region is a multidimensional space, parameterized by the coupling constants of the relevant scalar fields $\Phi_{\Delta_i,\Delta_i}(x)$ that are present in the conformal field theory of the fixed point under scrutiny. These operators are identified by the condition $x_i = 2\Delta_i < 2$. The fixed point action is unstable for the insertion of these operators, and any renormalization group flow that starts from a given fixed point can be obtained by a combination of the couplings of these relevant fields. If $\mathcal{S}^*$ is the conformal action of the fixed point and n is the number of its relevant fields, the most general deformation is given by

$$\mathcal{S} = \mathcal{S}^* + \sum_{i=1}^{n} \lambda_i \int \varphi_i(x)\, d^2x. \tag{15.1.1}$$

As discussed in Section 15.2, for what concerns the ultraviolet divergences encountered in the perturbative series of the theory (15.1.1), the quantum field theories defined by the relevant deformations of a conformal action are of the super-renormalizable type. In other words, the relevant operators do not influence the short-distance behavior of the system but, on the contrary, they drastically change the large-distance scales. The first effect of their presence is the breaking of the conformal invariance and the generation of a mass scale, a function of the coupling constants. The latter are, in fact, dimensional quantities, expressed in terms of a mass scale M by

$$M = \mathcal{D}_i\, \lambda_i^{\frac{1}{2-2\Delta_i}}, \tag{15.1.2}$$

where the coefficients $\mathcal{D}_i$ are pure numbers that can be fixed once we choose a normalization of the operators. In the following we adopt the conformal normalization, identified by the short-distance behavior of their two-point correlation function

$$\langle \varphi_i(r)\varphi_j(0)\rangle \simeq \frac{\delta_{ij}}{r^{2x_j}}, \qquad r \to 0. \tag{15.1.3}$$

Excluding the possibility of pathological cases, such as for instance the presence of limit cycles of the renormalization group, there are in general two different physical scenarios associated to the action (15.1.1):

1. In the first scenario, the final point of the renormalization group flow is also a fixed point associated to another conformal field theory. In this case, the quantum field theory associated to this RG flow has an ultraviolet behavior ruled by the conformal field theory $\mathcal{C}FT_1$ of the starting point, while its infrared behavior is controlled by the conformal field theory $\mathcal{C}FT_2$ of the final point. The occurrence of this scenario can be detected by studying the behavior of the two-point correlation functions $G_i(r) = \langle \varphi_i(r)\varphi_i(0)\rangle$: in this case they present a power law behavior in both regimes $r \to 0$ and $r \to \infty$

$$G_i(r) = \begin{cases} r^{-2x_i^{(1)}}, & r \to 0 \\ r^{-2x_i^{(2)}}, & r \to \infty \end{cases} \tag{15.1.4}$$

   with $x_i^{(1)} \neq x_i^{(2)}$. These two quantities are the anomalous dimensions of the field φ_i with respect to the initial and final conformal field theories respectively.

   Quantum field theories of this type have massless excitations, i.e. the physical correlation length of the problem is *infinite* all along the RG flow. However the conformal invariance is broken for the non-vanishing values of the $\beta_i(\{\lambda_j\})$ functions of the coupling constants, as we shall see in the following sections.
2. In the second scenario, which is by far the most common one, the system presents a *finite* correlation length ξ. In this case, the infrared behavior of the theory is ruled by a massive quantum field theory. Once again, the identification of this circumstance can be done by looking at the two-point correlation functions: in this case, for $r \to \infty$ they present an exponential decay while for $r \to 0$ thay have a power law behavior, determined by the initial conformal theory $\mathcal{C}FT_1$

$$G_i(r) = \begin{cases} r^{-2x_i}, & r \to 0 \\ e^{-m_i r}, & r \to \infty. \end{cases} \tag{15.1.5}$$

   In this expression $m_i = \xi^{-1}$ is the mass of the lightest particle that couples to the field φ_i.

From a geometrical point of view, the nature of the renormalization group flows in the multidimensional coupling constant space is show in Fig. 15.1.

The analysis of the off-critical theories poses a series of interesting questions, such as:

- Is there a way to predict whether a deformation of a conformal action gives rise to a massless or a massive theory?

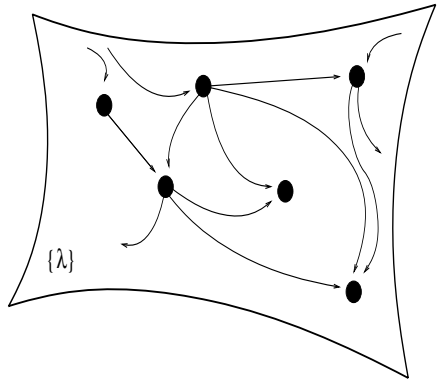


Fig. 15.1 *Renormalization group flows in the coupling constant space.*

- If the off-critical theory is massive, is it possible to determine its mass spectrum?
- Is it possible to characterize the operator content of the off-critical theory and its correlation functions?
- Do the off-critical correlation functions satisfy differential equations? Of what kind?
- Is it possible to determine the thermodynamics of these models?
- What are the relationships between the conformal data – such as, central charge, anomalous dimensions, and structure constants – and the off-critical data, such as the mass spectrum?

Presently there is no general answer to all these questions. However there is a series of important results that permit us to reach satisfactory control of the off-critical theories, at least from a perturbative point of view. It should be pointed out that the situation can be undoubtedly better for particular deformations: as we will see in the following chapters, certain off-critical theories are in fact severely constrained by the presence of infinite conserved charges. These theories correspond to integrable models that can be exactly solved by a formalism based on the S-matrix. Their study turns out to be decisive to solve some of the above-mentioned questions.

In this chapter we initially study the nature of the perturbative series associated to the perturbed action, reformulating the renormalization group equations that they give rise to. Later we discuss two general results of the RG flows, known as the *c-theorem* and Δ-theorem, that permit us to obtain extremely useful information on the theories of the initial and final fixed points.

15.2 Conformal Perturbation Theory

In the vicinity of the critical point, the action of the theory can always be expressed as (15.1.1). The unperturbed action corresponds to the conformal field theory of the fixed point, of which we know in principle all correlation functions. This allows us to define the perturbative series for any physical quantity away from criticality. For

simplicity, hereafter, we consider the case in which the deformation is made only by one relevant scalar field $\varphi(x)$ of conformal weights (Δ, Δ). The expectation value of any operator A of the perturbed theory is expressed by the series

$$\begin{aligned} \langle A \rangle_\lambda &= \frac{1}{Z_\lambda} \left\langle A \exp\left[-\lambda \int d^2x\, \varphi(x) \right] \right\rangle_0 \\ &= \frac{1}{Z_\lambda} \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} \int d^2x_1 \dots d^2x_n \, \langle A\, \varphi(x_1) \dots \varphi(x_n) \rangle_0 \end{aligned} \tag{15.2.1}$$

where

$$Z_\lambda = \left\langle \exp\left[-\lambda \int d^2x \varphi \right] \right\rangle_0 . \tag{15.2.2}$$

In this expression, $\langle \cdots \rangle_0$ are the correlation functions of the unperturbed conformal field theory. In the computation of the integrals of the perturbative series, there are, however, both ultraviolet and infrared divergences. They can be regularized by introducing an ultraviolet cut-off ϵ and an infrared cut-off R – quantities that finally have to be sent to the limits $\epsilon \to 0$ and $R \to \infty$.

For what concerns the ultraviolet properties of the perturbative series, the quantum field theories defined by deformations of the relevant fields are of the super-normalizable type. Therefore the ultraviolet divergences can be dealt with the standard renormalization methods: they lead to a redefinition of the local fields and, when the deformation operator is marginal (i.e. with conformal weight $\Delta = 1$), also to a renormalization of the couling constant.

The infrared divergences are of different type. They cannot be absorbed in the redefinition of the local quantities and, for this reason, they give rise to non-analytic expressions in the coupling constants. The physical origin of this phenomenon is easy to understand. In fact, the vacuum state of the deformed theory (as well as all other excited states) is not adiabatically related to the vacuum states of the conformal theory:[1] if, for instance, the perturbed system corresponds to a massive theory, the new Hilbert space is set by the Fock space of the multiparticle states, whereas the original Hilbert space is spanned by the Verma modules of the conformal states. In particular, the vacuum of the perturbed theory is the state without any particle excitations, whereas the vacuum of the unperturbed theory is characterized in a completely different way, since it is the state annihilated by all L_n with $n \geq -1$.

The different nature of the ultraviolet and infrared divergences permits their separate treatment, providing the key to controlling the theory perturbatively. Let's first discuss the ultraviolet properties and later the infrared ones.

Ultraviolet divergences. To understand the ultraviolet structure of the theory, let's consider initially what we can learn from the first-order calculation. Let $\Phi(0)$ be a field of the perturbed theory (to become later a renormalized field), obtained as a

[1] One can draw an analogy with a quantum mechanics problem. Consider a free particle system on a line, in which we switch on a potential like $g|x|$ that cuts off the free asymptotic states. The perturbed system has all and only bound states that are not adiabatically related to the energy eigenstates of the unperturbed system. In particular, their energies scale as a function of the coupling constant according to the non-analytic law $g^{2/3}$.

deformation of the field $\tilde{\Phi}(0)$ of the original conformal theory. Denote by X a generic product of other fields and consider the correlator $\langle X\Phi(0)\rangle$. Its perturbative definition is given by

$$\langle X\,\Phi(0)\rangle_\lambda \;\simeq\; \langle X\,\tilde{\Phi}(0)\rangle_0 - \lambda \int_{\epsilon<|x|<R} d^2x\,\langle X\,\tilde{\Phi}(0)\varphi(x)\rangle_0 + \cdots \tag{15.2.3}$$

This integral is ultraviolet divergent only if the operator product expansion

$$\varphi(x)\,\tilde{\Phi}(0) \;=\; \sum_k C^k_{\varphi\Phi}\,|x|^{2(\Delta_k-\Delta_\Phi-\Delta)}\,A_k(0) \tag{15.2.4}$$

contains the fields A_k with conformal weights Δ_k that fulfill the equation

$$\gamma_k \;\equiv\; \Delta_k - \Delta_\Phi - \Delta + 1 \le 0. \tag{15.2.5}$$

In this case, to obtain a finite expression at the first-order of the correlation functions it is sufficient to define the renormalized operator by

$$\Phi \;=\; \tilde{\Phi} + \lambda \sum_k b_k\,\epsilon^{2\gamma_k} A_k + \mathcal{O}(\lambda^2), \tag{15.2.6}$$

where

$$b_k \;=\; \pi\,\frac{C^k_{\varphi\Phi}}{\gamma_k}.$$

This formula shows that, in general, the renormalization procedure induces a mixing of the original operators with those of lower conformal weights.

Off-critical operator product expansion. Let's now analyze in more detail the renormalization procedure of the ultraviolet and infrared divergences by studying the two-point correlation function of the renormalized field $\Phi(x)$, whose perturbative expression is given by

$$\langle\Phi(x)\Phi(0)\rangle \;=\; \sum_{k=0}^{\infty} \frac{(-\lambda)^n}{n!}\;\langle\tilde{\Phi}(x)\tilde{\Phi}(0)\varphi(y_1)\ldots\varphi(y_n)\rangle_0\;d^2y_1\ldots d^2y_n. \tag{15.2.7}$$

To evaluate the behavior of this correlator (at least in the limit $|x|\to 0$, even though the final expression also holds for finite values of x), it is convenient to start from the operator product expansion

$$\Phi(x)\Phi(0) \;=\; \sum_k C^k_{\Phi\Phi}(x)\,A_k(0), \tag{15.2.8}$$

where A_k $(k = 0, 1, \ldots)$ is a complete set of fields and $C^k_{\Phi\Phi}(x)$ are their relative structure constants with the field Φ. Since the structure constants are local quantities, they are analytic functions of the coupling constant λ and therefore can be expanded in power series theoreof. The fields A_k are the renormalized expressions of the operators

$\tilde{A}_k$ present at the critical point: denoting by $(\Delta_k, \bar{\Delta}_k)$ their conformal weights, on the basis of a simple dimensional analysis argument, we have

$$C^k_{\Phi\Phi}(x) \;=\; x^{\Delta_k - 2\Delta_\Phi}\; \bar{x}^{\bar{\Delta}_k - 2bar\Delta_\Phi} \sum_{n=0}^{\infty} C^{k\,(n)}_{\Phi\Phi}\,(\lambda\, r^{2-2\Delta})^n, \tag{15.2.9}$$

where $r = (x\,\bar{x})^{1/2}$. However, the price to pay for the analyticity of the structure constants is the presence of the vacuum expectation values of some of the operators A_k. These are non-local quantities and therefore non-analytic with respect to the coupling constant. For dimensional reason, their expression is given by

$$\langle A_k(0)\rangle_\lambda \;=\; \lambda^{\frac{\Delta_k}{1-\Delta}}\; Q_k, \tag{15.2.10}$$

where Q_k are pure numbers. The set of these vacuum expectation values encodes important information on the infrared properties of the theory and cannot be determined by perturbation theory. Obviously many of them vanish for symmetry reasons as, for instance, those of fields with spin or those of the derivative fields. Selecting only those with vacuum expectation value different from zero, we have

$$\langle \Phi(x)\Phi(0)\rangle_\lambda \;=\; \sum_\nu C^\nu_{\Phi\Phi}(x)\,\langle A_\nu(0)\rangle_\lambda. \tag{15.2.11}$$

It is worth mentioning that the exact vacuum expectation values of the primary fields can be computed for integrable deformations of conformal theories and the relevant formulas will be presented in Chapter 20. Together with the perturbative expressions of the structure constants, discussed below, the vacuum expectation values permit the determination of the correlation functions.

Renormalization. Consider now the renormalization of these fields by analyzing the matrix elements

$$\begin{aligned} \tilde{I}^k_l(\lambda, R, \epsilon) \;&=\; \langle \tilde{A}^l(\infty)\tilde{A}_k(0)\rangle \qquad (15.2.12)\\ &=\; \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} \int_{\epsilon<|y_i|<R} \langle \tilde{A}^k(\infty)\,\tilde{A}_l(0)\,\varphi(y_1)\ldots\varphi(y_n)\rangle_0\; d^2y_1\ldots d^2y_n. \end{aligned}$$

By means of the ultraviolet and infrared cutoffs, all these integrals are finite and the quantities $\tilde{I}^k_l(\lambda, R, \epsilon)$ are regular functions of λ. Adopting the conformal normalization of the fields, we have

$$\tilde{I}^k_l(\lambda, R, \epsilon) \;=\; \delta^k_l + \mathcal{O}(\lambda). \tag{15.2.13}$$

From the invariance under rotations, the matrix $\tilde{I}^k_l$ is diagonal in the spin of the fields and it is therefore possible to consider any spin sector separately. For simplicity we discuss only the spin zero sector, that of the scalar fields.

In the limit $\epsilon \to 0$, the matrix elements $\tilde{I}^k_l(\lambda, R, \epsilon)$ become singular and it is natural to assume that they have a factorized form

$$\tilde{I}^k_l(\lambda, R, \epsilon) \;=\; \sum_{k'} U^k_{k'}(\lambda, \epsilon)\, I^{k'}_l(\lambda, R) \tag{15.2.14}$$

where $I_l^{k'}(\lambda, R)$ are the elements of the renormalized matrix whereas $U_{k'}^k(\lambda, \epsilon)$ are the elements of the renormalization matrix. Both are regular functions of λ. For dimensional reasons, we have

$$U_l^k(\lambda, \epsilon) \;=\; \sum_{n=0}^{\infty} \frac{U_l^{k\,(n)} (\lambda\, \epsilon^{2-2\Delta})^n}{\epsilon^{2(\Delta_l - \Delta_k)}}. \tag{15.2.15}$$

When $\epsilon \to 0$, it is necessary to keep only the terms with negative powers ϵ. For $\Delta < 1$, there is only a finite number of them. Organizing the fields A_i in increasing order of the conformal weights $\Delta_0 \leq \Delta_1 \leq \Delta_2 \ldots$, the matrix U_l^k assumes a triangular form, i.e.

$$U_l^k(\lambda, \epsilon) \;=\; 0, \qquad \Delta_k > \Delta_l. \tag{15.2.16}$$

The inverse matrix of U obviously has the same properties (15.2.15) and (15.2.16) as U and we can then define the renormalized fields as

$$A_k \;=\; \sum_l (U^{-1})_k^l\, \tilde{A}_l. \tag{15.2.17}$$

With the normalization $U_l^k(\lambda, \epsilon) \;=\; \delta_l^k + \mathcal{O}(\lambda)$, for the renormalized fields we have

$$A_k \;=\; \tilde{A}_k + \cdots \tag{15.2.18}$$

with a finite number of other terms, corresponding to the operators with lower conformal weights. In this way, we recover the result previously obtained to first order.

It is necessary to stress that the analysis done above can become more involved in the presences of "resonances" of the conformal weights, namely if it happens that $\Delta_k - \Delta_l \;=\; n(1 - \Delta)$. In this case, there are also logarithmic divergences and the factorized form (15.2.14) becomes ambiguous since it depends on an arbitary renormalization point. Although this is an interesting question, it is however outside the scope of the present analysis.

The elements of the renormalized matrix

$$I_l^k(\lambda, R) \;=\; \langle \tilde{A}^k(\infty) A_l(0) \rangle_\lambda, \tag{15.2.19}$$

are independent of the ultraviolet cut-off ϵ and they have the same structure as the matrix U

$$I_l^k(\lambda, R) \;=\; \sum_{n=0}^{\infty} \frac{I_l^{k\,(n)}\, (\lambda\, R^{2-2\Delta})^n}{R^{2(\Delta_l - \Delta_k)}}, \tag{15.2.20}$$

where, now, it is necessary to keep the terms with positive powers of R. The sum of this series produces a non-trivial dependence on R. Although it is not easy to find their exact expression, it is natural to assume that these series behave in a homogeneous way: this means that there exists the limit

$$\lim_{R \to \infty} \frac{I_l^k(\lambda, R)}{I_0^k(\lambda, R)} \;=\; \lambda^{\frac{\Delta_k}{1-\Delta}}\, Q_l^{(k)} \tag{15.2.21}$$

and that this limit corresponds to the vacuum expectation values of the operators.

It is also convenient to define the quantities

$$
\begin{aligned}
G^k_{\Phi\Phi}(\lambda, x, R) \;&=\; \langle \tilde{A}^k(\infty)\, \Phi(x)\, \Phi(0)\rangle_\lambda \qquad (15.2.22)\\
&=\; \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} \int_{|y_i|<R} \langle \tilde{A}^k(\infty)\, \Phi(x)\, \Phi(0)\, \varphi(y_1)\dots\varphi(y_n)\rangle_0 \; d^2y_1 \dots d^2y_n.
\end{aligned}
$$

Since Φ are the renormalized fields, all the integrals are ultraviolet convergent and it is not necessary to introduce the ultraviolet cut-off ϵ. Substituting this expression in the operator expansion (15.2.8), it yields

$$
C^k_{\Phi\Phi}(x) \;=\; \sum_l G^l_{\Phi\Phi}(\lambda, x, R)\, (I^{-1})^k_l(\lambda, R). \qquad (15.2.23)
$$

Structure constants. The quantity above is now finite in the limit $R \to \infty$ and allows us to compute the structure constants of the renormalized fields. As an explicit example, we present here their first-order term. Let $\mathbf{C}^k_{\Phi\Phi}$, $\mathbf{C}^k_{\varphi,\Phi}$, and $\mathbf{C}^k_{\varphi\, l}$ be the structure constants of the conformal theory, namely

$$
\begin{aligned}
\langle \tilde{A}^k(\infty)\Phi(x)\Phi(0)\rangle_0 &= \mathbf{C}^k_{\Phi\Phi}\, (x\,\bar{x})^{\Delta_k - 2\Delta_\Phi)}\\
\langle \tilde{A}^k(\infty)\varphi(x)\Phi(0)\rangle_0 &= \mathbf{C}^k_{\varphi\Phi}\, (x\,\bar{x})^{\Delta_k - \Delta - \Delta_\Phi)}\\
\langle \tilde{A}^k(\infty)\varphi(x)\tilde{A}_l(0)\rangle_0 &= \mathbf{C}^k_{\varphi\, l}\, (x\,\bar{x})^{\Delta_k - \Delta_l - \Delta)}.
\end{aligned} \qquad (15.2.24)
$$

At first order, we have

$$
\tilde{I}^k_l(\lambda, R, \epsilon) \;=\; \delta^k_l - \lambda\,\pi\, \mathbf{C}^k_{\varphi\, l}\, \frac{R^{2(\Delta_k - \Delta_l - \Delta + 1)} - \epsilon^{2(\Delta_k - \Delta_l - \Delta + 1)}}{\Delta_k - \Delta_l - \Delta + 1}, \qquad (15.2.25)
$$

and, therefore, at the same perturbative order in λ

$$
\begin{aligned}
I^k_l(\lambda, R) \;&=\; \delta^k_l - \lambda\,\pi, \mathbf{C}^k_{\varphi\, l}\, \frac{R^{2(\Delta_k - \Delta_l - \Delta + 1)}}{\Delta_k - \Delta_l - \Delta + 1}\\
U^k_l(\lambda, \epsilon) \;&=\; \delta^k_l + \lambda\,\pi, \mathbf{C}^k_{\varphi\, l}\, \frac{\epsilon^{2(\Delta_k - \Delta_l - \Delta + 1)}}{\Delta_k - \Delta_l - \Delta + 1}.
\end{aligned} \qquad (15.2.26)
$$

Hence, at first order in λ the structure constants are given by

$$
\begin{aligned}
C^k_{\Phi\Phi}(x) &= \mathbf{C}^k_{\Phi\Phi}\, (x\,\bar{x})^{\Delta_k - 2\Delta_\Phi} \qquad (15.2.27)\\
&\quad -\lambda \int_{|y|<R} \langle \tilde{A}^k(\infty)\, \varphi(y)\Phi(x)\Phi(0)\rangle_0 \, d^2y + \lambda\,\pi \sum_l \frac{\mathbf{C}^l_{\Phi\phi}\, \mathbf{C}^k_{\varphi\, l}\, R^{2(\Delta_k - \Delta_l - \Delta + 1)}}{\Delta_k - \Delta_l - \Delta + 1}.
\end{aligned}
$$

Substituting the operator expansion (15.2.8), it is easy to see that the last term of the previous expression is the one that cancels the infrared divergences, so that the final

expression that is finite at the first perturbative order is

$$C^k_{\Phi\Phi}(x) = r^{2(\Delta_k - 2\Delta_\Phi)} \times \left(\mathbf{C}^k_{\Phi\Phi} + \lambda\pi r^{2-2\Delta} \sum_l \left[\frac{\mathbf{C}^l_{\Phi\phi}\mathbf{C}^k_{\varphi\, l}}{\Delta_k - \Delta_l - \Delta + 1} - \frac{\mathbf{C}^l_{\varphi\Phi}\mathbf{C}^k_{\Phi\, l}}{\Delta_l - \Delta_\Phi - \Delta + 1} \right] + \mathcal{O}(\lambda^2) \right). \tag{15.2.28}$$

It should be said that this series is not particularly convenient from a practical point of view. The most efficient way to compute the first-order correction of the structure constants is in fact through the integral

$$C^{k\,(1)}_{\Phi\Phi}(r) \;=\; -\lambda \int' \langle \tilde{A}^k(\infty)\, \varphi(y)\Phi(x)\Phi(0)\rangle_0 \; d^2y, \tag{15.2.29}$$

where the index in the integral means that one has to neglect all divergent terms that appear in the limit $R \to \infty$. This is equivalent to regarding the integral as an analytic expression of the conformal weights of the fields. For instance, the integral of the three-point correlation function of the fields Φ_i, Φ_j, and Φ_k of conformal weights Δ_i, Δ_j, and Δ_k produces the analytic function

$$J_{ijk}(x) = \int' d^2y\, \langle \Phi_i(x)\Phi_j(0)\Phi_k(y)\rangle_0 \;=\; \mathbf{C}_{ijk}\, r^{2(1-\Delta_i-\Delta_j-\Delta_k)} \times 2\pi\, \frac{\Gamma(\Delta_i - \Delta_j - \Delta_k + 1)\,\Gamma(\Delta_j - \Delta_i - \Delta_k + 1)\,\Gamma(2\Delta_k - 1)}{\Gamma(2 - 2\Delta_k)\,\Gamma(\Delta_i + \Delta_k - \Delta_j)\,\Gamma(\Delta_j + \Delta_k - \Delta_i)}. \tag{15.2.30}$$

15.3 Example: The Two-point Function of the Yang–Lee Model

A simple application of the formalism developed in the previous section is the computation of the off-critical correlation function of the Yang–Lee model (see Section 14.5). Suppose we perturb the CFT action $\mathcal{S}_0$ of this model by adding the perturbation of the field φ with $\Delta = -1/5$

$$\mathcal{S} \;=\; \mathcal{S}_0 + ig \int d^2x\, \varphi(x).$$

For this model a sensible QFT is obtained if the coupling constant is purely imaginary as shown above, with $g > 0$. The off-critical two-point function of the field $\varphi(x)$ can be written as

$$G(r) \;=\; \langle \varphi(x)\varphi(0)\rangle \;=\; C^I_{\varphi\varphi}(r)\,\langle I\rangle + C^{\varphi}_{\varphi\varphi}(r)\,\langle \varphi(0)\rangle + C^{\bar{T}T}_{\varphi\varphi}(r)\,\langle :\bar{T}T:(0)\rangle + \cdots \tag{15.3.1}$$

The structure constants have a regular perturbative expansion that, taking into account the conformal weight of the field φ, is given by

$$\begin{aligned} C^I_{\varphi\varphi} \;&=\; r^{4/5}\left(1 + Q^I_1 t + Q^I_2 t^2 + \cdots\right) \\ C^{\varphi}_{\varphi\varphi} \;&=\; \mathbf{C}^{\varphi}_{\varphi\varphi}\left(1 + Q^{\varphi}_1 t + Q^{\varphi}_2 t^2 + \cdots\right) \\ C^{\bar{T}T}_{\varphi\varphi} \;&=\; \frac{1}{121} r^{24/5}\left(1 + Q^{\bar{T}T}_1 t + Q^{\bar{T}T}_2 t^2 + \cdots\right) \end{aligned}$$

where $t = gr^{12/5}$ is the dimensionless coupling constant and $\mathbf{C}^{\varphi}_{\varphi\varphi}$ is the (imaginary) structure constant of the CFT, given in eqn (14.5.8). All coefficients $Q^{\mathcal{O}}_i$ can in principle

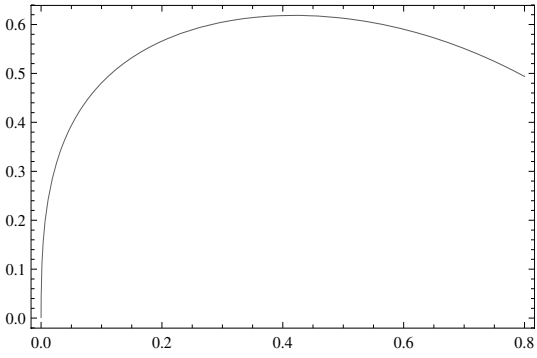


Fig. 15.2 *Plot of* $-G(r)$ *versus the distance* r *at* $g = 1$.

be computed using the perturbative scheme of the previous section: the numerical values of the first ones are

$$\begin{aligned} Q_1^I &= 0.319800\ldots \\ Q_1^\varphi &= 0.02122\ldots \end{aligned}$$

We also need the vacuum expectation value of the various fields, $\langle I\rangle$, $\langle\varphi(0)\rangle$, etc. entering eqn (15.3.1). Here we take into account only $\langle I\rangle$ and the vacuum expectation value of the field φ that can be computed exactly (see Chapter 20) with value

$$\begin{aligned} \langle\varphi\rangle &= ig^{-1/6}\,\frac{5}{24\sqrt{3}}\left(\frac{12}{25}\right)^{5/6}\left(\frac{12}{\pi}\right)^{1/6}\left(\frac{\Gamma\left(\frac{1}{3}\right)}{\Gamma\left(\frac{5}{6}\right)}\right)^2\left(\left|\frac{\Gamma\left(-\frac{1}{5}\right)\Gamma\left(\frac{3}{5}\right)}{\Gamma\left(\frac{2}{5}\right)\Gamma\left(\frac{6}{5}\right)}\right|\right)^{5/12} \\ &= 0.840184\ldots ig^{-1/6}. \end{aligned}$$

Substituting the expressions given above for the structure constants and the vacuum expectation values, one can obtain an estimate of the function $G(r)$ up to order $\mathcal{O}(r^{24/5})$. The plot of this function, with a minus sign in front, is given in Fig. 15.2.

This example is particularly significant to enlighten the role played by the vacuum expectation value of the various fields for the off-critical correlators. Notice that, because of the presence of a non-zero value of $\langle\varphi\rangle$, the two-point correlation function $G(r)$ behaves for small r not as $r^{4/5}$, as one could expect naively from CFT reasons, but rather as $r^{2/5}$. Furthermore, the competition between the terms coming from the families I and φ, which have opposite sign, produces the curve drawn above, which starts to bend for values of $r \simeq \xi$, where ξ is the finite correlation function of the off-critical model, whose exact value can be determined by the thermodynamic Bethe ansatz analysis of Chapter 19:

$$\xi = g^{-5/12}\left(\frac{25}{12}\right)^{5/12}\left(\frac{\pi}{12}\right)^{1/12}\left(\frac{\Gamma\left(\frac{5}{6}\right)}{\Gamma\left(\frac{1}{3}\right)}\right)\left(\left|\frac{\Gamma\left(\frac{2}{5}\right)\Gamma\left(\frac{6}{5}\right)}{\Gamma\left(-\frac{1}{5}\right)\Gamma\left(\frac{3}{5}\right)}\right|\right)^{5/24} = 0.37836\ldots\, g^{-5/12}.$$

15.4 Renormalization Group and β-functions

In this section we reconsider the renormalization group theory in the perspective pursued in this chapter. We also present the computation at lowest orders of the β-functions.

The key ideas of the renormalization group can be expressed as follows. Each theory is described by an action $\mathcal{S}(g, a)$ that is a function of a certain number of dimensionless coupling constants[2] $g = \{g_1, g_2, \ldots\}$ and an ultraviolet cut-off that cures the divergences coming from the short-distance operator expansions. The main hypothesis consists of the existence of a one-parameter family of flows in the manifold G of the coupling constants, $\mathcal{R}_t\, G \to G$, such that the quantum field theory described by the action $\mathcal{S}(\mathcal{R}_t g, e^t a)$ is equivalent to the theory described by the action $\mathcal{S}(g, a)$. In more detail, both theories give rise to the same result over a range of scales $|x| \gg e^t a$. Our main interest here is the two-dimensional theories but the results given below can be easily generalized to quantum field theories in higher dimensions. As a starting point, let's consider the correlation functions of the local fields $A_i(x)$, defined as usual by the functional integral

$$\langle A_1(x_1) \ldots A_n(x_n) \rangle \;=\; \int \mathcal{D}\varphi\, A_1(x_1) \ldots A_n(x_n)\, e^{-\mathcal{S}[\varphi]}, \tag{15.4.1}$$

where we have included in the definition of the action $\mathcal{S}$ an additive constant that ensures the correct normalization of the expression above (this permits us to avoid the introduction of the normalization factor Z^{-1}). Consider now the Ward identity coming from the substitution

$$x^\mu \to x'^\mu \;=\; x^\mu + \epsilon^\mu(x).$$

Since[3]

$$\delta\mathcal{S} \;=\; \int d^2x\, T_{\mu\nu}(x)\, \partial^\mu \epsilon^\nu(x).$$

and

$$A_i(x) \to A_i(x) + \delta A_i(x), \tag{15.4.2}$$

we have

$$\sum_{i=1}^{n} \langle A_1(x_1) \ldots \delta A_i(x_i) \ldots A_n(x) \rangle - \int d^2x \langle T_{\mu\nu}(x) A_1(x_1) \ldots A_n(x_n) \rangle\, \partial^\mu \epsilon^\nu(x) = 0. \tag{15.4.3}$$

For a global dilatation, the variation of the fields is given by

$$\delta A(x) \;=\; \epsilon \left(\frac{1}{2}\, x^\mu\, \partial_\mu + \hat{D} \right) A(x), \tag{15.4.4}$$

[2] Dimensionful couplings can be rescaled by appropriate powers of the cut-off a and expressed by their dimensionless version.

[3] With respect the formula of Chapter 10, we have absorbed a factor $1/(2\pi)$ in the definition of $T_{\mu\nu}$.

where $\hat{D}$ is the operator that implements the internal transformation of the fields under this transformation. Defining the trace of the stress–energy tensor

$$\Theta(x) \equiv T^{\mu}_{\mu}(x) \tag{15.4.5}$$

and substituting the last two equations in (15.4.3) we get

$$\sum_{i=1}^{n} \left\langle \left(\frac{1}{2}\, x^{\mu}\partial_{\mu} + \hat{D}_i \right) A_1 \dots A_n \right\rangle - \int d^2x \langle \Theta(x) A_1(x_1) \dots A_n(x_n) \rangle \;=\; 0. \tag{15.4.6}$$

Let's now discuss the nature of the action. We assume that it is given by an integral over a local expression of the fields

$$\mathcal{S}_g \;=\; \int d^2x\, \mathcal{L}_g, \tag{15.4.7}$$

which depends on a certain number of dimensionless couplings $\{g_a\} = \{g_1, g_2, \dots\}$. For instance, the perturbed conformal field theories (15.1.1) have, as coupling constants, those of the relevant fields, rescaled by the ultraviolet cut-off, i.e. $g_i \;=\; a^{2(1-\Delta_i)}\, \lambda_i$. From the property of $\mathcal{L}$, the derivative

$$\varphi_i(x) \;=\; \frac{\partial \mathcal{L}}{\partial g_i} \tag{15.4.8}$$

is a local field of the theory. If we consider only homogeneous and isotropic interactions, these are scalar fields and the space of all of them, $\mathcal{O}^{(0)}$, may be regarded as the tangent space to G at the point $\{g_1, g_2, \dots\}$. The fields $A_i(x)$ can also depend on the coupling constants and their variation, varying the couplings, is expressed in terms of an operator $\hat{B}_i$

$$\hat{B}_k\, A_i(x) \;=\; \frac{\partial}{\partial g_k} A_i(x). \tag{15.4.9}$$

The necessity to introduce a coupling dependence of the fields is obvious in view of the ultraviolet divergences and the implementation of their renormalization, as discussed in the previous section. For the correlation function (15.4.1) this implies

$$\begin{aligned} \frac{\partial}{\partial g_a} \langle A_1(x_1) \dots A_n(x_n) \rangle \;=\; & \sum_{i=1}^{n} \langle A_1(x_1) \dots \hat{B}_a A_i(x_i) \dots A_n(x_n) \rangle \\ & - \int d^2x\, \langle\, \varphi_a(x) A_1(x_1) \dots A_n(x_n) \rangle. \end{aligned} \tag{15.4.10}$$

The trace of the stress–energy tensor belongs to the space $\mathcal{O}^{(0)}$ and can be expressed in terms of the fields of this space as

$$\Theta(x) \;=\; \sum_{a} \beta^a(g)\, \varphi_a(x). \tag{15.4.11}$$

The coefficients $\beta_a(g)$ are the β-functions of the theory. To see this, note that if dt is the infinitesimal parameter of the dilatation, $x_\mu \to (1 + dt) x_\mu$, the trace $\Theta(x)$ can be

also defined as

$$\Theta(x) \;=\; \frac{d\mathcal{L}}{dt}, \tag{15.4.12}$$

an expression that can be written as

$$\Theta(x) \;=\; \sum_a \frac{\partial \mathcal{L}}{\partial g_a}\,\frac{\partial g_a}{\partial t} \;=\; \sum_a \frac{\partial g_a}{\partial t}\,\varphi_a(x). \tag{15.4.13}$$

Hence the β_a-functions express the variation of the coupling constant under a change of the length-scale

$$\beta^a(\{g\}) \;=\; \frac{\partial g_a}{\partial t}. \tag{15.4.14}$$

Combining now eqn (15.4.10) with eqn (15.4.11) and using the definition of the β_a functions, we arrive at the celebrated Callan–Symanzik equation

$$\begin{aligned}\sum_{i=1}^{n} \left\langle \left(\frac{1}{2}\, x_i^\mu \frac{\partial}{\partial x_i^\mu} + \hat{\gamma}^{(i)}(g) \right) A_1(x_1) \ldots A_n(x_n) \right\rangle \\ - \sum_a \beta^a(g)\, \frac{\partial}{\partial g_a} \langle A_1(x_1) \ldots A_n(x_n) \rangle \;=\; 0,\end{aligned} \tag{15.4.15}$$

where the linear operators $\hat{\gamma}^{(i)}(g)$, defined by

$$\hat{\gamma}(g) \;=\; D + \beta^a(g) B_a, \tag{15.4.16}$$

act on the fields $A_i(x)$. The operator $\hat{\gamma}$ is the matrix of the anomalous dimensions. Since the stress–energy tensor is a conserved quantity, it does not renormalize and therefore its anomalous dimension coincides with the canonical one

$$\hat{\gamma}(g)\Theta \;=\; 2\,\Theta. \tag{15.4.17}$$

In this way we easily obtain the action of $\hat{\gamma}(g)$ on the fields of the basis in the space $\mathcal{O}^{(0)}$

$$\hat{\gamma}(g)\,\varphi_a \;\equiv\; \gamma_a^b(g)\,\varphi_b \;=\; \left(2\delta_a^b - \frac{\partial \beta^b}{\partial g^a} \right) \varphi_b. \tag{15.4.18}$$

In the renormalized theory, every β_a-function and any matrix element of the operator $\hat{\gamma}$ do not depend on the initial point of the renormalization group $\mathcal{R}_0$. In particular, two field theories corresponding to coupling constants $g(t_1)$ and $g(t_2)$ that belong to the same integrated curve of their evolution equation

$$dg^a \;=\; \beta^a(g)\, dt, \tag{15.4.19}$$

differ only by a scale transformation of the length-scale $x_\mu \to e^{(t_1 - t_2)}\, x_\mu$. Therefore the scaling properties of the theory only depend on the vector fields $\beta^a(g)$. The simplest and, at the same time, the most important characteristic of these functions is associated to the fixed points g^* that satisfy the equation $\mathcal{R}_t g^* = g^*$. These are the points where the β-functions vanish, $\beta^a(g^*) \;=\; 0$. These conditions identify the critical points

of the system, where the correlation length diverges. If, in addition to the location of the fixed points, we also know the derivatives of the β at these points, eqn (15.4.18) allows us to compute the anomalous dimensions of the various operators. We will see an interesting example of this formalism in the next section.

Let's now discuss how to determine the first terms of the $\beta_a(g)$-functions by using the perturbative expansion of the partition function of the perturbed conformal theories, with an action given in eqn (15.1.1). It is convenient to take care firstly of the dimensionality of the coupling constants: introducing a as a length-scale and simultaneously as ultraviolet cut-off, they can be expressed as $\lambda_i = g_i a^{-2(1-\Delta_i)}$, where g_i are now the dimensionless couplings. In terms of these quantities, the perturbative expansion of Z is given by

$$Z = \int \mathcal{D}\varphi \, \exp\left[-\mathcal{S}^* - \sum_i g_i \int \frac{d^2x}{a^{2(1-\Delta_i)}} \varphi_i(x)\right] = Z^* \left[1 - \sum_i g_i \int \frac{d^2x}{a^{2(1-\Delta_i)}} \langle \varphi_i(x)\rangle + \frac{1}{2}\sum_{i,j} g_i\, g_j \int_{|x_1-x_2|>a} \langle \varphi_i(x_1)\varphi_j(x_2)\rangle \frac{d^2x_1}{a^{2(1-\Delta_i)}} \frac{d^2x_2}{a^{2(1-\Delta_j)}} + \cdots\right].$$

To find the β-functions, we shall address the question of how to change the coupling constants g_i under the infinitesimal scale transformation $a \to (1+\delta t)a$ in such a way that the partition function remains invariant.

Observe that, in the perturbative expansion of Z, the length-scale a appears both explicitly (in the factors $a^{2-2\Delta_i}$ of the denominators) and implicitly, as a cut-off of the integrals. If the rescaling of a is done with the infinitesimal parameter δt, there is an additive effect of the different dependences in the computation of the β-functions. Let's consider then the two different terms separately.

The first dependence of the coupling constants from a is simple to take into account. In fact, a change of a is compensated by the substitution

$$g_i \to (1+\delta t)^{2(1-\Delta_i)}\, g_i \simeq g_i + 2(1-\Delta_i) g_i\, \delta t. \tag{15.4.20}$$

The effect of a change of the cut-off in the integrals can instead be estimated by the operator product expansion. Consider, for instance, the second perturbative order. The integral, after a rescaling of a, can be written as

$$\int_{|x_1-x_2|>a(1+\delta t)} [\cdots] = \int_{|x_1-x_2|>a} [\cdots] - \int_{a<|x_1-x_2|<a(1+\delta t)} [\cdots]. \tag{15.4.21}$$

The first terms produces the original contribution in Z, and the second term can be computed through the operator expansion of the conformal theory

$$\sum_k \mathbf{C}_{ijk} a^{2(\Delta_k-\Delta_i-\Delta_j)} \int_{a<|x_1-x_2|<a(1+\delta t)} \langle \varphi_k(x_2)\rangle \frac{d^2x_1}{a^{2(1-\Delta_i)}} \frac{d^2x_2}{a^{2(1-\Delta_j)}}. \tag{15.4.22}$$

The integral over x_1 gives the area of the infinitesimal annulus, i.e. $2\pi a^2\, \delta t$. Taking into account the powers of a and the original negative sign of this contribution, see

eqn (15.4.21), in the partition function the term above gives

$$-\pi\delta t \sum_{i,j,k} \mathbf{C}_{ijk}\, g_i\, g_j \int \langle \varphi_k(x) \rangle \frac{d^2x}{a^{2(1-\Delta_k)}}. \tag{15.4.23}$$

The presence of this term in Z can then be compensated by a redefinition of the coupling constant g_k

$$g_k \to g_k - \pi \sum_{i,j} \mathbf{C}_{ijk} g_i\, g_j\, \delta t. \tag{15.4.24}$$

Gathering together the two contributions given in eqns (15.4.20) and (15.4.24), we arrive at the first terms of the β-functions

$$\frac{dg_k}{dt} \equiv \beta_k(g) = 2(1-\Delta_k)\, g_k - \pi \sum_{i,j} \mathbf{C}_{ijk}\, g_i\, g_j + \cdots \tag{15.4.25}$$

A few comments are in order:

- The β-functions rule the ultraviolet behavior of the field theories and therefore it is not surprising that their first coefficients are expressed by the conformal data, such as the conformal weights and structure constants of the primary fields. Note that when the coupling constant g_k corresponds to a relevant operator with $\Delta_k < 1$, the first term is responsible for the repulsive nature of the fixed point at the origin, $g_i = 0$. Vice versa, if the coupling constant corresponds to an irrelevant operator with $\Delta_k > 1$, the origin becomes an attractive fixed point.
- The formula (15.4.25) can be generalized to higher dimensional field theories, as long as we know the structure constants and the anomalous dimensions of the operators.
- The higher order terms of the β-functions can be computed, in principle, iterating the argument given above. It is easy, if fact, to see the iterative nature of the renormalization procedure: the terms of order $g_i^{n_i}\, g_j^{n_j}\, g_k^{n_k} \cdots$ influence the renormalization of the terms $g_i^{n_i-1}\, g_j^{n_j-1}\, g_k^{n_k+1} \cdots$. However their explicit computation soon becomes involved and will not be discussed here.
- Focusing attention only on the first perturbative terms of the β-functions given in eqn (15.4.25), it is important to establish the range of validity of these expressions. We expect that we can trust them when the coupling constants are those of the quasi-marginal operators, for which $|\epsilon_i| \sim \epsilon \ll 1$, where $\epsilon_i \equiv (1-\Delta_i)$. In these cases, in fact, the nonlinear terms become comparable with the linear term when $g_i \sim \epsilon$. This process can give rise to new fixed points in the region $g_i \sim \epsilon$, which is a range compatible with the perturbative nature of the approach itself. In the next section we will present an interesting realization of this situation.

  When the coupling constant corresponds to a strongly relevant operator, the eventual fixed points are localized at a finite distance from the origin, i.e. outside the perturbative regime of the formalism. In this case, the evolution equations of the coupling constants give, in general, only a qualitative indication of the renormalization group flows and caution must be used in extracting quantitative predictions.

- It is interesting to observe that, at the lowest perturbative orders, the evolution equations of the coupling constants are irrotational flows, generated by the gradients of a scalar function. In fact, they can be written as

$$\dot{g}_k \;=\; \frac{\partial}{\partial g_k}\tilde{C}(g), \tag{15.4.26}$$

where the function $\tilde{C}(g)$ is defined as

$$\tilde{C}(g) \;=\; \sum_i (1-\Delta_i)\, g_i^2 - \frac{\pi}{3}\sum_{ijl} \mathbf{C}_{ijl} g_i\, g_j\, g_l. \tag{15.4.27}$$

This observation will be useful in the discussion in Section 15.6.

15.5 C-theorem

For the deformations of the unitary conformal models there is an important theorem associated to the renormalization group flows. The theorem, due to A.B. Zamolodchikov, states the following: for a two-dimensional field theory that is unitary, invariant under rotations, and for which conservation of the stress-energy tensor holds, there exists a function of the coupling constants $C(\{\lambda_i\})$ that decreases along the flows, being stationary only at the fixed points. Its value at the fixed points coincides with the central charge c of the corresponding conformal field theories. There is a simple proof of this theorem. Let T, Θ, and $\overline{T}$ be the components of spin 2, 0, and -2, respectively, of the stress–energy tensor. For their dimensions and spins, the off-critical correlators of these quantities can be parameterized as

$$\begin{aligned}
\langle T(z,\overline{z})T(0,0)\rangle &= \frac{F\,(m\; z\overline{z})}{z^4},\\
\langle T(z,\overline{z})\Theta(0,0)\rangle &= \frac{G\,(m\; z\overline{z})}{z^3\overline{z}},\\
\langle \Theta(z,\overline{z})\Theta(0,0)\rangle &= \frac{H\,(m\; z\overline{z})}{z^2\overline{z}^2},
\end{aligned} \tag{15.5.1}$$

where m is a mass scale. Using the conservation law of the stress–energy tensor in complex coordinates (given in eqn (10.5.6))

$$\begin{aligned}
\partial_{\overline{z}} T + \tfrac{1}{4}\,\partial_z \Theta = 0,\\
\partial_z \overline{T} + \tfrac{1}{4}\,\partial_{\overline{z}} \Theta = 0,
\end{aligned} \tag{15.5.2}$$

we obtain the differential equations for the scalar functions F, G and H

$$\begin{aligned}
\dot{F} + \tfrac{1}{4}\,\Big(\dot{G} - 3G\Big) = 0;\\
\dot{G} - G + \tfrac{1}{4}\,\Big(\dot{H} - 2H\Big) = 0,
\end{aligned} \tag{15.5.3}$$

where $\tau = m^2 z\bar{z} = (mR)^2$ and

$$\dot{F} \;\equiv\; \frac{dF(x)}{d\log\tau}.$$

Defining

$$C \equiv 2F - G - \frac{3}{8}H, \tag{15.5.4}$$

we have

$$\dot{C} = -\frac{3}{4}H. \tag{15.5.5}$$

The hypothesis of unitarity implies that H is a positive quantity and therefore C is a non-decreasing function of the distance τ, at fixed values of the coupling constants. At the critical points, the trace vanishes, $\Theta = 0$, and consequently $G = H = 0$ with $F = \frac{1}{2}c$. Hence the function C assumes the value of the central charge of the corresponding conformal field theory.

The same theorem can be reformulated in terms of the coupling constants. In fact, fixing the parameter τ (for instance, $\tau = 1$), the quantities F, G, and H become functions of the coupling constants g. From the dimensionless nature of the function $C(R, g)$ and its independence of the cut-off, it satisfies the renormalization group equation

$$\left(\frac{1}{2}R\frac{\partial}{\partial R} - \sum_a \beta^a \frac{\partial}{\partial g_a}\right) C(R, g) = 0. \tag{15.5.6}$$

Using now eqn (15.4.11), we get

$$\beta^a \frac{\partial}{\partial g^a} C(g) = -\frac{3}{4}\, G_{ab}(g)\, \beta^a(g)\, \beta^b(g), \tag{15.5.7}$$

where

$$G_{ab}(g) = G_{ab}(1, g), \qquad G_{ab}(z\bar{z}, g) = (m\, z\, \bar{z})^2\, \langle \varphi_a(z, \bar{z}) \varphi(0, 0) \rangle$$

is a symmetric matrix that is positive definite by the unitarity of the theory. As a by-product of this result, we see that $G_{ab}(g)$ may be regarded as a metric tensor in the space G of the coupling constants, with line element $ds^2 = G_{ab}(g) dg^a\, dg^b$.

The c-theorem admits also an integral formulation that ends up in a sum rule. Integrating eqn (15.5.5) from the ultraviolet fixed point at $r = 0$ and the infrared fixed point at $r = \infty$, and denoting by $\Delta c = c_1 - c_2$ the difference of central charges of the two conformal theories emerging in these limits, we have the equivalent expressions

$$\begin{aligned} \Delta c &= \frac{3}{4} \int_0^\infty d(r^2)\, r^2\, \langle \Theta(r) \Theta(0) \rangle \\ &= \frac{3}{4\pi} \int d^2 r\, r^2\, \langle \Theta(r) \Theta(0) \rangle \\ &= \frac{3}{2} \int dr\, r^3\, \langle \Theta(r) \Theta(0) \rangle. \end{aligned} \tag{15.5.8}$$

This formula remarkably links the second moment of the off-critical correlation function of Θ to the variation of the central charges along the renormalization group flow (an infrared massive theory corresponds to $c_2 = 0$).

The expression above can also be written in terms of the correlation function of the perturbing field. Consider, in fact, a conformal field theory perturbed, for simplicity, by only one relevant scalar field φ, with conformal weight Δ

$$\mathcal{S} \,=\, \mathcal{S}^* + \lambda \int d^2x\, \varphi(x). \tag{15.5.9}$$

Let's examine the renormalization of the analytic component $T(z)$ of the stress–energy tensor. To the first order we have

$$\langle T(z) \cdots \rangle_\lambda \,=\, \langle T(z) \cdots \rangle_0 - \lambda \int d^2z_1\, \langle T(z) \varphi(z_1, \bar{z}_1) \cdots \rangle + \cdots$$

From the operator expansion

$$\begin{aligned} T(z)\varphi(z_1, \bar{z}_1) \,&=\, \frac{\Delta}{(z - z_1)^2}\, \varphi(z_1, \bar{z}_1) + \frac{1}{z - z_1}\, \partial_z \varphi(z_1, \bar{z}_1) + \cdots \\ &=\, \frac{\Delta}{(z - z_1)^2}\, \varphi(z, \bar{z}) + \frac{1 - \Delta}{z - z_1}\, \partial_z \varphi(z, \bar{z}) + \cdots \end{aligned}$$

one can see that the integral is ultraviolet divergent and needs to be regularized. This can be done by inserting in the integral the step function[4] $H((z - z_1)(\bar{z} - \bar{z}_1) - a^2))$, where a is the ultraviolet cut-off. The most singular term vanishes after the angular integration but, for the presence of the function H and the cut-off a, the quantity $\partial_{\bar{z}} T$ is no longer zero. In fact,

$$\begin{aligned} \partial_{\bar{z}} T \,&=\, -\lambda \int \frac{1 - \Delta}{(z - z_1)} (z - z_1)\, \partial_z \varphi(z, \bar{z})\, \delta(|z - z_1|^2 - a^2)\, d^2 z_1 \\ &=\, -\pi\, \lambda\, (1 - \Delta)\, \partial_z \varphi. \end{aligned} \tag{15.5.10}$$

Since the stress–energy tensor satisfies the conservation law

$$\partial_z\, T \,+\, \frac{1}{4}\, \partial_z\, \Theta \,=\, 0,$$

comparing with the equation above we get

$$\Theta \,=\, 4\pi\, \lambda (1 - \Delta)\, \varphi(z, \bar{z}) + \cdots \tag{15.5.11}$$

Note that this expression can be directly recovered from eqn (15.4.12), for Θ is the conjugate field to the scale transformation $x \to tx$: in the action (15.5.9), $d^2x \to t^2\, d^2x$ while $\varphi \to t^{2\Delta}\, \varphi$, therefore taking the derivative with respect to t, we obtain eqn (15.5.11), with the additional factor 2π that takes into account the conformal normalization of the operators. Let's now see some interesting applications of the c-theorem.

[4] $H(x) = 0$ if $x < 0$, while $H(x) = 1$ if $x > 0$.

15.6 Applications of the c-theorem

In this section we initially study the renormalizion group flow associated to the deformation $\Phi_{1,3}$ of the minimal conformal models. Later, we present two applications of the sum rule of the c-theorem, relative to the Ising model and the lagrangian theory of the Sine–Gordon model.

15.6.1 Minimal Conformal Models $\mathcal{M}_p$ perturbed by the $\Phi_{1,3}$ Operator

The first significant application of the c-theorem is in the study of the unitary minimal models $\mathcal{M}_p$ perturbed by the relevant operator $\Phi_{1,3}$

$$\mathcal{S} \;=\; \mathcal{S}_p + \lambda \int d^2x\, \Phi_{1,3}(x). \tag{15.6.1}$$

$\Phi_{1,3}$ is an operator characterized by two specific properties. The first property is related to the operator product expansion with itself, that has the skeleton form

$$\Phi_{1,3} \times \Phi_{1,3} \;=\; \mathbf{1} + \mathbf{C}_1\, \Phi_{1,3} + \mathbf{C}_2\, \Phi_{1,5}. \tag{15.6.2}$$

Since $\Phi_{1,5}$ is an irrelevant operator, the operator expansion above implies the renormalization of the field $\Phi_{1,3}$, which does not mix with any other fields. The second property is related to its conformal weight

$$\Delta_{1,3} \;=\; 1 - \frac{2}{p+1} \;\equiv\; 1 - \epsilon. \tag{15.6.3}$$

For p sufficiently large, $\Phi_{1,3}$ is a quasi-marginal operator and therefore we are in the condition of the validity of eqn (15.4.25). The structure constant $\mathbf{C}_1$ tends to a finite limit for $p \to \infty$: using its exact expression (obtained by substituting in eqn (11.5.53) the indices (n, m) with $(1, 3)$), one has

$$\begin{aligned}\mathbf{C}_1(\epsilon) = \frac{4}{\sqrt{3}} \frac{(1-2\epsilon)^2}{(1-\epsilon)(1-3\epsilon/2)} \left[\frac{\Gamma(1-\epsilon/2)\,\Gamma(1+3\epsilon/2)}{\Gamma(1+\epsilon/2)\,\Gamma(1-3\epsilon/2)}\right]^{1/2} \\ \times \frac{\Gamma(1-2\epsilon)\,\Gamma^2(1+\epsilon)}{\Gamma(1+2\epsilon)\,\Gamma^2(1-\epsilon)} \;=\; \frac{4}{\sqrt{3}} \left(1 - \frac{3\epsilon}{2} + \cdots\right).\end{aligned} \tag{15.6.4}$$

The β-function of the dimensionless coupling constant g associated to λ is then

$$\dot{g} \;=\; \beta(g) \;=\; 2\epsilon\, g - \pi\, \mathbf{C}_1\, g^2 + \cdots \tag{15.6.5}$$

The plot of this function, given in Fig. 15.3, shows the existence of a new fixed point. As shown in Section 15.4, the β-function can be written as the gradient of a scalar function

$$\dot{g} \;=\; \frac{\partial}{\partial g} \tilde{C}(g),$$

where $\tilde{C}(g)$, in this case, is given by

$$\tilde{C}(g) \;=\; \epsilon\, g^2 - \frac{\pi}{3} \mathbf{C}_1\, g^3. \tag{15.6.6}$$

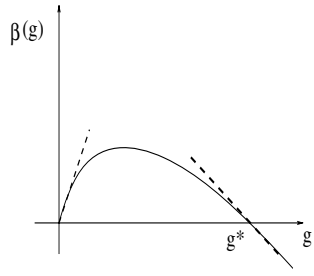


Fig. 15.3 *β-function relative to $\Phi_{1,3}$ deformation of the unitary minimal models. The slope of the tangents at the fixed points is directly related to the anomalous dimension of the operator.*

This expression allows us to easily compute the function $C(g)$ entering the c-theorem. In fact, both functions have the same stationary points and, at this perturbative order, they must be proportional to each other, so

$$C(g) \;=\; c + \alpha\, \tilde{C}(g) + \mathcal{O}(g^4).$$

The proportionality constant can be fixed by a perturbative computation of C, using eqns (15.5.5) and (15.5.11):

$$\begin{aligned} C &= c - \frac{3}{4}(4\pi)^2\,(1-\Delta)^2\, g^2 \int_0^1 \frac{r^4}{r^{4\Delta}}\,\frac{d(r^2)}{r^2} + \cdots \\ &= c - 6\pi^2\,\epsilon g^2 + \cdots \end{aligned} \tag{15.6.7}$$

i.e. $\alpha = -6\pi^2$.

The fixed point of the β-function is at $g^* = 2\epsilon/\pi\mathbf{C}_1$, and this value is compatible with the perturbation expansion. Substituing it in $C(g)$ we get an estimate of the central charge of the new fixed point

$$C(g^*) \;=\; c - \frac{8\epsilon^3}{\mathbf{C}_1^2} + \mathcal{O}(\epsilon^4). \tag{15.6.8}$$

Substituting in this expression $\epsilon \simeq 2/p$ and the value of the structure constant given above, we have

$$C(g^*) \;=\; c(p) - \frac{12}{p^3}. \tag{15.6.9}$$

At this perturbative order, the new value of the central charge coincides with that of the unitary minimal model $\mathcal{M}_{p-1}$. Hence, the deformation $\Phi_{1,3}$ associated to a positive sign of the coupling constant gives rise, for large values of p, to a massless RG flow between two nearest conformal field theories

$$\mathcal{M}_p \to \mathcal{M}_{p-1} \tag{15.6.10}$$

This scenario is compatible with the Landau–Ginzburg formulation of these unitary minimal models. In fact, in the Landau–Ginzburg formulation, the field $\Phi_{1,3}$ of the

model $\mathcal{M}_p$ corresponds to the operator $:\varphi^{2(p-1)-2}:$. Once we switch on this deformation, the highest power $:\varphi^{2(p-1)}:$ that defines the original conformal theory $\mathcal{M}_p$ becomes irrelevant. Consequently, the infrared dynamics of the perturbed system is described by a Landau–Ginzburg model with interaction $:\varphi^{2(p-2)}:$, i.e. the one that corresponds to the conformal theory $\mathcal{M}_{p-1}$.

In this RG flow there is also an evolution of the other fields, so that they occupy different positions in the Kac table of the starting and ending conformal field theory. Hence, also the anomalous dimensions of the operator change accordingly, and their variation can be computed using eqn (15.4.18), evaluated at $g = g^*$. Let's compute, for instance, the anomalous dimension at the new fixed point g^* of the original field $\Phi_{1,3}$: the derivative of β at this fixed point is

$$\frac{\partial \beta}{\partial g}(g^*) \,=\, -2\,\epsilon \tag{15.6.11}$$

and therefore the conformal weight of the field at this new fixed point is $\Delta' = 1 + 2/(p+1)$. At this perturbative order, it coincides with the conformal weight of the irrelevant field $\Phi_{3,1}$ of the conformal model $\mathcal{M}_{p-1}$. Hence, in the RG flow, the operator $\Phi^{(p)}_{1,3}$ of $\mathcal{M}_p$ transforms in to the operators $\Phi^{(p-1)}_{3,1}$ of $\mathcal{M}_{p-1}$.

In addition to the field $\Phi_{1,3}$, one can also study the evolution of the other fields of the Kac table. Consider, for instance, those along the main diagonal of the Kac table of the starting conformal theory, $\Phi^{(p)}_{n,n}$. To follow their evolution, some preliminary data are needed. Their operator product expansion with $\Phi^{(p)}_{1,3}$ reads

$$\Phi^{(p)}_{n,n}\,\Phi^{(p)}_{1,3} \,=\, \mathbf{C}^{(n,n)}_{(n,n),(1,3)}\,\left[\Phi^{(p)}_{n,n}\right] + \mathbf{C}^{(n,n+2)}_{(n,n),(1,3)}\,\left[\Phi^{(p)}_{n,n+2}\right] + \mathbf{C}^{(n,n-2)}_{(n,n),(1,3)}\,\left[\Phi^{(p)}_{n,n-2}\right].$$

For large p, the structure constants tend to

$$\begin{aligned}
\mathbf{C}^{(n,n)}_{(n,n),(1,3)} &= \frac{(n-p)^2(p+1)}{2\sqrt{3}(p-1)} + \mathcal{O}(\epsilon) \quad (n \le p) \\
\mathbf{C}^{(n,n)}_{(n,n),(1,3)} &= \frac{(n^2-1)}{8\sqrt{3}}\,\epsilon^2 + \mathcal{O}(\epsilon^3) \quad (n \ll p) \\
\mathbf{C}^{(n,n+2)}_{(n,n),(1,3)} &= \left(\frac{p+2}{3p}\right)^{1/2} + \mathcal{O}(\epsilon) \\
\mathbf{C}^{(n,n-2)}_{(n,n),(1,3)} &= \frac{(n^2-1)^{1/2}}{\sqrt{3}n} + \mathcal{O}(\epsilon).
\end{aligned}$$

The conformal weights of the operators $\Phi^{(p)}_{n,n}$ are

$$\Delta^{(p)}_{(n,n)} \,=\, \frac{n^2-1}{4p(p+1)} \,=\, \frac{n^2-1}{16}\,\epsilon^2\,\left(1 + \frac{\epsilon}{2} + \cdots\right), \tag{15.6.12}$$

and, for $n \ll p$, they are strongly relevant. Their operator expansion with $\Phi_{1,3}$ shows that they do not mix with any other operators. To determine in which fields they transform along the RG flow, one must compute their anomalous dimensions at $g = g^*$

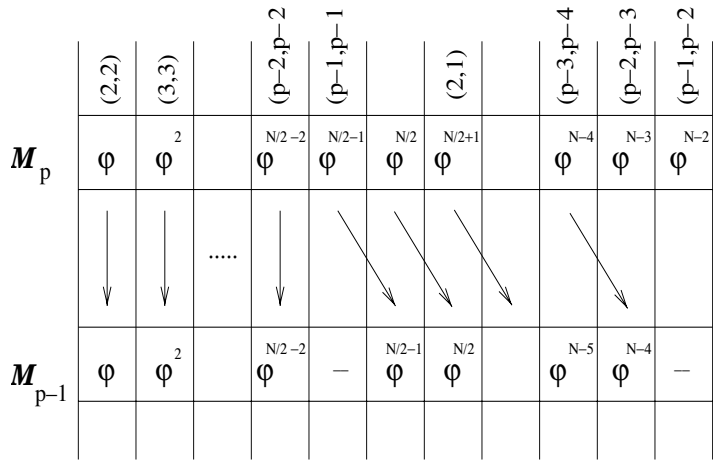


Fig. 15.4 *Mappings of the conformal fields in the massless RG flow* $\mathcal{M}_p \to \mathcal{M}_{p-1}$ *induced by* $\Phi^{(p)}_{1,3}$. *Here* $N = 2(p-1)$.

in terms of the derivative of $\beta^{(n)}$-functions of these fields with respect the coupling constant g. Using the general formula (15.4.25), one obtains

$$\frac{\partial \beta^{(n)}}{\partial g} \;=\; -2\pi\, \mathbf{C}^{(n,n)}_{(n,n),(1,3)}\, g + \cdots \tag{15.6.13}$$

so that, at the new fixed points, the anomalous dimensions are

$$\begin{aligned} \hat{\gamma}(g^*)\, \Phi_{n,n} \;&=\; \gamma_{(n,n)}(g^*)\, \Phi_{n,n} \;=\; \left[2\Delta_{(n,n)} + 2\pi \mathbf{C}^{(n,n)}_{(n,n),(1,3)}\, g^* \right]\, \Phi_{n,n} \\ &=\; 2\,\epsilon^2\, \frac{n^2-1}{16} \left(1 + \frac{3\epsilon}{2} + \cdots \right) \;=\; \frac{n^2-1}{4p(p-1)} + \mathcal{O}(\epsilon^4). \end{aligned}$$

This implies the mappings

$$\Phi^{(p)}_{n,n} \to \Phi^{(p-1)}_{n,n}, \tag{15.6.14}$$

The analysis can be extended to all other fields and the final result is summarized in Fig. 15.4.

15.6.2 Ising Model at Temperature $T \neq T_c$

Consider the two-dimensional Ising model in its fermionic formulation at $T \neq T_c$. At the critical point the action is

$$\mathcal{S}^* \;=\; \int d^2x \left[\psi\, \partial_{\bar{z}}\, \psi + \bar{\psi}\, \partial_z\, \bar{\psi} \right], \tag{15.6.15}$$

and the perturbation that moves the system away from the critical point is given by the mass term $i\,m \int \bar{\psi}\,\psi\, d^2x$. To compute the correlator of $\Theta(x) = i\,m\,\bar{\psi}\,\psi$ we need

the propagator of the massive fermionic field

$$\begin{aligned}
\langle\, \bar\psi(z,\bar z)\psi(0,0)\,\rangle &= -i\,m \int \frac{d^2p}{(2\pi)^2}\,\frac{e^{\frac{i}{2}(p\bar z+\bar p z)}}{p^2+m^2} = -i\,\frac{m}{2\pi}\,K_0(mr)\\
\langle\, \psi(z,\bar z)\psi(0,0)\,\rangle &= -i \int \frac{d^2p}{(2\pi)^2}\,\frac{\bar p\, e^{\frac{i}{2}(p\bar z+\bar p z)}}{p^2+m^2} = 2\,\partial_z\,\frac{1}{2\pi}\,K_0(mr)\\
&= -\frac{m}{2\pi}\,\frac{\bar z}{z}\,K_1(mr) \qquad (15.6.16)\\
\langle\, \bar\psi(z,\bar z)\bar\psi(0,0)\,\rangle &= -\frac{m}{2\pi}\,\frac{z}{\bar z}\,K_1(mr)
\end{aligned}$$

with $r = \sqrt{z\bar z}$, where $K_i(x)$ are the modified Bessel functions. Applying Wick's theorem, we have

$$\begin{aligned}
\langle\Theta(r)\Theta(0,0)\rangle &= -m^2\,\langle\bar\psi(r)\psi(r)\bar\psi(0)\psi(0)\rangle\\
&= -\,|\langle\bar\psi(r)\psi(0)\rangle|^2 + \langle\psi(r)\psi(0)\rangle\,\langle\bar\psi(r)\bar\psi(0)\rangle \qquad (15.6.17)\\
&= \left(\frac{m^2}{2\pi}\right)^2 \left[K_1^2(mr) - K_0^2(mr)\right].
\end{aligned}$$

Substituting this expression into the sum rule (15.5.8) and computing the integral, we get

$$\Delta c = \frac{1}{2}. \qquad (15.6.18)$$

Since $c = \frac{1}{2}$ is the central charge of the critical Ising model, we have an explicit check that the perturbed theory has central charge $c = 0$, i.e. a purely massive field theory.

15.6.3 A Lagrangian Theory: The Sine–Gordon Model

Another interesting example of the c-theorem sum rule comes from a lagrangian theory with a varying coupling constant. Consider the Sine–Gordon model, with lagrangian

$$\mathcal{L} = :\frac{1}{2}(\partial_\mu\varphi)^2 + \frac{m^2}{\beta^2}(\cos\beta\varphi - 1): \qquad (15.6.19)$$

Let's restrict attention to the range $\beta^2 < 8\pi$, where $\cos\beta\varphi$ is a relevant operator. This massive model (with $c = 0$) can be regarded as a deformation of the free massless bosonic theory with $c = 1$. Adopting this interpretation, we have

$$\lambda\varphi \equiv \epsilon(x) = \frac{m^2}{\beta^2} :(\cos\beta\varphi - 1):$$

with anomalous dimension

$$2\Delta = \frac{\beta^2}{4\pi}.$$

Equation (15.5.8) becomes

$$\Delta c = 3\pi \left(2 - \frac{\beta^2}{4\pi}\right)^2 \int d^2x \mid x \mid^2 \langle\epsilon(x)\epsilon(0)\rangle. \qquad (15.6.20)$$

Since the left-hand side of (15.6.20) does not depend on β (in particular, it is identically equal to 1), the same must hold for the right-hand side. Note that the sum rule is

already saturated at zero order in β by the term that corresponds to the free massive theory. At this order we have in fact

$$\epsilon(r) \;=\; \frac{m^2}{2}\varphi^2(r), \qquad \Delta = 0;$$
$$\langle \epsilon(r)\epsilon(0)\rangle \;=\; \frac{m^4}{2}\langle \varphi(r)\varphi(0)\rangle^2 \;=\; \frac{m^4}{8\pi^2}\, K_0^2(mr),$$

and therefore

$$\Delta c_0 \;=\; 3\pi \frac{m^4}{2\pi^2}\int d^2x \mid x \mid^2 K_0^2(m \mid x \mid) \;=\; 3\int_0^\infty dR\, R^3\, K_0^2(R) \;=\; 1. \qquad (15.6.21)$$

This implies that, expanding in power series with respect to β^2, on the right-hand side of (15.6.20) all coefficients but the constant must vanish. Let's check the validity of this conclusion on the first non-trivial term: taking the first derivative with respect to β^2 and imposing $\beta^2 = 0$ we have

$$\frac{1}{3\pi}\frac{d(\Delta c)}{d\beta^2}\bigg|_{\beta^2=0} \;=\; -\frac{1}{\pi}\int d^2x \mid x \mid^2 \langle \epsilon(x)\epsilon(0)\rangle + 4\int d^2x \mid x \mid^2 \frac{d}{d\beta^2}\langle \epsilon(x)\epsilon(0)\rangle.$$

The first term has already been computed in the free massive theory. To compute the second term, let's expand $\cos(\beta\varphi)$ up to fourth order and then use Wick's theorem, with the result

$$\frac{1}{3\pi}\frac{d(\Delta c)}{d\beta^2}\mid_{\beta^2=0} = \; -\frac{1}{12\pi^2} + \frac{m^6}{(2\pi)^4}\int d^2x\, d^2z \mid x \mid^2 K_0^2(m \mid x - z \mid)\, K_0^2(m \mid z \mid). \qquad (15.6.22)$$

The last integral can be easily computed: changing the variable

$$\mid x - z \mid \to \mid t \mid,$$

it is expressed by the product of the integrals

$$2\pi \textstyle\int_0^\infty R^3\, K_0^2(R)\, dR \;=\; \frac{2\pi}{3};$$
$$2\pi \textstyle\int_0^\infty R\, K_0^2(R)\, dR \;=\; \pi$$

(the term (15.6.22) with the scalar product $\vec{x}\cdot\vec{z}$ vanishes after the angular intergration). Inserting this expression in (15.6.22), one can see that the variation of the central charge by varying β^2 is effectively zero, as it should be. The perturbative check can be easily generalized to the next order and it is natural to conjecture its validity to all perturbative orders. Hence, using the c-theorem we can generate in this case an infinite number of identities that involve the integrals of the correlation functions of the Sine–Gordon model.

15.7 Δ-theorem

The c-theorem provides useful information on the RG flow induced by the relevant fields of a conformal field theory. There is another theorem that permits us to follow

directly the change of the anomalous dimensions of the various fields. Its formulation is due to Delfino, Simonetti and Cardy.

Consider the off-critical correlators of the components $T(z,\bar{z})$ and $\Theta(z,\bar{z})$ of the stress–energy tensor with a field Φ, where the latter is the deformed primary field of the perturbed conformal theory. These correlators can be parameterized as

$$\begin{aligned}\langle T(z,\bar{z})\Phi(0,0)\rangle &= \frac{U(mz\,\bar{z})}{z^2},\\ \langle \Theta(z,\bar{z})\Phi(0,0)\rangle &= \frac{V(mz\,\bar{z})}{z\bar{z}}.\end{aligned}$$

Using the conservation of the stress–energy tensor, eqn (15.5.2), we arrive at the differential equation

$$\dot{D} = \frac{1}{4}V, \tag{15.7.1}$$

where $D = U + \frac{1}{4}V$ and the dot denotes the logarithmic derivative $z\bar{z}\,\frac{d}{dz\bar{z}}$. Since the trace Θ is related to the perturbing field by the relation (15.5.11), the short-distance expansion of the function V is determined by the operator product expansion with the perturbing field Φ, i.e.

$$V(x) \simeq 2\pi\,\lambda\,(2-2\Delta)\,C^0_{\varphi\Phi}\,|x|^{2(\Delta_0-\Delta_\Phi-\Delta+1}\,\langle A_0\rangle, \tag{15.7.2}$$

where A_0 is the most relevant operator that appears in this expansion. It is necessary to distinguish two cases:

1. If
$$\Delta_0 - \Delta_\Phi - \Delta + 1 > 0, \tag{15.7.3}$$
then $V(x)$ vanishes in the conformal limit $x \to 0$. In this case the function D is stationary at the fixed point and coincides with U. If the operator Φ does not mix with other fields under renormalization, we can straightforward by use its operator expansion with T and for the function U we have
$$U(x) \simeq \Delta_\Phi\,\langle\Phi\rangle. \tag{15.7.4}$$
If the perturbed theory is associated to massless flow to another conformal field theory, the same analysis can be repeated near the other fixed point and, integrating over all distance scales, we establish the sum rule
$$\Delta^{uv}_\Phi - \Delta^{ir}_\Phi = -\frac{1}{4\pi\langle\Phi\rangle}\int d^2x\,\langle\Theta(x)\Phi(0)\rangle. \tag{15.7.5}$$
If the RG flow leads to a massive theory, we have instead $\Delta^{ir}_\Phi = 0$.
2. If we have instead
$$\Delta_0 - \Delta_\Phi - \Delta + 1 < 0, \tag{15.7.6}$$
the function $V(x)$ does not vanish at the origin and the attempt to use eqn (15.7.5) fails for the divergence of the integral. This is a simple consequence of the mixing of the operator Φ under renormalization. In this case, the

function $U(x)$ does not present the behavior of eqn (15.7.4) at short distances. Its correct behavior is obtained by directly integrating eqn (15.7.1), namely

$$U(x) \simeq \pi\lambda(1-\Delta)\, C^0_{\varphi\Phi}\, \frac{1-\gamma_0}{\gamma_0}\, |x|^{2\gamma_0} \quad \gamma_0 < 0$$
$$U(x) \simeq 2\pi\lambda(1-\Delta)\, C^0_{\varphi\Phi}\, \langle A_0 \rangle\, \log|x| \quad \gamma_0 = 0$$

where we introduce the notation $\gamma_0 = \Delta_0 - \Delta_\Phi - \Delta + 1$.

It is worth stressing that the Δ-theorem, as expressed by eqn (15.7.5), can be easily generalized to quantum field theories of any dimensions when the integral converges both in the ultraviolet and the infrared regions. Indeed it simply expresses the Ward identity relative to the trace of the stress–energy tensor of any field theory, i.e. the field responsible for the global scale transformations.

References and Further Reading

The splitting of the ultraviolet and infrared properties of a quantum field theory has a long history. It has been applied, for instance, to study the vacuum of quantum chromodynamics. For this topic see:

M. A. Shifman, A.I. Vainshtein, V. I Zakharov, *QCD and resonance physics. Sum rules*, Nucl. Phys. B 147 (1979), 385.

In two dimensions the perturbative formalism was developed by Al.B. Zamolodchikov in the paper:

Al.B. Zamolodchikov, *Two point correlation function in scaling Lee–Yang model*, Nucl. Phys. B 348 (1991), 619.

For a discussion of the renormalization group in two-dimensional systems see:

A.B. Zamolodchikov, *Renormalization group and perturbation theory about fixed points in two-dimensional field theory*, Sov. J. Nucl. Phys. 45 (1987), 1090.

A.W.W. Ludwig, J.L. Cardy, *Perturbative evaluation of the conformal anomaly at new critical points with applications to random systems*, Nucl. Phys. B 285 (1987), 687.

The computation of the β-functions with the formalism of the operator product expansion is discussed in:

J.L. Cardy, *Conformal invariance and statistical mechanics*, in *Fields, Strings and Critical Phenomena* (Les Houches 1988), ed. E. Brezin and J. Zinn-Justin.

The c-theorem has been proved by A.B. Zamolodchikov in the paper:

A.B. Zamolodchikov, *Irreversibility of the flux of the renormalization group in a 2D field theory*, JETP Lett. 43 (1986), 730.

Its integral version is due to J.L. Cardy and can be found in:

J. Cardy, *Central charge and universal combinations of amplitudes in two-dimensional theories away from criticality*, Phys. Rev. Lett. 60 (1998), 2709.

The Δ-theorem has been proved by G. Delfino, P. Simonetti and J.L. Cardy in the article:

G. Delfino, P. Simonetti, J.L. Cardy, *Asymptotic factorization of form-factors in two-dimensional quantum field theory*, Phys. Lett. B 387 (1996), 327.

16
Integrable Quantum Field Theories

This belief is handed down in Beersheba: that, suspended in the heavens, there exists another Beersheba, where the city's most elevated virtues and sentiments are poised, and that if the terrestrial Beersheba will take the celestial one as its model the two cities will become one.

Italo Calvino, *Invisible Cities*

16.1 Introduction

An integrable quantum field theory is characterized by an infinite number of conserved charges. In classical mechanics, the existence of a sufficient number of integrals of motion allows us to pass from the initial coordinates and momenta to the angle-action variables, thus finding the exact solution of the equation of motion by quadrature. Similarly, if in a quantum field theory there are an infinite number of conservation laws, we can derive the exact mass spectrum of its excitations, the S-matrix of the scattering processes, the correlation functions, the thermodynamics, and so on, in short its exact solution. For reasons that will become clearer later, non-trivial integrable quantum field theories can only occur in $(1+1)$ dimensions.[1] In higher dimensions, in fact, they are either free theories or models with non-local interactions. Hence we focus our attention only on two-dimensional models.

In $(1+1)$ dimensions, using the complex notation to denote the analytic and anti-analytic indices of tensor quantities, the conservation law of a current with components (T_{s+1}, Θ_{s-1}) is written as

$$\partial_{\bar{z}} T_{s+1} \;=\; \partial_z \Theta_{s-1}, \tag{16.1.1}$$

and this leads to the conservation of the charges

$$\mathcal{Q}_s \;=\; \oint \left[T_{s+1}\, dz + \Theta_{s-1} d\bar{z} \right]. \tag{16.1.2}$$

The integer index s that identifies the integrals of motion is called the *spin* of the operator $\mathcal{Q}_s$. The value $s = 1$ always corresponds to the stress–energy tensor, with

[1] A remark on the notation: in the following we will use the terminology "$(1+1)$ dimensions" if we want to stress the Minkowski version of a two-dimensional quantum field theory, while we will use the terminology "two dimensions" to denote both a generic two-dimensional quantum field theory or its euclidean version.

$\mathcal{Q}_1 = \mathcal{P}$, where $\mathcal{P} = E + P$ is the analytic part of the total momentum of the system. The set of values of s is specific to each integrable model, as we will see in the examples presented in this chapter.

In the previous chapters we stress that an important aspect of the two-dimensional conformal theories is the splitting of the analytic and anti-analytic sectors. If we take, in this case, as T_{s+1} any independent descendent field at the level $(s+1)$ of the identity conformal family, it is easy to see that the conservation laws (16.1.1) are identically satisfied for, in conformal theories, T_{s+1} is a purely analytic field satisfying $\partial_{\bar{z}} T_{s+1} = 0$. Hence, all two-dimensional conformal theories have an infinite number of conservation laws and therefore can also be considered as integrable models.

Perturbing the conformal theories by means of the insertion of one or more relevant fields, there is a breaking of the factorization of the analytic/anti-analytic sectors. Consequently, there is in general the destruction of all the hierarchy of the conserved currents of the conformal point. As we have seen in the previous chapter, the analysis of the models away from criticality can always be carried out by perturbative techniques, but this approach rarely leads to an exact solution of the model. It is therefore a circumstance of the utmost importance that some particular deformations of the critical action lead to the definition of integrable models also away from criticality. This possibility opens in fact more interesting scenarios than the perturbative approach, since it allows us to solve exactly the statistical models also away from criticality.

In the first part of this chapter we discuss the integrable quantum field theories that are associated to a lagrangian density. In the second part, we examine the conditions on the deformations of a conformal theory that lead to the existence of conserved currents and an integrable theory away from the critical point. The physical consequences of this remarkable circumstance will be the objects of our study in the following chapters.

16.2 The Sinh–Gordon Model

Consider the two-dimensional euclidean space and the lagrangian theory of the so-called Sinh-Gordon model. Its action is given by[2]

$$\mathcal{S} \;=\; \int d^2x \left[\frac{1}{16\pi} \left(\partial_\mu \phi\right)^2 + \frac{\mu^2}{g^2} \cosh\left(\frac{g}{\sqrt{8\pi}} \phi(x) \right) \right]. \tag{16.2.1}$$

With this normalization, in the limit $\mu \to 0$ the two-point correlation function is

$$\langle \phi(x) \phi(y) \rangle \;=\; -2 \log(x^2 + y^2). \tag{16.2.2}$$

The Sinh–Gordon model provides the simplest example of a general class of integrable models, the so-called Toda field theories discussed in more detail in Section 16.6. The action (16.2.1) is invariant under the Z_2 transformation $\phi \to -\phi$. The potential of the lagrangian

$$V(\phi) \;=\; \frac{\mu^2}{g^2} \cosh\left(\frac{g}{\sqrt{8\pi}} \phi \right),$$

[2]The numerical factors of this definition are chosen in such a way as to match the notation of the general Toda field theories, discussed below.

has a unique minimum at the origin, with a quadratic curvature equal to μ^2. Hence, in Minkowski space, this theory describes the interactions of a relativistic particle of (bare) mass μ. In euclidean space, though, there are several equivalent ways to consider this field theory according to the different splittings of the action

$$\mathcal{S} \;=\; \mathcal{S}_0 + \mathcal{S}_I. \tag{16.2.3}$$

In this expression $\mathcal{S}_0$ plays the role of the unperturbed action, with $\mathcal{S}_I$ of its deformation. This splitting leads to different expressions of the central charge of the conformal theory that emerges in the ultraviolet regime of the Sinh–Gordon model. Which central charge will be selected is in fact related to the choice of $\mathcal{S}_0$, while the central charge in the infrared regime is always $C_{ir} = 0$ since, independently of the various splittings, the theory is always massive in its infrared regime.

One should not be surprised that the action (16.2.1) can describe different ultraviolet fixed points. As a matter of fact, the value of the central charge is not linked to the lagrangian but is instead related to the definition that we assume for the associated stress–energy $T^{\mu\nu}(\mathrm{x})$. This is an operator intrinsically defined up to a total divergence: if we denote by $\tilde{T}^{\mu\nu}$ the stress–energy tensor coming from Noether's theorem, there is in fact a one-parameter family (labeled by the parameter α) of stress–energy tensors associated to the *same* lagrangian given by

$$\begin{aligned} T_{\mu\nu}(x) \;&=\; \tilde{T}_{\mu\nu}(x) \;+\; \alpha\,(\partial_\mu\partial_\nu \;-\; g_{\mu\nu}\Box)\,\phi(x), \\ \tilde{T}_{\mu\nu}(x) \;&=\; \left[\partial_\mu\phi\,\partial_\nu\phi \;-\; \eta_{\mu\nu}\left(\frac{1}{2}(\partial\phi)^2 \;-\; V(\phi)\right)\right]. \end{aligned} \tag{16.2.4}$$

We recall that, in the modified Coulomb gas formalism, the presence of α is equivalent to introducing a charge at infinity and, as shown in Chapter 11, this changes the conformal properties of the fields and leads to different values of the central charge. Let's see the four different ways of interpreting the theory (16.2.1), each of them corresponding to a particular choice of the parameter α.

Feynman perturbation approach. The first approach is based on standard perturbation theory defined by the Feynman graphs. In this case, first we expand in power series in g the hyperbolic term present in the lagrangian, identifying as $\mathcal{S}_0$ the expression

$$\mathcal{S}_0 \;=\; \frac{1}{16\pi}\int d^2x\,\left[\,(\partial_\mu\phi)^2 + \mu^2\phi^2)\right], \tag{16.2.5}$$

and as $\mathcal{S}_I$ all the other terms of the series expansion. In this approach, the Sinh–Gordon model appears as a Landau–Ginzburg theory with an infinite number of interaction terms, all even in the field ϕ. Since in two dimensions the field ϕ is dimensionless, the theory is renormalizable. The only ultraviolet divergences come from the *tadpole* graphs, such as those shown in Fig. 16.1. Calling Λ the ultraviolet cut-off, these divergences can be removed at once by redefining the exponentials in terms of the normal

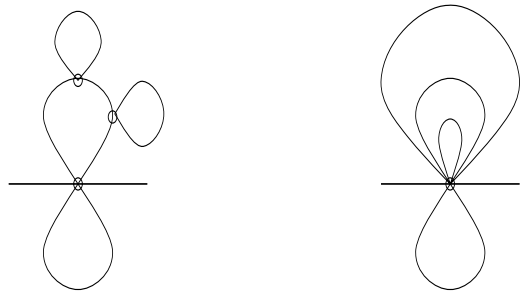


Fig. 16.1 *Tadpole diagrams entering the two-point correlation function.*

order product[3] with respect to an arbitrary mass scale μ

$$e^{\pm \frac{g}{\sqrt{8\pi}}\phi(x)} \rightarrow: e^{\pm \frac{g}{\sqrt{8\pi}}\phi(x)} := \left(\frac{\Lambda}{\mu}\right)^{g^2/8\pi} e^{\pm \frac{g}{\sqrt{8\pi}}\phi(x)}. \tag{16.2.6}$$

The removal of the divergences is equivalent to the renormalization of the mass term μ^2. This choice of $\mathcal{S}_0$ corresponds to $\alpha = 0$ in eqn (16.2.4). The corresponding ultraviolet central charge is then $C_{uv} = 1$, as one can easily derive by applying the c-theorem discussed in the previous chapter.

Deformation of gaussian action. An alternative way to consider the Sinh–Gordon model consists of taking as $\mathcal{S}_0$ only the kinetic term and as $\mathcal{S}_I$ the Z_2 invariant combination of the vertex operators

$$S_O = \frac{1}{8\pi}\int d^2x(\partial_\mu\phi)^2, \qquad S_I = \frac{\mu^2}{g^2}\int d^2x \left(e^{\frac{g}{\sqrt{8\pi}}\phi} + e^{-\frac{g}{\sqrt{8\pi}}\phi}\right). \tag{16.2.7}$$

In this case, S_0 is explicitly associated to the conformal theory of a free gaussian bosonic field (discussed in Chapter 12), whose ultraviolet central charge is $C_{uv} = 1$. Also in this case $\alpha = 0$, but with $\tilde{T}_{\mu\nu}$ only expressed by the kinetic part $\partial_\mu\phi\partial_\nu\phi$, while all the remaining terms are considered as part of its trace.

Deformation of Liouville action I. A third way to look at the Sinh–Gordon model is to take as $\mathcal{S}_0$ the Liouville action

$$\mathcal{S}_0 = \int d^2x \left[\frac{1}{16\pi}(\partial_\mu\phi)^2 + \lambda\, e^{\frac{g}{\sqrt{8\pi}}\phi}\right], \tag{16.2.8}$$

and as $\mathcal{S}_I$ the deformation given by the relevant operator $e^{-\frac{g}{\sqrt{8\pi}}\phi}$. Although the Liouville action is formally invariant under conformal transformations, its correct quantization requires the introduction of a charge at infinity

$$Q_+ = \frac{1}{2}\left(\frac{g}{\sqrt{8\pi}} + \frac{\sqrt{8\pi}}{g}\right). \tag{16.2.9}$$

[3]This definition of the normal order prohibits a propagator starting and ending at the same point. It is similar to the definition adopted in Section 11.5.1 for the vertex operators, even though in this case the conformal weights of the exponential operators are negative, $\Delta = -\frac{g^2}{8\pi}$.

This quantization method is done along the lines of the Coulomb gas approach discussed in Section 11.5.1. The charge at infinity shifts the value of the ultraviolet central charge, which is no longer $C_{uv} = 1$ but

$$C_{uv} \;=\; 1 + 24\, Q_+^2. \tag{16.2.10}$$

There is also a shift of the conformal weights of the vertex operators $e^{\alpha\phi(x)}$, now given by

$$\Delta_+^{(\alpha)} \;=\; -\alpha^2 + 2\alpha\, Q_+. \tag{16.2.11}$$

The reason to quantize the Liouville theory with the charge at infinity should now be clear: in fact, it is only in this way that the vertex operator $e^{\frac{g}{\sqrt{8\pi}}\phi(x)}$, present in the action (16.2.8), acquires a conformal weight equal to 1, so that the action (16.2.8) becomes conformally invariant at the quantum level. For the perturbing operator $e^{-\frac{g}{\sqrt{8\pi}}\phi(x)}$, its new conformal weight is instead

$$\Delta_- \;=\; -1 - \frac{g^2}{4\pi}. \tag{16.2.12}$$

Deformation of Liouville action II. In this Liouville approach, we can obviously exchange the role played by the two exponentials, namely we can take as $\mathcal{S}_0$ the Liouville theory defined by the other exponential

$$\mathcal{S}_0 \;=\; \int d^2x \left[\frac{1}{16\pi} (\partial_\mu \phi)^2 + \lambda e^{-\frac{g}{\sqrt{8\pi}}\,\phi} \right], \tag{16.2.13}$$

and as perturbation $\mathcal{S}_I$ the one induced by the vertex operator $e^{\frac{g}{\sqrt{8\pi}}\phi(x)}$. As in the previous case, this Liouville action needs a charge at infinity for its correct quantization, this time given by

$$Q_- \;=\; -\frac{1}{2} \left(\frac{g}{\sqrt{8\pi}} + \frac{\sqrt{8\pi}}{g} \right). \tag{16.2.14}$$

This charge at infinity modifies the values of the central charge and the conformal weights of the vertex operators $e^{\alpha\phi(x)}$, which can be obtained by substituting $Q_+ \to Q_-$ in the previous formulas (16.2.10) and (16.2.11). In this scheme the operator $e^{-\frac{g}{\sqrt{8\pi}}\phi(x)}$ has conformal weight equal to 1, while that of $e^{\frac{g}{\sqrt{8\pi}}\phi(x)}$ is

$$\Delta_+ \;=\; -1 - \frac{g^2}{4\pi}. \tag{16.2.15}$$

Integrability. It is worth stressing that, independently of how we interpret the theory in the ultraviolet regime, the Sinh–Gordon model enjoys a fundamental property: it is an integrable model. To obtain the classical expression of the conserved charges $\mathcal{Q}_s$,

it is convenient to rescale for simplicity the field as $\phi(x) \to \frac{g\phi}{\sqrt{8\pi}}$ and impose $\mu = 1$. Let's also introduce the light-cone coordinates σ and τ:

$$\sigma \;=\; \frac{1}{2}(x-t); \quad \tau \;=\; \frac{1}{2}(x+t).$$

In these coordinates the equation of motion becomes

$$\partial_\sigma\, \partial_\tau\, \phi(\sigma,\tau) \;=\; \sinh(\phi). \tag{16.2.16}$$

There is a conserved charge $\mathcal{Q}_s$ if there exists a current with components (J^0_s, J^1_s) satisfying the equation $\partial_\mu J^\mu_s = 0$. This can be written in light-cone coordinates defining $J^0_s = T_{s+1} + \Theta_{s-1}$ and $J^1_s = T_{s+1} - \Theta_{s-1}$. For the densities $T_{s+1}[\phi]$ and $\Theta_{s-1}[\phi]$ we have

$$\frac{\partial}{\partial\sigma} T_{s+1}[\phi] \;=\; \frac{\partial}{\partial\tau}\Theta_{s-1}[\phi]. \tag{16.2.17}$$

The index s refers to the spin of this current, related to the difference of the partial derivatives ∂^n_τ and ∂^k_σ present in the expression of the densites, $s = n - k - 1$. The charge Q_s

$$\mathcal{Q}_s \;=\; \int_{-\infty}^{\infty} dx\, J^0_s \;=\; \int \left[T_{s+1}\, d\tau + \Theta_{s-1}\, d\sigma\right] \tag{16.2.18}$$

is a conserved quantity since, by eqn (16.2.17), it satisfies

$$\frac{d\mathcal{Q}_s}{dt} \;=\; 0. \tag{16.2.19}$$

To explicitly find the densities $T_{s+1}[\phi]$ and $\Theta_{s-1}[\phi]$, let's define the field $\hat\phi(\sigma,\tau)$, the solution of the so-called *Bäcklund transformations*

$$\begin{aligned} \partial_\sigma(\hat\phi - \phi) \;&=\; 2\,\epsilon\, \sinh\left(\frac{1}{2}(\hat\phi + \phi)\right), \\ \partial_\tau(\hat\phi + \phi) \;&=\; \frac{2}{\epsilon}\, \sinh\left(\frac{1}{2}(\hat\phi - \phi)\right). \end{aligned} \tag{16.2.20}$$

Assuming that $\phi(\sigma,\tau)$ is a solution of the equation of motion, eqns (16.2.20) provide another solution. In fact, acting with ∂_τ of the first of them and using the second equation, we have

$$\partial_\tau\, \partial_\sigma(\hat\phi - \phi) \;=\; 2\, \sinh\frac{1}{2}(\hat\phi - \phi)\, \cosh\frac{1}{2}(\hat\phi + \phi) \;=\; \Big[\sinh(\hat\phi) - \sinh(\phi)\Big].$$

The field $\hat\phi(z,\bar z,\epsilon)$ can be expressed in power series of the parameter ϵ

$$\hat\phi(\sigma,\tau,\epsilon) \;=\; \sum_{n=0}^{\infty} \phi^{(n)}(\sigma,\tau)\, \epsilon^n \tag{16.2.21}$$

where $\phi^{(n)}(\sigma,\tau)$ can be computed by plugging it into (16.2.20) and comparing term by term in ϵ. For the first terms we have[4]

$$\begin{aligned} \hat{\phi}^{(0)} &= \phi & , \quad \phi^{(1)} &= 2\phi_\tau, \\ \hat{\phi}^{(2)} &= 2\phi_{\tau\tau} & , \quad \phi^{(3)} &= 2\phi_{\tau\tau\tau} - \phi_\tau^3/3, \\ \hat{\phi}^{(4)} &= 2\phi_{\tau\tau\tau\tau} - 2\phi_\tau^2\,\phi_{\tau\tau} , & \cdots \end{aligned} \tag{16.2.22}$$

The existence of this series expression gives us the possibility of obtaining an infinite number of conservation laws starting from a finite number of them. To this end we can use, for instance

$$\left(\frac{1}{2}\psi_\sigma^2\right)_\tau + (1-\cosh\psi)_\sigma = 0, \tag{16.2.23}$$

or a similar equation

$$\left(\frac{1}{2}\psi_\tau^2\right)_\sigma + (1-\cosh\psi)_\tau = 0, \tag{16.2.24}$$

whose validity can be easily checked employing the equation of motion (16.2.16) satisfied by ϕ. Using, for instance, eqn (16.2.24) and substituting eqn (16.2.22), we obtain an infinite number of conserved densities. The first non-trivial expressions (i.e. those that cannot be expressed as total derivative) are

$$\begin{aligned} T_2 &= \frac{1}{2}\phi_\tau^2 \\ T_4 &= 2\phi_{\tau\tau}^2 + 2\phi_\tau\,\phi_{\tau\tau\tau} \\ T_6 &= 2\phi_{\tau\tau\tau}^2 + 4\phi_{\tau\tau}\,\phi_{\tau\tau\tau\tau} - 6\phi_{\tau\tau}^2\phi_\tau^2 - 2\phi_\tau^3\phi_{\tau\tau\tau} + 2\phi_\tau\phi_{\tau\tau\tau\tau\tau}. \end{aligned} \tag{16.2.25}$$

In general, it can be proved that non-trivial conservation laws are obtained for all odd values of s

$$s = 1, 3, 5, \ldots \tag{16.2.26}$$

The set of these values of s constitutes the spectrum of the conserved charges. It is also possible to show that the classical expressions of the conserved currents, suitably modified, keep their meaning also at the quantum level and that the corresponding charges are in involution, i.e. they commute with each other

$$[\,\mathcal{Q}_s, \mathcal{Q}_{s'}\,] = 0. \tag{16.2.27}$$

In the next chapter we will see how the involution nature of the conserved charges and the spectrum of s strongly influence the structure of the massive excitations and their dynamics.

[4] $\phi_\tau \equiv \partial_\tau\phi$ e $\phi_\sigma \equiv \partial_\sigma\phi$.

16.3 The Sine–Gordon Model

Under the analytic continuation $g \to ig$, the Sinh–Gordon model becomes the Sine–Gordon model, with the euclidean action given by

$$\mathcal{S} = \int d^2x \left[\frac{1}{16\pi} (\partial_\mu \phi)^2 - \frac{\mu^2}{g^2} \cos(\frac{g}{\sqrt{8\pi}} \phi(x)) \right]. \tag{16.3.1}$$

As seen in Chapter 12, this theory enters the bosonization procedure. However, there is a wider range of phenomena where the Sine–Gordon model plays a crucial role: it describes, for instance, the dislocation of a crystal subjected to an external force; in quantum optics, it describes the propagation of a light wave in a material made of two quantum levels and, in superconductivity, it plays an important role in the theory of magnetic flux of a Josephson junction.

For the Sine–Gordon model we can also adopt the four different approaches discussed for the Sinh–Gordon model to study its ultraviolet limit. Because of their similarity, we will not repeat their analysis here. However, it is important to note the following circumstance: adopting the quantization scheme of the Liouville theory, since now the exponentials are complex, we get an ultraviolet central charge less than 1

$$C_{uv} = c = 1 - 24Q^2, \quad Q = \pm \left(\frac{g}{\sqrt{8\pi}} - \frac{g}{\sqrt{8\pi}} \right) \tag{16.3.2}$$

(the sign of Q depends on the particular choice of the exponential that enters the Liouville action). This implies that, with an appropriate choice of the coupling constant g, the central charge c can take the values of the minimal models $\mathcal{M}_{p,q}$. When this happens, it is easy to see that the perturbing operator of the Liouville action corresponds to the operator $\Phi_{1,3}$ of the minimal models.

The Sine-Gordon model has an infinite sequence of conserved charges in involution that ensures both its classical and quantum integrability. It is useful to provide an explicit sequence of the first quantum densities T_{s+1}, in the quantization scheme of the Liouville theory, adopting the complex coordinates $z = t + ix$ and $\bar{z} = t - ix$. If $T(z)$ is the analytic component of the stress–energy tensor associated to the Liouville theory

$$T(z) = -\frac{1}{2}(\partial_z \phi)^2 + iQ\, \partial_z^2 \phi, \tag{16.3.3}$$

and $(AB)(z)$ denotes the normal order of two operators as defined by their operator product expansion

$$(AB)(z) = \oint_z \frac{dw}{w - z} A(w)B(z),$$

we have

$$\begin{aligned} T_2 &= T, \\ T_4 &= (T^2) \\ T_6 &= (T(T^2)) + \frac{c+2}{12}(T\, \partial_z^2 T), \\ T_8 &= (T(T(T^2))) + \frac{c+8}{6}(T(T\partial_z^2 T)) + \frac{1}{180}(c^2 + 4c - 101)(T\partial_z^4 T). \\ \cdots &= \cdots \end{aligned} \tag{16.3.4}$$

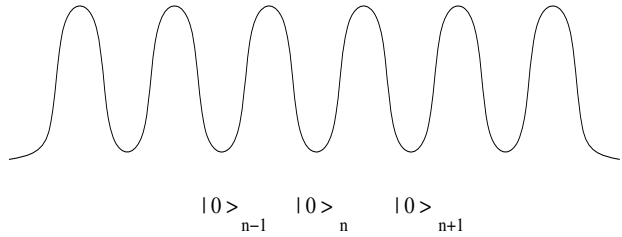


Fig. 16.2 *Potential of the Sine–Gordon model and sequence of the infinite equivalent vacua.*

The dynamics of the model is better understood if we consider its formulation in Minkowski space. To simplify the notation, let's rescale the field as $\phi \to \phi/\sqrt{8\pi}$, so that the action in Minkowski space becomes

$$\mathcal{S} \,=\, \int d^2x \,\left[\frac{1}{2}(\partial_\mu \phi)^2 + \frac{\mu^2}{g^2}\cos g\phi(x)\right].$$

The potential of the theory

$$V(\phi) \,=\, \frac{\mu^2}{g^2}\,[1-\cos(\,g\phi)] \tag{16.3.5}$$

(to which we have added, for convenience, a constant) presents an infinite series of degenerate minima placed at $\phi = 2\pi n/g$ $(n = 0, \pm 1, \ldots)$, as shown in Fig. 16.2. In the quantum version, they correspond to an infinite family of equivalent vacua, denoted by $|\,0\,\rangle_n$. Around each minimum, the potential has a quadratic concavity μ^2 that can be associated to the mass of the scalar particle created out of the vacuum by the field ϕ. This scalar particle does not however, exhaust, the spectrum of the excitations of the the model. In the Sine–Gordon model there are in fact topological excitations of finite energy, associated to those field configurations that interpolate between two degenerate vacua.

Topological excitations. The topological excitations can be specified by two integers (n_1, n_2) that label the vacua $2\pi n_1/g$ and $2\pi n_2/g$ reached by the field $\phi(x)$ at $x \to \pm\infty$. Let's define then the topological charge

$$Q_t \,=\, n_1 - n_2 \,=\, \frac{1}{2\pi g}\int_{-\infty}^{\infty} dx\,\frac{\partial \phi}{\partial x}. \tag{16.3.6}$$

From the periodicity of $V(\phi)$, the field ϕ is defined modulo $2\pi/g$, i.e. at any given point x the value of the field can be changed as $\phi(x) \to \phi(x) + 2\pi k$, maintaining, though, the continuity of its configurations. The topological charge is insensitive to these transformations as long as we keep fixed the final values assumed by the fields.

Let's write down the energy of a generic configuration of $\phi(x,t)$

$$E[\phi] \,=\, \int_{-\infty}^{\infty} dx\,\left[\frac{1}{2}\left(\frac{\partial \phi}{\partial t^2}\right)^2 + \left(\frac{\partial \phi}{\partial x^2}\right)^2 + V(\phi)\right], \tag{16.3.7}$$

and the equation of motion of the model

$$\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} = \frac{\partial V}{\partial \phi}. \tag{16.3.8}$$

The classical expression of the elementary topological configurations, i.e. those associated to $Q_t = \pm 1$, can be obtained by looking at the static solutions of the equation of motion. In this case the first term on the right-hand side vanishes and the equation of motion reduces to

$$\frac{\partial^2 \phi}{\partial x^2} = -\frac{\partial V}{\partial \phi}.$$

This expression coincides, formally, with the equation of motion of classical mechanics of a fictitious particle described by the coordinate $\phi(x)$ and subjected to the potential $-V(\phi)$ (note the change of sign in the potential). In this interpretation, the original variable x in the field $\phi(x)$ plays the role of time coordinate of the classical particle. As with any classical system subjected to a conservative force, it has an integral of motion given by its mechanical energy (which must not to be confused with $E[\phi]$), given by

$$W = \frac{1}{2}\left(\frac{d\phi}{dx}\right)^2 - V(\phi). \tag{16.3.9}$$

The value of the constant of motion W can be immediately determined. In fact, if we require that the static solutions $\phi(x)$ have a finite energy $E[\phi]$, we must have at $x \to \pm\infty$ both $V(\phi) \to 0$ and $(\partial\phi/\partial x) \to 0$. In analogy with the newtonian motion of the particle, this means that the particle at $x \pm \infty$ has to be in one of the maxima of the potential $-V(\phi)$ and, furthermore, that its velocity has to vanish both at the starting and ending points. For the constant W we have $W = 0$. Instead of solving the second-order equation of motion (16.3.8), for the static solutions we can take advantage of the mechanical analogy and find the solution by quadrature from eqn (16.3.9):

$$\frac{d\phi}{dx} = \pm\sqrt{2\,V(\phi)} \qquad \to \qquad (x - x_0) = \pm \int_{\phi(x_0)}^{\phi(x)} \frac{d\bar{\phi}}{\sqrt{2V(\bar{\phi})}}, \tag{16.3.10}$$

where x_0 is an arbitrary constant of integration. Performing the integral, with the explicit expression of $V(\phi)$ given in (16.3.5), we get

$$\bar{\phi}(x) = \pm 4/g \arctan\left[\exp m(x - x_0)\right]. \tag{16.3.11}$$

The first solution, the one with the positive sign, has topological charge $Q_t = 1$ and corresponds to a *soliton* that interpolates between the vacuum $\bar{\phi} = 0$ and the next one at $2\pi/g$ or, equivalently, between a generic pair of vacua $2\pi n/g$ and $2\pi(n+1)/g$. The second solution, the one with the negative sign, has instead $Q_t = -1$ and corresponds to an *antisoliton* that interpolates between a generic pair of vacua $2\pi n/g$ e $2\pi(n-1)/g$,

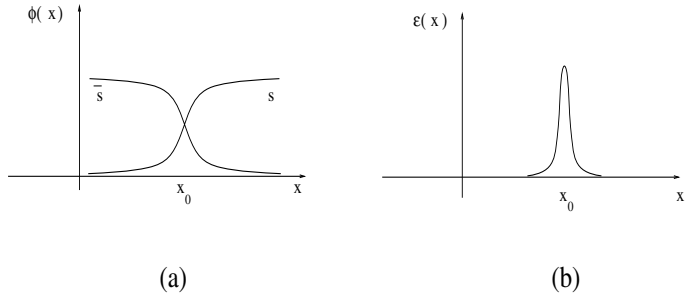


Fig. 16.3 *Solitonic solutions and their energy density* $\epsilon(x)$.

as shown in Fig. 16.3. The origin of the terminology is in the peculiar form assumed by the energy density $\epsilon(x)$ of these solutions entering the formula

$$E[\bar{\phi}] \;=\; \int_{-\infty}^{\infty} dx\, \epsilon(x), \qquad \epsilon(x) \;=\; \frac{4\mu^2}{g^2}\, \frac{1}{\cosh^2 m(x-x_0)}. \tag{16.3.12}$$

As shown in Fig. 16.3, $\epsilon(x)$ has a shape strongly localized at x_0 that rapidly decreases to zero outside an interval large of order $1/m$. For this localization property, the solitonic solutions of the Sine–Gordon model can be interpreted as particle excitations of the system and the energy (16.3.12) of the static solution corresponds to the mass M_s of the soliton/antisoliton

$$M_s \;=\; \frac{8\mu^2}{g^2}. \tag{16.3.13}$$

The non-perturbative nature of the solitonic solutions is revealed by the dependence on the coupling constant, placed in the denominator of the expression above. The particle nature of these excitations is further confirmed by using the Lorentz invariance of the equation of motion: given a static solution $\bar{\phi}(x)$, we can use a Lorentz transformation[5] to transform it into a solution that moves with velocity v

$$\bar{\phi}(x) \to \bar{\phi}\left[\frac{m(x-x_0)-vt}{\sqrt{1-v^2}}\right].$$

It is easy to check that this expression indeed satisfies the equation of motion (16.3.8) and substituting it in (16.3.7), we get

$$E[\bar{\phi}(x,t)] \;=\; \frac{M_s}{\sqrt{1-v^2}}.$$

Hence we recover the Einstein relationship that links the mass and the energy of a particle. The solitons are then particle excitations of the system that, in the classical description, appear as waves that propagate in the medium without dispersion or dissipation, always keeping their shape intact.

[5]The velocity is measured in units of the light velocity, so that its limiting value is $v = 1$.

Time-dependent solutions. The Sine–Gordon model admits exact solutions also in other topological sectors, although they are time-dependent expressions. For instance, a solution with $Q_t = 0$ is

$$\bar{\phi}_{s\bar{s}}(x,t) \;=\; \frac{4}{g}\arctan\left(\frac{\sinh(mvt/\sqrt{1-v^2})}{v\,\cosh(mx/\sqrt{1-v^2})}\right). \tag{16.3.14}$$

It has the peculiar property of tending, for $t \to \pm\infty$ to a configuration made of a soliton and an antisoliton

$$\bar{\phi}_{s\bar{s}}(x,t) \;\to\; \bar{\phi}_s\left(\frac{x+v(t\pm\Delta_{s\bar{s}}/2)}{\sqrt{1-v^2}}\right)+\bar{\phi}_{\bar{s}}\left(\frac{x-v(t\pm\Delta_{s\bar{s}}/2)}{\sqrt{1-v^2}}\right), \qquad t\to\pm\infty.$$

When the time varies, this solution describes an elastic scattering process, whose only effect is a negative time shift $\Delta_{s\bar{s}} \equiv (1-v^2)v\,\log v$ of the propagation of the soliton and the antisoliton with respect to their free propagation. The elasticity of the scattering processes is a common characteristic in all other topological sectors. For instance, in the sector with topological charge $Q_t = 2$, a solution of the equation of motion is given by

$$\bar{\phi}_{ss}(x,t) \;=\; \frac{4}{g}\arctan\left(\frac{v\,\sinh(x/\sqrt{1-v^2})}{\cosh(vt/\sqrt{1-v^2})}\right). \tag{16.3.15}$$

At any given time, it interpolates between the vacua $-2\pi/g$ and $2\pi/g$. It can then be interpreted as a configuration made of two solitons. Following the time evolution of this solution, one realizes that it corresponds to the elastic scattering of the two solitons, since at $t \to \pm\infty$ it becomes

$$\bar{\phi}_{ss}(x,t) \;\to\; \bar{\phi}_s\left(\frac{x+v(t\pm\Delta_{ss}/2)}{\sqrt{1-v^2}}\right)+\bar{\phi}_{s}\left(\frac{x-v(t\pm\Delta_{ss}/2)}{\sqrt{1-v^2}}\right), \qquad t\to\pm\infty.$$

Also in this case, the only effect of the interaction is a time shift Δ_{ss}, although positive this time (see Problem 4).

In conclusion, the Sine–Gordon theory is an integrable theory that has multisolitonic solutions that describe purely elastic scattering processes. The elasticity of these scattering processes is a consequence of the infinite number of conserved charges of the model. To compute the complete spectrum of the excitations of its quantum version it is necessary to study the S-matrix of the scattering processes, a subject that will be addressed in Chapter 18.

16.4 The Bullogh–Dodd Model

The Bullogh-Dodd model is another lagrangian system that is integrable both at the classical and at the quantum level. Its euclidean theory is defined by

$$\mathcal{S} \;=\; \int d^2x\,\left\{\frac{1}{16\pi}\,(\partial_\mu\phi)^2+\frac{\mu^2}{6g^2}\left[2\,e^{\frac{g}{\sqrt{8\pi}}\phi}+e^{-2\frac{g}{\sqrt{8\pi}}\phi}\right]\right\}. \tag{16.4.1}$$

It may be considered as a deformation of the Liouville theory

$$\mathcal{S}_0 \;=\; \int d^2x \left\{ \frac{1}{16\pi} \left(\partial_\mu \phi\right)^2 + \frac{\mu^2}{3g^2}\, e^{\frac{g}{\sqrt{8\pi}}\phi} \right\}, \tag{16.4.2}$$

by means of the exponential $e^{-2\frac{g}{\sqrt{8\pi}}\phi}$. As in the Sinh–Gordon model, the quantization of this theory requires the introduction of a charge at infinity, in this case expressed by

$$Q_+ \;=\; \frac{1}{2}\left(\frac{g}{\sqrt{8\pi}} + \frac{\sqrt{8\pi}}{g} \right).$$

This leads to an ultraviolet central charge equal to

$$C_{uv} \;=\; 1 + 24 Q_+^2. \tag{16.4.3}$$

The conformal weight $\Delta(\alpha)$ of the exponentials $e^{\alpha\phi}$ is given by

$$\Delta(\alpha) \;=\; -\alpha^2 + 2\alpha\, Q_+. \tag{16.4.4}$$

In this way, the exponential in (16.4.2) has a conformal weight equal to 1, while the other exponential has a conformal weight

$$\Delta[e^{-2\frac{g}{\sqrt{8\pi}}\phi}] \;=\; -2 - \frac{3}{4\pi} g^2.$$

By the analytic continuation $g \to ig$, the central charge (16.4.3) becomes less than 1 and, by an opportune choice of the coupling constant g, we can match the central charges of the minimal models. In this case, it is easy to show that the operator $e^{-2\sqrt{2}g\phi}$ corresponds to the operator $\Phi_{1,2}$ of the conformal minimal models. However, contrary to what happens in the Sinh–Gordon model, the substitution $g \to ig$ makes, this time, the action (16.4.1) a complex quantity and therefore it is not obvious how it can give rise, as indeed it does, to a consistent physical theory.

Vice versa, as a starting Liouville theory we can assume

$$\mathcal{S}_0 \;=\; \int d^2x \left\{ \frac{1}{16\pi} \left(\partial_\mu \phi\right)^2 + \frac{\mu^2}{6g^2}\, e^{-2\frac{g}{\sqrt{8\pi}}\,\phi} \right\}. \tag{16.4.5}$$

In this case the charge at infinity is

$$Q_- \;=\; -\frac{1}{2}\left(\frac{g}{\sqrt{2\pi}} + \frac{\sqrt{2\pi}}{g} \right).$$

The central charge and the conformal weights have the same expressions as above, once we make the substitution $Q_+ \to Q_-$. In this way the exponential present in the action has a conformal weight equal to 1, while

$$\Delta[e^{\frac{g}{\sqrt{8\pi}}\,\phi}] \;=\; -\frac{1}{2} - \frac{3g^2}{8\pi}.$$

With the analytic continuation $g \to ig$ and with the choice of g that matches the central charge with that of conformal minimal models, the perturbing operator $e^{\sqrt{2}g\phi}$can be identified with the field $\Phi_{2,1}$ of the minimal models.

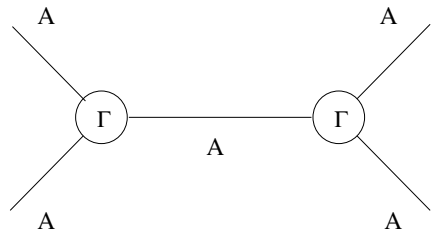


Fig. 16.4 *Scattering amplitudes of the Bullogh–Dodd theory.*

To discuss the perturbative quantization of the theory, based on Minkowski space and the Feynman graphs, it is convenient to scale the field and the coupling constant in such a way that the action reads

$$\mathcal{L} \,=\, \frac{1}{2}(\partial_\mu \phi)^2 - \frac{\mu^2}{6g^2}\left(2e^{g\phi} + e^{-2g\phi}\right), \tag{16.4.6}$$

where μ is a mass parameter and g the coupling constant. Also this model belongs to the Toda field theories,[6] discussed in Section 16.6. The series expansion of the exponential terms gives rise to the n-leg interaction vertices

$$V(\phi) \,=\, \frac{\mu^2}{6g^2}\left(2e^{g\phi} + e^{-2g\phi}\right) \,=\, \frac{\mu^2}{2\,g^2} + \sum_{k=2}^{\infty} \frac{g_k}{k!}\phi^k,$$

where

$$g_k \,=\, \frac{\mu^2}{3} g^{k-2}\left[1 + (-1)^k 2^{k-1}\right].$$

Also in this case the renormalization of the divergences met in the perturbative series reduces to eliminate the tadpoles. This is equivalent to renormalizing the mass term μ as

$$\mu^2 \,\to \mu^2 \left(\frac{\Lambda}{\mu}\right)^{g^2/4\pi}.$$

This theory is supported by an infinite number of conserved charges, whose spectrum of spin s is given by

$$s \,=\, 1, 5, 7, 11, 13, \ldots \tag{16.4.7}$$

i.e. all odd integer numbers apart from multiples of 3.

The perturbative particle content of the theory consists of a particle, denoted by A, that takes part in scattering processes in which it appears as a bound state of itself. This is a simple consequence of the ϕ^3 vertex in $V(\phi)$, which gives rise to scattering processes such as those shown in Fig. 16.4. Here we anticipate that this perturbative scenario will be confirmed by the exact S-matrix of this model discussed in Chapter 18.

[6]This model can be obtained by the folding with respect to a Z_2 symmetry of the affine simply laced algebra A_2. The resulting algebra is denoted by BC_1 and its Coxeter number is $h = 3$.

16.5 Integrability versus Non-integrability

One may wonder what is so special about the Sinh–Gordon, Sine–Gordon, or Bullogh–Dodd models with respect to other two-dimensional field theories. The answer to this question can be given both at the classical and the quantum level. Let's first discuss the classical aspects.

The Sine–Gordon model is not the only theory to possess static topological configurations. If we examine further the argument used to find the solitonic solutions, we realize that it is sufficient that the theory simply possesses two degenerate next neighbor vacua. From this point of view, even the ϕ^4 theory, in the phase in which the Z_2 is spontaneously broken, should have solitonic excitations. The potential $\mathcal{U}(\phi)$ of this theory

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \mathcal{U}(\phi), \quad \mathcal{U} = \frac{\lambda}{4}\left(\phi^2 - \frac{\mu^2}{\lambda}\right).$$

has in fact, two degenerate minima at $\phi_\pm = \pm m/\sqrt{\lambda}$. This is indeed the case, and the explicit expression of the solitons of this theory is obtained by inserting $\mathcal{U}(\phi)$ in (16.3.10)

$$\bar{\phi}_(x) = \pm\frac{m}{\sqrt{\lambda}} \tanh\left[\frac{m}{\sqrt{2}}(x - x_0)\right]. \tag{16.5.1}$$

The soliton mass is obtained by substituting their energy density

$$\epsilon(x) = \frac{m^4}{2\lambda}\frac{1}{\cosh^4(m(x-x_0)/\sqrt{2}},$$

in $E[\phi]$ given in (16.3.7)

$$M = \frac{2\sqrt{2}}{3}\frac{m^3}{\lambda}.$$

Hence, even in the ϕ^4 theory there are solitonic phenomena of non-perturbative nature.

If the configurations of the static solitons of the Sine–Gordon and the ϕ^4 theory may appear very similar,[7] their differences show up when we consider the multisolitonic configurations. In fact, in the Sine–Gordon model these configurations have the properties of preserving the shape that they have at $t \to -\infty$ also at $t \to +\infty$. This is not the case for the ϕ^4 theory. In other words, the scattering processes that take place in the ϕ^4 theory are inelastic: the initial particles, identified as the multisolitons present at $t \to -\infty$, lose their identity during the time evolution.

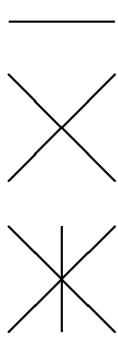
This classical situation has a quantum analog: this permits us to easily appreciate the significant difference that exists between the Sine–Gordon and ϕ^4 theory or, more generally, its difference with respect to any other theory invariant under a Z_2 symmetry. As will be discussed in detail in the next chapter, quantum integrable theories have the same peculiar features already noticed at the classical level, i.e. the elasticity of the scattering processes. This is indeed a peculiar aspect for relativistic quantum field

[7]Note that in the ϕ^4 theory there are no topological sectors with topological charge $|Q_t| > 1$, since there are only two vacua. The multisoliton sequences of this theory consist of only alternate configurations of solitons and antisolitons.

theories since, for purely kinematic reasons, the number of particles is not a conserved quantity: for instance, one can always create $4, 6, 8, \ldots$ particles simply by increasing the energy in the center of mass of two colliding particles. In light of this remark, following an argument by P. Dorey, let's start from the more general two-dimensional Z_2 invariant lagrangian theory

$$\mathcal{L} = \frac{1}{2}\left[(\partial_\mu \phi)^2 - \mu^2 \phi^2\right] - \frac{g_4}{4!}\phi^4 - \frac{g_6}{6!}\phi^6 - \cdots \tag{16.5.2}$$

and let's find out the conditions on the coefficients $g_4, g_6, \ldots$ that prevent the production processes. Let's analyze the simplest case, i.e. a process in which two initial particles gives rise to four final particles. The Feynman rules are



$= i/(p^2 - \mu^2 + i\epsilon)$

$= -ig_4$

$= -ig_6$

etc. Applying these rules, let's compute the tree level processes in which we have $2 \to 4$ particles. Suppose, for simplicity, that the initial particles have just the energy to create the four out-coming particles. Working in the center of mass reference frame, the momenta (p^0, p^1) of the initial particles, satisfying the on-shell relation $(p^0)^2 - (p^1)^2 = \mu^2$, are then $(2\mu, \pm\sqrt{3}\mu)$. The total energy is $E_t = 4\mu$, a value sufficient to create the four final particles. Since these four particles will all be at rest, their common value of the momenta is $(\mu, 0)$. Using the conservation of the total momentum at each vertex, for the graphs of Fig. 16.5 we have

$$\begin{aligned}
(a) &\to \; i\,\frac{g_4^2}{32\mu^2} \\
(b) &\to -i\,\frac{g_4^2}{96\mu^2} \\
(c) &\to -i\,\frac{g_6}{48\mu^2}.
\end{aligned}$$

The total amplitude is

$$(a) + (b) + (c) \;=\; \frac{i}{48\mu^2}\,(g_4^2 - g_6), \tag{16.5.3}$$

so that, choosing $g_6 = g_4^2$, we can *dynamically* suppress this production process.

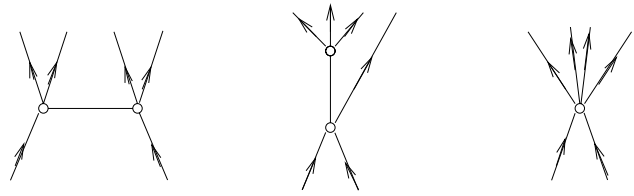


Fig. 16.5 *Feynman graphs at the tree level for the production process* $2 \to 4$.

Generalizing the analysis to the production process $2 \to 6$ and requiring its dynamical absence, one derives the further condition $g_6 = g_4^3$. Carrying out the same analysis for the higher particle production processes, one arrives at the following result: the only lagrangian theories with a Z_2 symmetry that *dynamically* suppress all production processes are represented by the potentials

$$U(\phi) \;=\; \mu^2 \left[\frac{1}{2}\phi^2 \pm \frac{g^2}{4!}\phi^4 + \frac{g^4}{6!}\phi^6 \pm \frac{g^6}{8!}\phi^8 + \cdots \right]. \tag{16.5.4}$$

It is easy to recognize that these potentials either correspond to that of the Sinh–Gordon model (when all the signs are chosen positive) or of the Sine–Gordon model (with the choice of alternate signs).

Repeating the same analysis with the most general Landau–Ginzburg potential, which also presents odd powers of the field ϕ, the one that is selected by the absence of production processes is given by

$$\begin{aligned} U(\phi) \;&=\; \mu^2 \left[\frac{1}{2}\phi^2 - \frac{g}{6}\phi^3 + \frac{g^2}{8}\phi^4 - \frac{g^3}{24}\phi^5 + \cdots \right] \\ &=\; \frac{\mu^2}{6g^2} \left[2e^{g\phi} + 2\, e^{-2g\phi} - 3 \right], \end{aligned} \tag{16.5.5}$$

namely, the Bullogh–Dodd model!

In conclusion, the only relativistic two-dimensional lagrangian theories which involve only one scalar field and that are integrable both at the classical and at the quantum level are the Sinh–Gordon and the Sine–Gordon theories (if there is a Z_2 symmetry) or the Bullogh–Dodd theory.

16.6 The Toda Field Theories

A generalization of the models encountered so far is provided by the Toda field theories. These theories can be constructed using the Lie algebras discussed in the appendix of Chapter 13. In the following we mainly focus our attention on the simply laced algebras A_n, D_n and E_n, i.e. those with simple roots of the same length. The Toda field theories based on the non-simply laced algebras can be defined, at least at the classical level, by an identification of the roots using the symmetry properties of the Dynkin diagram. This is the so-called *folding* procedure, as shown in Table 16.1a and Table 16.1b below.

Table 16.1a: *Foldings of the Dynkin diagrams of the simply laced algebras: the principal series. Near the roots there are the numbers q_i.*

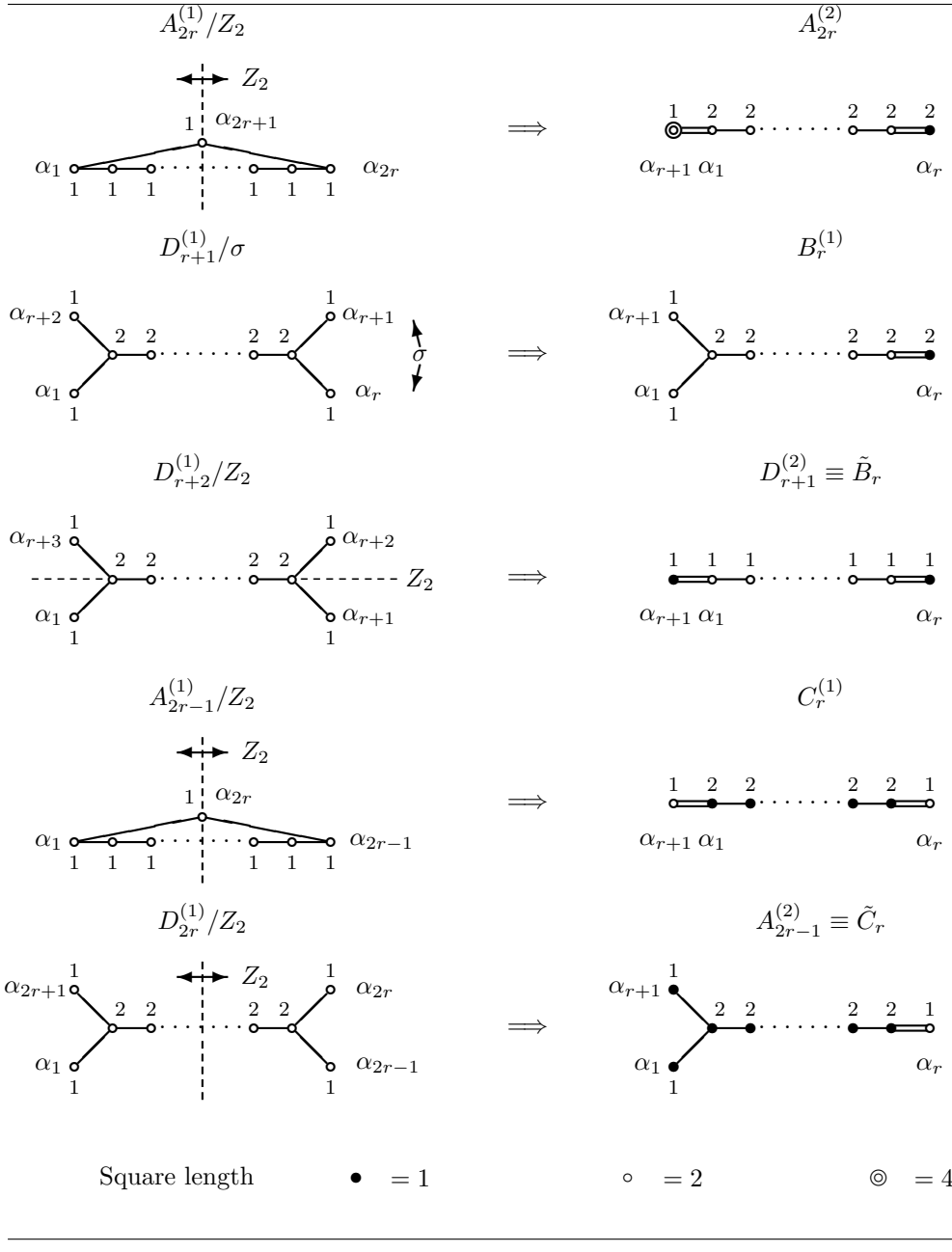
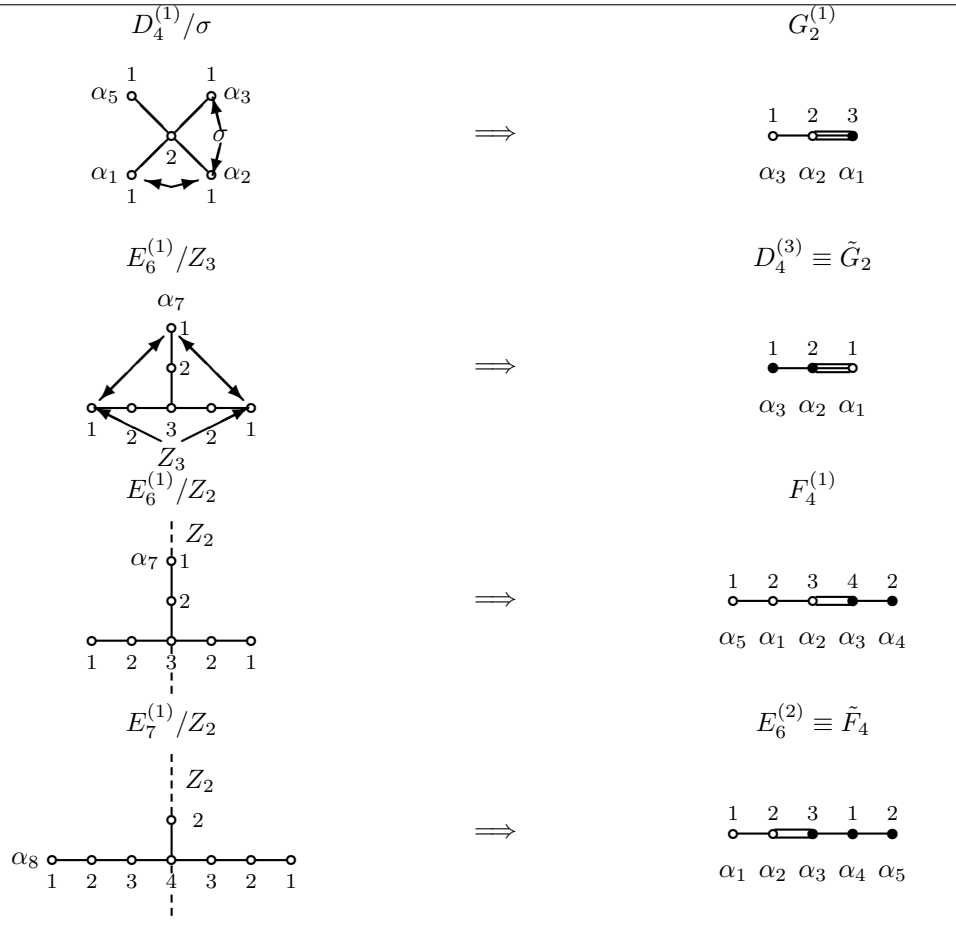


Table 16.1b: *Foldings of the Dynkin diagrams of the simply laced algebras: the exceptional series. Near the roots there are the numbers q_i.*



The Toda field theory associated to a Lie algebra G of rank r is a lagrangian model of r bosonic fields, collected in a vector $\phi = (\phi_1, \ldots, \phi_r)$, given by

$$\mathcal{S} = \int d^2x \left\{ \frac{1}{8\pi} (\partial_\mu \phi) \cdot (\partial^\mu \phi) + \frac{\mu^2}{16\beta^2} \sum_{i=1}^{r+1} q_i [\exp(\beta \alpha_i \cdot \phi) - 1] \right\}, \qquad (16.6.1)$$

where μ^2 and β are real parameters. The set $\{\alpha_i\}_{i=1}^r$ is given the simple roots of G, with their norm equal to 2. The set of the integer numbers $\{q_i\}$ is different for each

algebra and it is related to the definition of the *maximal root* of the algebra, given by[8]

$$\alpha_{r+1} \;=\; -\sum_{i=1}^{r} q_i\,\alpha_i. \tag{16.6.2}$$

The extended set of roots, obtained by adding the maximal root, form the Dynkin diagram of the *affine Lie algebras*. For these systems, imposing $q_{r+1} = 1$, we have

$$\sum_{i=1}^{r+1} q_i \alpha_i = 0, \qquad \sum_{i=1}^{r+1} q_i = h \tag{16.6.3}$$

where, for the simply laced algebras, h coincides with ψ, the Coxeter number of G.

The exponential with the maximal root is responsible for the massive nature of these field theories. Also for the Toda field theories we can adopt two different ways of looking at the action (16.6.1), depending on the choice of S_0. So, taking for S_0 the action that excludes the maximal root, one has a generalized Liouville theory

$$\mathcal{S}_0 \;=\; \int d^2x \left[\frac{1}{8\pi}\,(\partial_\mu\phi)\cdot(\partial^\mu\phi) + \frac{\mu^2}{16\beta^2}\sum_{i=1}^{r} q_i[\exp(\beta\alpha_i\cdot\phi) - 1]\right]. \tag{16.6.4}$$

Analogously to the cases previously analyzed, these actions describe conformal models. Their quantization requires a set of charges at infinity, encoded in the vector

$$\vec{Q} \;=\; (\beta + 1/\beta)\;\vec{\rho}, \qquad \vec{\rho} \;=\; \frac{1}{2}\sum_{\alpha>0}\alpha. \tag{16.6.5}$$

The analytic component of the stress–energy tensor, given by

$$T(z) \;=\; -\frac{1}{2}(\partial_z\phi)^2 + Q\cdot\partial_z^2\phi,$$

gives rise to the central charge

$$C \;=\; r\,\left[1 + h(h+1)(\beta + 1/\beta)^2\right]. \tag{16.6.6}$$

The second way to approach the Toda field theories consists of using the Feynman perturbation theory. As in previous cases, all the perturbative divergences of these theories come from the tadpole diagrams, which can be eliminated by defining the

[8]For a non-simply laced algebra there is also the possibility to extend the Dynkin diagram of the original theory by adding the shorter maximal root. These are the so-called *twisted* algebras and denoted by $\tilde{G}$. In all the non-twisted models h is equal to the Coxeter number ψ, while for the twisted ones h is equal either to the dual Coxeter number $\tilde{\psi}$ of the same algebra or of another non-simply laced algebra.

normal order of the exponential operators. This induces a renormalization of the mass parameter μ^2.

$$\mu^2 \;\to\; \mu^2 \left(\frac{\Lambda^2}{\mu^2}\right)^{\frac{\beta^2}{4\pi}\frac{\tilde{h}}{h}} \tag{16.6.7}$$

where

$$\tilde{h} \;=\; \frac{1}{2}\sum_{a=1}^{r}\sum_{i=1}^{r+1} q_i \alpha_i^a \alpha_i^a. \tag{16.6.8}$$

In the simply laced algebras $\tilde{h} = h$ and these two numbers get simplified in (16.6.7).

In the Feynman perturbative approach it is necessary to determine the classical values of the masses of the various particles A_a, as coming from the quadratic terms of the lagrangian

$$M_{ab}^2 \;=\; \mu^2 \sum_{i=1}^{r+1} q_i \alpha_i^a \alpha_i^b. \tag{16.6.9}$$

Mass spectrum. The classical mass spectrum is determined by the zeros of the characteristic equation

$$\|M^2 - x \cdot \mathbf{1}\| = 0. \tag{16.6.10}$$

The left-hand side of (16.6.10) is a polynomial of order r, whose general form is

$$\mathcal{P}(x) \;=\; x^r - p_1 x^{r-1} - p_2 x^{r-2} - \ldots - p_r.$$

The first coefficient p_1 is simply the trace of M^2 and, for the simply laced algebra, this is simply twice their Coxeter number. The other coefficients p_i can be expressed in terms of the trace of higher powers of M^2. To simplify the notation, let's impose $\mathcal{M} = M^2$. Their expression is then

$$k\, p_k \;=\; a_k - p_1\, a_{k-1} - \cdots p_{k-1}\, a_1, \tag{16.6.11}$$

where

$$\begin{array}{l} a_1 = \mathrm{Tr}\,\mathcal{M} = \sum_i m_i^2 \\ a_2 = \mathrm{Tr}\,\mathcal{M}^2 = \sum_i m_i^4 \\ \ldots \quad \ldots \\ a_n = \mathrm{Tr}\,\mathcal{M}^n = \sum_i m_i^{2n}. \end{array} \tag{16.6.12}$$

It is convenient to introduce a matrix $\mathcal{N}$ directly linked to the Dynkin diagrams. Its matrix elements are given by

$$\mathcal{N}_{ij} = (q_i \alpha_i, \alpha_j) = \sum_{k=1}^{n} q_i \alpha_i^k \alpha_j^k.$$

It is easy to prove that

$$\mathrm{Tr}\,\mathcal{M}^s = \mathrm{Tr}\,\mathcal{N}^s, \qquad s = 1, 2, \cdots n.$$

Hence, the characteristic equation of $\mathcal{M}$ coincides with that of $\mathcal{N}$. However, $\mathcal{N}$ is a $(n+1)\times(n+1)$ matrix while $\mathcal{M}$ is an $n \times n$ matrix. Since α_0 is expressed by a linear

combination of the other roots, $\mathcal{N}$ is then a singular matrix: one of its eigenvalues vanishes, whereas the others coincide with the eigenvalues of $\mathcal{M}$.

In the basis of its eigenvectors, $M^2_{ij} = \mu_i^2 \delta^{ij}$. The mass spectrum is degenerate if the group of automorphisms of the Dynkin diagram is non-trivial. In this case it may be convenient to organize the particles pairwise, associated to complex conjugate fields. In the simply laced algebra, a remarkable result is that the masses can be organized in a vector

$$\mathbf{m} \,=\, (m_1, m_2, \cdots m_r),$$

which are the eigenvectors of the incidence matrix I of the algebra $\mathcal{G}$. It is defined by $I = 2 - C$, where C is the Cartan matrix. In fact, $\mathbf{m}$ is the Perron–Frobenius eigenvector of I and its components can thus be associated directly to the dots of the Dynkin diagram itself. On the other hand, since the dots of the Dynkin diagram are also associated to the fundamental representation of the simply laced algebra, we arrive at the interesting conclusion that there is a correspondence between the particles of mass m_i and the relative representations of $\mathcal{G}$. This will be a useful observation in the future discussion of the scattering processes of these theories.

Let's now discuss in detail the mass spectrum of the various Toda field theories, with the final result of this analysis summarized in Tables 16.2 and 16.3 below.

16.6.1 $A_n^{(1)}$ Series

For this series, the matrix $\mathcal{N}$ reduces to the Cartan matrix of the affine Lie algebras. The characteristic equation associated to $\mathcal{N}$ is given by

$$\mathcal{Q}_{n+1}(x) = \begin{vmatrix} 2-x & -1 & 0 & \cdots & 0 & 0 & -1 \\ -1 & 2-x & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2-x & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 & 2-x & -1 \\ -1 & 0 & 0 & \cdots & 0 & -1 & 2-x \end{vmatrix}.$$

Imposing $2y = 2 - x$, it is possible to show that

$$\mathcal{Q}_{n+1} \,=\, 2\,(\mathcal{T}_{n+1}(y) - 1), \tag{16.6.13}$$

where $\mathcal{T}_{n+1}(y)$ is the Chebyshev polynomial of the first type

$$\mathcal{T}_{n+1}(\cos\theta) \,=\, \cos(n+1)\theta.$$

The mass spectrum of the series $A_n^{(1)}$ is given by the n non-vanishing roots of the equation $\mathcal{T}_{n+1}(y) = 1$, namely

$$m_k^2 \,=\, 4\sin^2\frac{k\pi}{n+1} \qquad k = 1, 2, \cdots n. \tag{16.6.14}$$

16.6.2 $D_n^{(1)}$ Series

For this series, we have

$$\mathcal{N} = \begin{vmatrix} 4 & -2 & 0 \cdots & 0 & 0 & -2 & -2 \\ -2 & 4 & 2 \cdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & -1 & 0 \cdots & 0 & 0 & 2 & 0 \\ -1 & 0 & 0 \cdots & 0 & 0 & 0 & 2 \end{vmatrix}.$$

The characteristic equation has the form

$$\|\mathcal{N} - x \cdot \mathbf{1}\| = 2^{n+2}(y-1)(2y-1)^2\,\mathcal{U}_{n-2}(y) = 0, \tag{16.6.15}$$

where $x = 4(1-y)$ and $\mathcal{U}_m$ is the Chebyshev polynomial of the second kind. The roots of (16.6.15) are given by

$$\begin{array}{rcl} y_{n+1} = 1 & \rightarrow & x_{n+1} = 0 \\ y_n = \frac{1}{2} & \rightarrow & x_n = 2 \\ y_{n-1} = \frac{1}{2} & \rightarrow & x_{n-1} = 2 \end{array} \tag{16.6.16}$$

and by $\mathcal{U}_{n-2}(y) = 0$, i.e.

$$y_k = \cos\frac{k\pi}{n-1} \quad \rightarrow \quad x_k = 8\sin^2\frac{k\pi}{2(n-1)}, \qquad k = 1, 2, \cdots n-2. \tag{16.6.17}$$

The first root in (16.6.16) is irrelevant for the spectrum. The spectrum is reported in Table 16.2.

16.6.3 E_n Series

The analysis of these exceptional algebras has to be done separately for each of them.

1. The characteristic equation for the E_6 algebra is

$$\begin{aligned} \|\mathcal{M} - x \cdot \mathbf{1}\| &= x^6 - 24x^5 + 216x^4 - 936x^3 + 2052x^2 - 2160x + 864 \\ &= [x^2 - 12x + 24]\,[x^2 - 6x + 6]^2. \end{aligned} \tag{16.6.18}$$

   There are two doublets of degenerate particles, plus two other particles of different masses. The spectrum is given in Table 16.2.

2. The characteristic equation of the Toda field theory based on E_7 is

$$\begin{aligned} \|\mathcal{M} - x \cdot \mathbf{1}\| &= x^7 - 36x^6 + 504x^5 - 3552x^4 \\ &\quad +13536x^3 - 27648x^2 + 27648x - 10368 \\ &= [x-6]\,[x^3 - 18x^2 + 72x - 72] \\ &\quad \times [x^3 - 12x^2 + 36x - 24]. \end{aligned} \tag{16.6.19}$$

   The mass spectrum can be found in Table 16.2. Thanks to the Z_2 automorphism of the Dynkin diagram of the affine E_7 algebra, the particles can be classified into even and odd particles with respect this Z_2 symmetry.

Table 16.2: *Masses of the Toda field theories related by the folding procedure.*

$A_{2r}^{(1)}$ $2M\sin(\frac{\pi i}{2r+1}),\quad 1\le i\le 2r$	$A_{2r}^{(2)}$ $4M\sin(\frac{\pi i}{2r+1}),\quad 1\le i\le r$
$D_{r+1}^{(1)}$ $M, M, 2M\sin(\frac{\pi i}{2r}),\quad 1\le i\le r-1$	$B_r^{(1)}$ $M, 2M\sin(\frac{\pi i}{2r}),\quad 1\le i\le r-1$
$D_{r+2}^{(1)}$ $M, M, 2M\sin(\frac{\pi i}{2r+2}),\quad 1\le i\le r$	$D_{r+1}^{(2)}\equiv\tilde{B}_r$ $\sqrt{2}M\sin(\frac{\pi i}{2r+2}),\quad 1\le i\le r$
$A_{2r-1}^{(1)}$ $2M\sin(\frac{\pi i}{2r}),\quad 1\le i\le 2r-1$	$C_r^{(1)}$ $2M\sin\left(\frac{\pi i}{2r}\right),\quad 1\le i\le r$
$D_{2r}^{(1)}$ $M, M, 2M\sin(\frac{\pi i}{2(2r-1)}),\ 1\le i\le 2r-2$	$A_{2r-1}^{(2)}\equiv\tilde{C}_r$ $\frac{M}{\sqrt{2}}, \sqrt{2}M\sin(\frac{\pi i}{(2r-1)}),\ 1\le i\le r-1$
$D_4^{(1)}$ $M, M, M, \sqrt{3}M$	$G_2^{(1)}$ $M, \sqrt{3}M$
$E_6^{(1)}$ $m_1 = m_{\overline{1}} = M$ $m_2 = m_{\overline{2}} = 2M\ \cos(\frac{\pi}{12})$ $m_3 = 2M\ \cos(\frac{\pi}{4})$ $m_4 = 4M\ \cos(\frac{\pi}{12})\cos(\frac{\pi}{4})$	$D_4^{(3)}\equiv\tilde{G}_2$ $m_3,\ m_4$ $F_4^{(1)}$ $m_1,\ m_2,\ m_3,\ m_4$
$E_7^{(1)}$ $m_1 = M$ $m_2 = 2M\cos(\frac{5\pi}{18})$ $m_3 = 2M\cos(\frac{\pi}{9})$ $m_4 = 2M\cos(\frac{\pi}{18})$ $m_5 = 4M\cos(\frac{5\pi}{18})\cos(\frac{\pi}{18})$ $m_6 = 4M\cos(\frac{\pi}{9})\cos(\frac{2\pi}{9})$ $m_7 = 4M\cos(\frac{\pi}{18})\cos(\frac{\pi}{9})$	$E_6^{(2)}\equiv\tilde{F}_4$ $m_2,\ m_4,\ m_5,\ m_6$

3. For the Toda field theory on E_8 we have

$$\begin{aligned}\|\mathcal{M}-x\,1\| &= x^8-60x^7+1440x^6-18000x^5+1257440x^4\\ &\quad -518400x^3+1166400x^2-1296000x+518400\\ &= [x^4-30x^3+240x^2-720x+720]\\ &\quad \times[x^4-30x^3+300x^2-1080x+720]\end{aligned}\tag{16.6.20}$$

The masses of this theory are reported in Table 16.3.

These cases cover all the simply laced Toda field theories. A similar analysis can also be done for those defined by the non-simply laced algebra by using the foldings, and the final results are collected in Table 16.2.

Table 16.3: *Mass spectrum of the Toda field theory* $E_8^{(1)}$.

$E_8^{(1)}$
$m_1 = M$
$m_2 = 2M\cos(\frac{\pi}{5})$
$m_3 = 2M\cos(\frac{\pi}{30})$
$m_4 = 2m_2\cos(\frac{7\pi}{30})$
$m_5 = 2m_2\cos(\frac{2\pi}{15})$
$m_6 = 2m_2\cos(\frac{\pi}{30})$
$m_7 = 4m_2\cos 2(\frac{\pi}{5})\cos(\frac{7\pi}{30})$
$m_8 = 4m_2\cos(\frac{\pi}{5})\cos(\frac{2\pi}{15})$

Coupling constants. After the mass term, the next perturbation consists of the three-particle coupling constants

$$f^{abc} = \mu^2\beta\sum_i q_i\alpha_i^a\alpha_i^b\alpha_i^c. \tag{16.6.21}$$

These expressions enjoy a series of interesting geometrical properties. First of all, it is possible to prove that they vanish if it is impossible to draw a triangle with sides of m_a, m_b, and m_c whose internal angles are rational fractions of π. This can be seen as a natural consequence of the algebraic nature of the values of the masses. Moreover, the quantities f^{abc} vanish if they do not respect a discrete symmetry of the affine Dynkin diagram. Consider, for instance, the symmetry Z_2 of $E_7^{(1)}$: if two of the indices of f^{abc} refer to two even particles and the third one to an odd particle, this coupling constant clearly vanishes. Finally, when they are different from zero, the quantities f^{abc} are proportional to the area $\mathcal{A}^{abc}$ of the aforementioned mass triangle. For the simply laced algebra we have

$$\left|f^{abc}\right| = \frac{4\,\mu^2\,\beta}{\sqrt{h}}\,\mathcal{A}^{abc}. \tag{16.6.22}$$

Obviously the non-vanishing values of f^{abc} indicate the possible scattering processes in which the particles and their bound states enter. In fact, with $f^{abc} \neq 0$ we can have the process in Fig. 16.6: in the collision, the initial particles A_a and A_b form a bound state A_c that decays into the same particles as the final state. From the symmetry of the indices of f^{abc} we immediately infer that the same scenario occurs for the processes of the crossed channels, namely the particle A_a can be regarded as bound state of the particle A_b and A_c, and the particle A_b may be regarded as bound state of the particles A_a and A_c.

There are other n-particle vertices of the pertubation theory coming from the series expansions of the exponential terms. Interestingly enough, they admit a geometrical interpretations in terms of the moments of a distribuition of a set of positive electric charges $\{q_i\}$, placed at the points indicated by the vectors α_i. Adopting this interpretation, the total charge of the system is h. The condition (16.6.3) is then nothing else but the definition of the center reference frame of the charges while the diagonalization

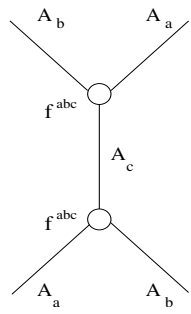


Fig. 16.6 *Scattering process of the particles A_a and A_b that gives rise to the bound state given by the particle A_c.*

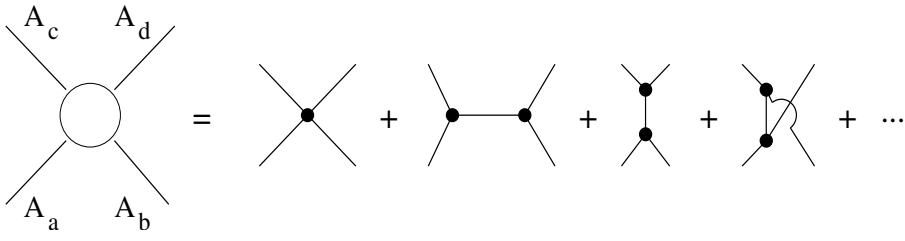


Fig. 16.7 *Scattering process of the particles A_a and A_b with final state given by the particles A_c and A_d. If $A_c\, A_d \neq A_a\, A_b$, the four-particle vertex f^{abcd} of the first graph cancels the sum of the other Feynman graphs whose internal propagators are made of all the particles allowed by the three-particle vertices $f^{ijk} \neq 0$.*

of (16.6.9) is equivalent to the choice of the coordinates along the principal axes of the ellipsoid defined by the quadrupole moments.

The Toda field theories possess an infinite set of conserved currents, both at the classical and quantum level. For our scope, rather than their explicit expressions, it is sufficient to know the spectrum of their spins s. This is given by the Coxeter exponents of the Lie algebra under investigation, modulo the Coxeter number. The sets of these values is given in Table 16.4. The quantum integrability of the Toda field theories has an extremely important consequence for the scattering processes in which are involved the particles A_i of these theories, namely their elasticity. This property is already manifest at the tree level of the scattering processes of two initial particles $A_a\, A_b$ going into two final particles $A_c\, A_d$: indeed, the only non-vanishing amplitudes are those in which the final particles coincide with the initial ones. At the lowest order in the coupling constant β, this amplitude is ruled by the sum of the Feynman graphs shown in Fig. 16.7 and this sum vanishes unless the final particles $A_c\, A_d$ are equal to the initial ones (in the propagators of the last three diagrams it enters all particles that are compatible with $f^{ijk} \neq 0$).

Table 16.4: *Coxeter numbers and Coxeter exponents of the affine Dynkin diagrams.*

Algebra	ψ	Exponents
$A_r^{(1)}$	$r+1$	$1, 2, \cdots, r$
$A_{2r}^{(2)} \equiv A_{2r}/Z_2$	$4r+2$	$1, 3, 5, \cdots, 2r-1, 2r+3, \cdots, 4r+1$
$B_r^{(1)}$	$2r$	$1, 3, 5, \cdots, 2r-1$
$\tilde{B}_r \equiv D_{r+1}^{(2)}$	$2r+2$	$1, 3, 5, \cdots, 2r+1$
$C_r^{(1)}$	$2r$	$1, 3, 5, \cdots, 2r-1$
$\tilde{C}_r \equiv A_{2r-1}^{(2)}$	$4r-2$	$1, 3, 5, \cdots, 4r-3$
$D_r^{(1)}$	$2r-2$	$1, 3, 5, \cdots, 2r-3, r-1$
$E_6^{(1)}$	12	$1, 4, 5, 7, 8, 11$
$E_7^{(1)}$	18	$1, 5, 7, 9, 11, 13, 17$
$E_8^{(1)}$	30	$1, 7, 11, 13, 17, 19, 23, 29$
$G_2^{(1)}$	6	$1, 5$
$\tilde{G}_2 \equiv D_4^{(3)}$	12	$1, 5, 7, 11$
$F_4^{(1)}$	12	$1, 5, 7, 11$
$\tilde{F}_4 \equiv E_6^{(2)}$	18	$1, 5, 7, 11, 13, 17$

16.7 Toda Field Theories with Imaginary Coupling Constant

If we make the analytic continuation $\beta \to i\beta$ in the previous action of the Toda field theories, we arrive, in general, at a complex action (the only real case is for the algebra $SU(2)$ that gives rise to the Sine–Gordon model). Even though the interpretation of these theories having a complex action is problematic from the point of view of a standard quantum field theory quantization, it can nevertheless be shown that, for particular values of β, an opportune restriction of their Hilbert space leads to the definition of consistent models. Note that, with this transformation, the Liouville part of these theories is associated to a conformal field theory with a value of the central charge less than the rank r of the algebra. Choosing the discrete values

$$\beta^2 = \frac{p}{p+1} \quad p = k+h, k+h+1, \ldots$$

for the central charge we have

$$c = r\left[1 - \frac{h(h+1)}{p(p+1)}\right]. \tag{16.7.1}$$

This value corresponds to the conformal theory constructed on the coset[9]

$$\frac{G_k \times G_1}{G_{k+1}}.$$

[9]To compare with the formulas of Chapter 13 one should recall that the dimension $|G|$ of the algebra is related to its rank r and the Coxeter number ψ by the relation $|G| = r(\psi+1)$.

In these theories, the vertex operator associated to the maximal root

$$V_{\alpha_{max}} = e^{i\beta\alpha_{r+1}\cdot\phi},$$

i.e. the perturbing operator of the conformal theory, has conformal weight

$$\Delta_{\alpha_{max}} = 1 - \frac{h}{k+h+1}. \tag{16.7.2}$$

Let's discuss some significant examples.

- Taking the algebra E_8 and $k = 1$, we have

$$c = \frac{1}{2}, \qquad \Delta_{max} = \frac{1}{16}. \tag{16.7.3}$$

  Hence the Toda field theory associated to this value of the imaginary coupling constant corresponds to the magnetic deformation of the Ising model.
- Taking the algebra E_7 and $k = 1$, we have

$$c = \frac{7}{10}, \qquad \Delta_{max} = \frac{1}{10}. \tag{16.7.4}$$

  Hence, in this case, the relative Toda field theory with imaginary coupling constant corresponds to the thermal deformation of the tricritical Ising model.
- With the algebra E_6 and $k = 1$, we have

$$c = \frac{6}{7}, \qquad \Delta_{max} = \frac{1}{7}. \tag{16.7.5}$$

  This theory corresponds to the thermal deformation of the tricritical three-state Potts model.

16.8 Deformation of Conformal Conservation Laws

In this section we set a criterion to establish whether a deformation of a conformal theory gives rise to an integrable model or not, away from criticality. To first order in the coupling constant, this criterion is based on the operator product expansion and on the formula of the conformal characters. When the integrals of motion belong to the conformal family of the identity operator, the corresponding analysis can be carried out in a purely algebraic way.

16.8.1 Operator Product Expansion

Consider a conformal minimal model $\mathcal{M}_{p,q}$ that is deformed by a relevant primary scalar field $\Phi_{lk}(z,\bar{z}) = \phi_{lk}(z)\bar{\phi}_{lk}(\bar{z})$, with anomalous dimension $x = 2\Delta < 2$. The perturbed action is

$$\mathcal{S} = \mathcal{S}_0 + \lambda \int \Phi_{lk}(z,\bar{z})\, d^2z.$$

Let $\mathcal{C}_{s+1}(z)$ be a conserved current of the conformal model $\mathcal{M}_{p,q}$ ($\partial_{\bar{z}}\,\mathcal{C}_s(z) = 0$) of spin $s+1$ (which we assume to be either an integer or fractional number), local with respect to Φ_{lk}:

$$C_{s+1}(z)\Phi_{lk}(w,\bar{w}) \;=\; \sum_{n=2}^{m} \frac{d_{lk}^{(n)}}{(z-w)^n}\Phi_{lk}^{(n)}(w,\bar{w}) + \frac{1}{z-w}B_{lk}(w,\bar{w}) + \cdots \qquad (16.8.1)$$

where n is an integer, $\Phi_{lk}^{(n)}$ and B_{lk} are the descendent fields of Φ_{lk}, while $d_{lk}^{(n)}$ denote here the structure constants of this operator product expansion. The Ward identity for the current $\mathcal{C}_s(z,\bar{z})$ can be expressed in terms of the conformal Ward identity

$$\begin{aligned}\langle\, C_{s+1}(z,\bar{z})\cdots\rangle \;=\; &\langle C_{s+1}(z)\cdots\rangle_0 \qquad (16.8.2)\\ &+\lambda\int dw\,d\bar{w}\,\langle C_{s+1}(z)\Phi_{lk}(w,\bar{w})\cdots\rangle_0 \;+\mathcal{O}(\lambda^2).\end{aligned}$$

To first order in λ, eqns (16.8.1) and (16.8.2), together with the identity

$$\partial_{\bar{z}}\,\frac{1}{z-w+i\epsilon} \;=\; \delta(z-w)\delta(\bar{z}-\bar{w}),$$

give rise to

$$\partial_{\bar{z}}\;C_{s+1}(z,\bar{z}) \;=\; \lambda\left(\; B_{lk}(z,\bar{z}) - d_{lk}^{(2)}\partial_z\Phi_{lk}^{(2)}\right). \qquad (16.8.3)$$

The existence of a conservation law away from the critical point only depends on whether B_{lk} is a total derivative with respect to z. The simplest example is provided by the stress–energy tensor: if $\mathcal{C}_2 = T$, then

$$B_{lk} - d_{lk}^{(2)}\partial_z\Phi_{lk}^{(2)} \;=\; (1-\Delta)\;\partial_z\Phi_{lk}(z,\bar{z})$$

and, in this case, we have

$$\partial_{\bar{z}}\;T(z,\bar{z}) \;=\; -\frac{1}{4}\partial_z\;\Theta, \qquad \Theta \;=\; -4\lambda\;(1-\Delta)\Phi_{lk}(z,\bar{z}).$$

The corresponding conserved charge is expressed by

$$\mathcal{Q}_1 \;=\; \int (T\,dz + \frac{1}{4}\Theta\,d\bar{z}).$$

Let's see some other significant examples.

1. The minimal model $\mathcal{M}_{4,5}$ corresponds to the universality class of the tricritical Ising model. On the other hand, this model is also the first of the superconformal series. Let's choose then as $\mathcal{C}_s$ the supercurrent $G_{3/2}$ of spin $s = \frac{3}{2}$ and as deformation the vacancy density, i.e. the operator $\Phi_{13} = \Phi_{\frac{3}{5},\frac{3}{5}}$. In the following we

will use the notation $\Phi_{\Delta,\bar{\Delta}}$ for the conformal fields. The supersymmetric operator product expansion

$$G(z_1)\Phi_{\frac{3}{5},\frac{3}{5}}(z_2,\bar{z}_2) \qquad = \left(\frac{1}{5z_{12}^2}+\frac{1}{z_{12}}\partial_2\right)\Phi_{\frac{1}{10},\frac{3}{5}}(z_2,\bar{z}_2)+\cdots$$

($z_{12}\equiv z_1-z_2$) leads to the conservation law

$$\partial_{\bar{z}}G(z,\bar{z}) \;=\; \partial_z\bar{\Psi}(z,\bar{z}), \qquad \bar{\Psi}(z,\bar{z}) \;=\; \frac{4}{5}\;\lambda\;\Phi_{\frac{1}{10},\frac{3}{5}}(z,\bar{z}).$$

The corresponding charge has spin $s=\frac{1}{2}$

$$\mathcal{Q}_{\frac{1}{2}} \;\equiv\; Q \;=\; \int (G\;dz \;+\; \bar{\Psi}\;d\bar{z}).$$

Using the operator product expansion

$$G(z_1)G(z_2) \;=\; \frac{2}{z_{12}}\;T(z_2)+\cdots$$
$$G(z_1)\;\Phi_{\frac{1}{10},\frac{1}{10}}(z_2,\bar{z}_2) \;=\; \frac{1}{z_{12}}\;\Phi_{\frac{3}{5},\frac{1}{10}}(z_2,\bar{z}_2)+\cdots$$

it is easy to show that

$$Q^2 = \int dz_1dz_2\;G(z_1)G(z_2)+\frac{4}{5}\lambda\int d\bar{z}_1d\bar{z}_2\left\{G(z_1\bar{z}_1),\Phi_{\frac{1}{10},\frac{3}{5}}(z_2,\bar{z}_2)\right\}$$
$$= \int (2\;T\;dz+\frac{4}{5}\lambda\;\Phi_{\frac{3}{5},\frac{3}{5}}\;d\bar{z}) \;=\; 2\mathcal{P}.$$

In addition to Q, one can similarly prove the conservation of $\bar{Q}$, which is constructed starting with the anti-analytic component $\bar{G}_{3/2}$ of the supercurrent. In this case we have $\bar{Q}^2=2\bar{\mathcal{P}}$, where $\bar{\mathcal{P}}=E-P$. Finally

$$Q\;\bar{Q}+\bar{Q}\;Q \;=\; \frac{4}{5}\;\lambda\;\int\left[\left(\;\partial_z\;\Phi_{\frac{1}{10},\frac{1}{10}}\right)\;dz+\left(\partial_{\bar{z}}\;\Phi_{\frac{1}{10},\frac{1}{10}}\right)d\bar{z}\right] \;=\; \mathcal{T}. \qquad (16.8.4)$$

The right-hand side of this equation is the topological charge $\mathcal{T}$. In fact, the tricritical Ising model perturbed by the vacancy density operator is driven, for

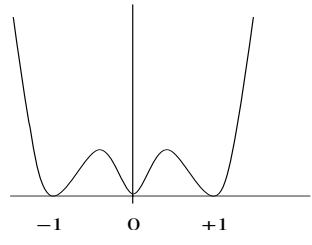


Fig. 16.8 *Effective potential of the tricritical Ising model perturbed by* $\Phi_{\frac{3}{5},\frac{3}{5}}$ *with* $\lambda<0$. *The off-critical model has solitonic excitations that interpolate between two nearest vacua.*

$\lambda < 0$, in a phase where there are three degenerate vacua, as shown in Fig. 16.8. The system has therefore solitonic excitations that interpolate between two nearest vacua and that are characterized by their topological charge. The integrability of this theory implies the elasticity of the scattering processes in which are involved the solitons. The charges $\{Q, \bar{Q}, P, \bar{P}\}$ generate a global supersymmetry of this model away from criticality.

2. The universality class of the tricritical three-state Potts model corresponds to a subalgebra of the minimal mode $\mathcal{M}_{6,7}$, as the universality class of the three-state Potts model corresponds to a subalgebra of $\mathcal{M}_{5,6}$. Let's choose as $\mathcal{C}_s$ the chiral field W of spin $s = 5$ and as deformation $\Phi_{12}(z, \bar{z}) = \Phi_{\frac{1}{7},\frac{1}{7}}$. The operator expansion

$$\mathcal{W}(z_1)\, \Phi_{\frac{1}{7},\frac{1}{7}}(z_2) \;=\; \left(\frac{w_0}{z_{12}^2} + \frac{1}{z_{12}} \partial_2 \right) \Phi_{\frac{22}{7},\frac{1}{7}}(z_2, \bar{z}_2) + \cdots$$

(where w_0 is a constant) gives rise to a conserved charge of spin 4

$$\mathcal{Q}_4 \;=\; \int (\, \mathcal{W}\, dz + \Lambda\, d\bar{z}), \qquad \Lambda \;=\; (w_0 - \frac{2}{7})\, \Phi_{\frac{22}{7},\frac{1}{7}}.$$

In this case, Φ_{12} is the scaling operator corresponding to the energy density of the lattice model. Hence its insertion into the action moves the temperature of the system away from its critical value. This perturbation preserves the permutation symmetry $S_3 = Z_2 \otimes Z_3$ of the model, generated by C (the charge conjugate operator) and ϑ, with

$$C^2 \;=\; \vartheta^3 \;=\; 1.$$

$\mathcal{Q}_4$ is an odd operator under C, i.e. $C\, \mathcal{Q}_4\, C \;=\; -\mathcal{Q}_4$, while the first conserved charge given by the total momentum $\mathcal{P}$ is an even operator, $C\, \mathcal{P}\, C \;=\; \mathcal{P}$.

16.8.2 Integrals of Motion of the Identity Family

It is possible to set up an efficient algebraic method to identify the integrals of motion coming from the conformal family of the identity operator. We need first to recall that in the conformal space the Virasoro operator L_{-1} acts as a derivative with respect to the analytic coordinate, i.e. $L_{-1} \to \partial_z$. Let's define $\hat{\Lambda}_{s+1} = \Lambda_{s+1}/L_{-1}\Lambda_s$ as the space of quasi-primary operators at the level $s+1$ of the conformal family $[\mathbf{I}]$ of the identity. Let $T_{s+1}^{(k)}$ be the vector basis of this space: their expressions consists of appropriate polynomials in L_{-n}:

$$T_{s+1}^{(k)} \;=\; L_{n_1}\, L_{-n_2} \cdots L_{-n_k}\, \mathbf{I}\,, \qquad \sum_i n_i \;=\; s+1 \tag{16.8.5}$$

with the first representatives given in Table 16.5.

The eventual conserved current will be constructed in terms of linear combinations of these vectors of the basis. Note that the operator $\partial_{\bar{z}}$ can be interpreted as a mapping

Table 16.5: *Dimensionality and vectors of the basis of Λ_n for $n \leq 6$.*

s	0	1	2	3	4	5	6
dim $\hat{\Lambda}_n$	1	0	1	0	1	0	2
Vectors	$\mathbf{I}$	–	$L_{-2}\mathbf{I}$	–	$T_4 = L_{-2}^2\mathbf{I}$	–	$T_6^{(1)} = L_{-2}^3\mathbf{I}$ $T_6^{(2)} = L_{-3}^2\mathbf{I}$

from the space $\hat{\Lambda}_{s+1}$ to the space of the operators at the level s of the perturbing field[10]

$$\partial_{\bar{z}} T_{s+1}^{(k)}(z,\bar{z}) \,=\, \lambda\, R_s^{(k)}(z,\bar{z}), \qquad \partial_{\bar{z}} : \hat{\Lambda}_{s+1} \to \Phi_s, \tag{16.8.6}$$

with the operator $R_s^{(k)}$ explicitly expressed by

$$R_s^{(k)}(z,\bar{z}) \,=\, \oint_z \frac{d\xi}{2\pi i}\, T_{s+1}^{(k)}(z)\, \Phi_{lk}(\xi,\bar{z}).$$

Since the contour integral of two operators corresponds to computing their commutator (see Chapter 10), we also have

$$R_s^{(k)}(z,\bar{z}) \,=\, \left[T_{s+1}^{(k)}(z), \int d\xi\, \Phi_{lk}(\xi,\bar{z})\right].$$

In addition to $\partial_{\bar{z}}$ we can also introduce an infinite family of operators D_n that map the family Λ of the identity operator into the space of the perturbing field

$$D_n\, \Lambda(z,\bar{z}) \,\equiv\, \oint_z \frac{d\xi}{2\pi i} \Lambda(z)\, (\xi - z)^n\, \Phi_{lk}(\xi,\bar{z}), \tag{16.8.7}$$

with $D_0 \,=\, \partial_{\bar{z}}$. Since the primary field Φ_{lk} satisfies

$$[L_n, \Phi_{lk}(\xi,\bar{\xi})] \,=\, \left[(\xi - z)^{n+1}\, \partial_\xi + \Delta\, (n+1)\, (\xi - z)^n\right]\, \Phi_{lk}(\xi,\bar{\xi}),$$

we have the relations

$$\begin{aligned} [L_n, D_m] \,&=\, -\,(m + (1-\Delta)(n+1))\; D_{n+m}, \\ D_{-m}\mathbf{I} \,&=\, \frac{1}{(m+1)!}\, L_{-1}^{m+1}\, \Phi_{lk}(z,\bar{z}). \end{aligned} \tag{16.8.8}$$

These equations allow us to easily compute $R_s^{(k)}$. For instance, choosing $T_2 \,=\, T \,=\, L_{-2}\mathbf{I}$, we have

$$\partial_{\bar{z}} T \,=\, \lambda D_0\, L_{-2}\, \mathbf{I} \,=\, \lambda(\Delta - 1)\, D_{-2}\, \mathbf{I} \,=\, \lambda\, (\Delta - 1)\, L_{-1}\, \Phi_{lk}(z,\bar{z}),$$

and, since $L_{-1}\,[\ldots] \,=\, \partial_z\,[\ldots]$, we recover the conservation law of the stress–energy tensor.

[10] We recall that the spin s measures the difference between the analytic and anti-analytic indices of the densities.

Consider now the quasi-primary field of spin 4 of the identity family $T_4 = (T^2) = L_{-2}^2 \mathbf{I}$. Let's compute $\partial_{\bar{z}} T_4$ with the rules given above:

$$\partial_{\bar{z}} T_4 = \lambda D_0 L_{-2} L_{-2} \mathbf{I} = \lambda(\Delta - 1)(D_{-2} L_{-2} + L_{-2} D_{-2}) \mathbf{I}$$

$$= \lambda(\Delta - 1) \left(2L_{-2} L_{-1} + \frac{\Delta - 3}{6} L_{-1}^3 \right) \Phi_{lk}$$

$$= \lambda(\Delta - 1) \left(-2L_{-3} + 2L_{-1} L_{-2} + \frac{\Delta - 3}{6} L_{-1}^3 \right) \Phi_{lk}$$

For a generic operator Φ_{lk}, the right-hand side is not a total derivative for the presence of the operator L_{-3} and, consequently, there is no conservation law. However, if the perturbing field coincides with the operator $\Phi_{1,3}$, the null-vector equation of this field at level 3

$$\left(L_{-3} - \frac{2}{\Delta + 2} L_{-1} L_{-2} + \frac{1}{(\Delta + 1)(\Delta + 2)} L_{-1}^3 \right) \Phi_{1,3} = 0, \tag{16.8.9}$$

allows us to re-express L_{-3}, arriving then at the conservation law

$$\partial_{\bar{z}} T_4 = \partial_z \Theta_2,$$

with

$$\Theta_2 = \lambda \frac{\Delta - 1}{\Delta + 2} \left\{ 2\Delta L_{-2} + \frac{(\Delta - 2)(\Delta - 1)(\Delta + 3)}{6(\Delta + 1)} L_{-1}^2 \right\} \Phi_{1,3}.$$

The conserved charge $\mathcal{Q}_3$ commutes with $\mathcal{Q}_1$, as can be shown using the commutation relations of the L_n's. Using eqn (16.8.9) and the other null-vector equations satisfied by $\Phi_{1,3}$, it is possible to prove that there are infinite conserved currents for all odd integer values of the spin s. Their expressions coincide with the analogous expressions of the Sine–Gordon model, eqn (16.3.4), a fact that should not be surprising in the light of the relationship between the Sine–Gordon model and the $\Phi_{1,3}$ deformation of the minimal models.

If the perturbing field is either $\Phi_{1,2}$ or $\Phi_{2,1}$, the first non-trivial conservation law is obtained by the following linear combination of the quasi-primary fields of spin 6:

$$T_6 = T_6^{(1)} + a\, T_6^{(2)}, \qquad a = \frac{18}{2\Delta + 1} + \Delta - 2. \tag{16.8.10}$$

For the null-vector equation of these fields

$$\left(L_{-2} - \frac{3}{2(2\Delta + 1)} \right) \Phi = 0$$

we have in fact

$$\partial_{\bar{z}} T_6 = \partial \Theta_4.$$

The explicit expression Θ_4 is proposed as an exercise in Problem 5. For these operators, it can be shown that other conserved currents are obtained for the values of the spin

$$s = 1, 5 \pmod 6. \tag{16.8.11}$$

16.8.3 Counting Argument

All the examples discussed above have illustrated the importance of the operator product expansion for defining the conserved currents with lower values of the spin. An extremely powerful method to establish a sufficient condition of their existence, without bothering to explicitly compute them, has been introduced by A.B. Zamolodchikov. It goes under the name of a *counting argument*. The following discussion focuses on the conserved currents coming from the identity operator although analogous results can be easily established by considering other conserved currents coming from conformal families of other generators that are local with respect to the perturbing field Φ.

Let $\hat{\Lambda}_{s+1}$ be the space of quasi-primary descendant fields of the identity operator and $\hat{\Phi}_s$ the quotient space at level s of the perturbing field

$$\hat{\Phi}_s \;=\; \Phi_s / L_{-1}\, \Phi_{s-1}.$$

The linear map

$$\partial_{\bar{z}} \quad : \quad \hat{T}_{s+1} \rightarrow \lambda\, \hat{\Phi}_s$$

clearly has a non-zero kernel when

$$\dim \hat{T}_{s+1} \;>\; \dim \hat{\Phi}_s. \tag{16.8.12}$$

If this condition is fulfilled, then there are necessarily some fields $T_{s+1}(z,\bar{z}) \in \hat{T}_{s+1}$ and $\Phi_{s-1}(z,\bar{z}) \in \hat{\Phi}_{s-1}$ such that

$$\partial_z\, T_{s+1}(z,\bar{z}) \;=\; \lambda\, \partial_{\bar{z}}\, \Phi_{s-1}(z,\bar{z}),$$

i.e. there is a conserved current of spin s. It is easy to check the condition (16.8.12) by computing the dimension of the involved spaces by means of the conformal characters

$$\sum_{s=0}^{\infty} q^s \dim \hat{T}_n \;=\; (1-q)\, \tilde{\chi}_{1,1}(q) + q,$$
$$\sum_{s=0}^{\infty} q^{s+\Delta_{kl}} \dim(\hat{\Phi}_{k,l})_s \;=\; (1-q)\, \tilde{\chi}_{k,l}(q),$$

where

$$\tilde{\chi}_{r,s}(q) \;=\; q^{\frac{(c-1)}{24} - \Delta_{r,s}}\, \chi_{r,s}(q),$$

with $\chi_{r,s}(q)$ the character of the field $\Phi_{r,s}$, whose explicit expression was presented in Chapter 11.

The counting argument provides useful information on the structure of the conserved currents only for values of low enough spin.[11] Using the counting argument it is easy to prove the existence of non-trivial integrals of motion for the deformations of the minimal models induced by the operators $\Phi_{1,3}$, $\Phi_{1,2}$, and $\Phi_{2,1}$. Hence, these deformations always define integrable models away from criticality.

[11] The reason is that the dimension of the higher level spaces of $\Phi_{r,s}$ asymptotically grows faster than the dimension of the same spaces coming from the identity operator.

16.8.4 Examples

Let's present some examples of the application of the counting argument.

- The first example comes from a model analyzed in the previous section, i.e. the thermal deformation of the tricritical three-state Potts model. In this case, there are two classes of conserved charges. The first class has its origin in the family of the identity operator, while the second class comes from the descendants of $\mathcal{W}$. These two classes are also distinguished by their quantum number under the charge conjugate operator C. The result is

$$\begin{aligned} \dim \hat{T}_{s+1} &> \dim (\hat{\Phi}_{\frac{1}{7},\frac{1}{7}})_s \quad \text{for} \quad s = 1,5,7,11 \qquad (C\text{even}) \\ \dim \hat{W}_{s+1} &> \dim (\hat{\Phi}_{\frac{22}{7},\frac{1}{7}})_s \quad \text{for} \quad s = 4,8 \qquad (C\text{odd}). \end{aligned}$$

  In light of these results, it is natural to conjecture that the spectrum of the spin of the conserved charges is given by

$$s \;=\; 1,4,5,7,8,11 \quad (\text{mod } 12). \tag{16.8.13}$$

  These values of the spin coincide with the Coxeter exponents of E_6, modulo the Coxeter number of this algebra. The presence of this algebra should not be surprising for the additional symmetry of this model, which can also be defined in terms of the coset $(E_6)_1 \otimes (E_6)_1/(E_6)_2$.
- An analogous computation for the tricritical Ising model ($\mathcal{M}_{4,5}$) perturbed by the energy operator $\Phi_{1,2} = \Phi_{\frac{1}{10},\frac{1}{10}}$ gives

$$\dim \hat{T}_{s+1} > \dim (\hat{\Phi}_{\frac{1}{10},\frac{1}{10}})_s \qquad \text{for} \quad s = 1,5,7,9,11,13 \ .$$

  These values of s coincide with the first Coxeter exponents of E_7. It is natural to conjecture that the full spectrum of the spins of the conserved charges is given in this case by

$$s = 1,5,7,9,11,13,17 \quad (\text{mod } 18) \tag{16.8.14}$$

  where 18 is the Coxeter number of E_7. This structure of the spins is obviously related to the coset realization $\frac{(E_7)_1 \otimes (E_7)_1}{(E_7)_2}$ of the model.
- For the Ising model ($\mathcal{M}_{3,4}$) perturbed by the magnetization operator $\Phi_{1,2} = \Phi_{\frac{1}{16},\frac{1}{16}}$, we have

$$\dim \hat{T}_{s+1} > \dim (\hat{\Phi}_{\frac{1}{16}\frac{1}{16}})_s \qquad \text{for} \quad s = 1,7,11,13,17,19,$$

  namely, the first representatives of the infinite series of the Coxeter exponents of E_8, modulo the Coxeter number $h = 30$ of this algebra.

$$s = 1,7,11,13,17,19,23,29 \quad (\text{mod } 30). \tag{16.8.15}$$

  This is not a coincidence, since the Ising model can also be defined in terms of the coset $\frac{(E_8)_1 \otimes (E_8)_1}{(E_8)_2}$.

16.9 Multiple Deformations of Conformal Field Theories

Till now we have analyzed the conformal models deformed by only one relevant operator. One may wonder if the analysis above can be generalized to deformations made of several fields. For instance, in the Ising model, there are two deformations – the thermal and the magnetization deformations – that separately give rise to two different integrable models. Are there, in this model, other lines[12] that are integrable in the plane $(h, T - T_c)$? The same question can be formulated for other models too, as for instance, for the tricritical Ising model where the two deformations $\Phi_{1,3}$ and $\Phi_{1,2}$ are individually integrable deformations.

Although presently there is no final answer to this question, an explicit computation to identify possible conserved currents with low values of the spin s gives a negative answer. The essential reason lies in the different null-vector structures that support the single deformations. This negative result leads us to be pessimistic about the possibility that there exists conserved currents of higher spin. To present this computation, let's first recall the derivation of a conservation law $\mathcal{C}_s \in \hat{T}_s$ when there is a single deformation, restricting attention to the unitary theory. Considering higher order perturbation terms, we have in general

$$\partial_{\bar{z}}\, \mathcal{C}_s(z,\bar{z}) = \lambda\, B^{(1)}_{lk}(z,\bar{z}) + \cdots \lambda^n\, B^{(n)}_{lk}(z,\bar{z}) + \ldots \tag{16.9.1}$$

Taking into account the dimensionality of the coupling constant, a dimensional analysis fixes the scaling dimensions of the operators $B^{(n)}_{lk}(z,\bar{z})$, given by

$$[s - n(1-\Delta), 1 - n(1-\Delta)].$$

Since $\Delta < 1$, there exists an integer n_c such that, for all $n > n_c$ the conformal weight of $B^{(n)}_{lk}(z,\bar{z})$ becomes negative. However, the absence of operators with negative conformal weights in the unitary minimal models forces the series (16.9.1) necessarily to stop (as a matter of fact, in most cases only the first term survives). If we now consider the deformations made by two operators with conformal weights Δ_1 and Δ_2 (and with corresponding couplings λ_1 and λ_2), the generalization of eqn (16.9.1) is

$$\partial_{\bar{z}}\, \mathcal{C}_s(z,\bar{z}) \;=\; \sum_{n,m=1} \lambda_1^n \lambda_2^m\, B^{(n,m)}_{lk}(z,\bar{z}). \tag{16.9.2}$$

The conformal weights of the quantities $B^{(n,m)}_{lk}$ are

$$[s - n(1-\Delta_1) - m(1-\Delta_2), 1 - n(1-\Delta_1) - m(1-\Delta_2)].$$

This series must truncate, for the same reason given above. Moreover, at least in the two explicit examples considered here, the Ising and the tricritical Ising model, the series splits into two independent expressions, one that is only a function of λ_1, with the other of λ_2. The reason is simple: in fact, the analytic conformal weight must

[12] If there exists an integrable point, this necessarily belongs to a renormalization group flow and therefore the Ising model would be integrable also along this line, see Fig. 16.9.

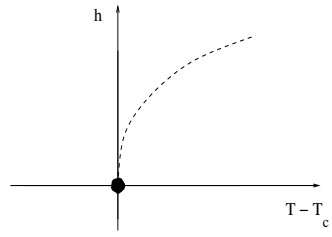


Fig. 16.9 *Space of the coupling constants of the Ising model near the critical point, here placed at the origin. The thermal and magnetic axes define two separate integrable models. Another potential integrable point in the plane will belong to a renormalization group flow (dashed line), so that the model would be integrable all along this curve.*

coincide with one of the conformal weights present in the Kac tables of the model. For the Ising model perturbed both by the energy and magnetization fields, we must have

$$1 - n\,\frac{1}{2} - m\,\frac{15}{16} = \Delta_r, \tag{16.9.3}$$

for some Δ_r of this model. However, possible values of Δ_r are only $\Delta_r = \{0, \frac{1}{2}, \frac{1}{16}\}$ and it is therefore impossible to have both n and m different from zero at the same time. The same situation occurs for the tricritical Ising model perturbed by the energy and vacancy densities, $\Phi_{\frac{1}{10}\frac{1}{10}}$ and $\Phi_{\frac{3}{5}\frac{3}{5}}$.

Therefore for these models, eqn (16.9.2) is expressed by the direct sum of the contribution of both terms. If there exists a conserved current, this should appear at the common level of the conserved currents of both deformations. Concerning the field $\Phi_{\frac{1}{2}\frac{1}{2}}$ of the Ising model and the field $\Phi_{\frac{3}{5}\frac{3}{5}}$ of the tricritical Ising model, both are $\Phi_{1,3}$ operators and therefore their associated conserved currents exist for

$$s = 1, 3, 5, 7, \ldots$$

For the magnetic deformation of the Ising model the spectrum of the conserved currents is given by the Coxeter exponents of E_8

$$s \;=\; 1, 7, 11, 13, 19, 23, 29 \pmod{30}.$$

For the tricritical Ising model, the spectrum of the conserved currents associated to the second operator $\Phi_{\frac{1}{10},\frac{1}{10}}$ coincides with the Coxeter exponents of E_7

$$s \;=\; 1, 5, 7, 9, 11, 13, 17 \pmod{18}.$$

Hence, in both models, the common values of the spins of their double deformation coincide with the Coxeter exponents of the corresponding E_n algebra.

In the following we explicitly show that there are no conserved currents of a double deformation of these models for the lowest values of s. As mentioned above, this result underlines the absence of the integrability of these statistical models under their multiple deformations.

16.9.1 The Tricritical Ising Model

We start the analysis from this model because there may exist a conserved current at level $s = 5$, whereas for the Ising model we shall consider at least $s = 7$.

The explicit expression of the conserved current $\mathcal{C}_6^{(1)}$ of the Φ_{13} deformation of the minimal model $\mathcal{M}_{p,p+1}$ coincides with the corresponding expression of the Sine–Gordon model

$$T_6^{(1)} \;=\; (T(T^2)) + \frac{9}{40}\,(T\partial^2\,T), \tag{16.9.4}$$

where we have substituted $c = 7/10$. Applying $\partial_{\bar{z}}$ to (16.9.4) and using the algebraic formalism of the operators D_n we have

$$\begin{aligned}\partial_{\bar{z}}\, C_6 = \lambda_1\,(1+\Delta_{13})[5\;L_{-5} - 4\;L_{-2}L_{-3}]\Phi_{13}\\ +L_{-1}[\cdots].\end{aligned}$$

The first term on the right-hand side is indeed zero for the Φ_{13} deformation, as a consequence of the null-vector equation of this operator at level 3. Hence $C_6^{(1)}$ is the conserved quantity under the $\Phi_{1,3}$ deformation. We need to check then if $T_6^{(1)}$ is still a conserved quantity if we perturb the model by means of the second operator $\Phi_{1,2} = \Phi_{\frac{1}{10}\,\frac{1}{10}}$. Repeating the previous steps, we get

$$\begin{aligned}\partial_{\bar{z}} C_6 = \lambda_2\,(1+\Delta_{12})\,[9\;L_{-5} - 6\;L_{-2}L_{-3}]\,\Phi_{12}\\ +L_{-1}[\cdots].\end{aligned} \tag{16.9.5}$$

In this case, however, the null-vector equation satisfied by the operator Φ_{12}

$$\left[L_{-2} - \frac{5}{42}\,L_{-1}^2\right]\Phi_{12} = 0$$

does not lead to the vanishing of the right-hand side of eqn (16.9.5)! As a matter of fact, the explicit expression of the conserved current under the Φ_{12} deformation is given by eqn (16.8.10)

$$T_6^{(2)} = (T(T^2)) + \frac{131}{10}\,(T\partial^2\,T), \tag{16.9.6}$$

which does not coincide with (16.9.4). Hence, the final conclusion of this computation is the absence of a conserved current of spin $s = 5$ for a multiple deformation of the tricritical Ising model. A similar analysis can be done also for the level $s = 7$, with a negative result as well.

16.9.2 The Ising Model

For the Ising model in an external magnetic field and at $T \neq T_c$ the first common value of the spin for both deformations is $s = 7$. The explicit expression of the conserved current under the $\Phi_{1,3}$ deformation is

$$C_8 \;=\; (T(T(T^2))) + \frac{c+8}{6}(T(T\partial_z^2 T)) + \frac{1}{180}(c^2+4c-101)(T\partial_z^4 T) \tag{16.9.7}$$

with $c = \frac{1}{2}$. Repeating the same steps of the computation shown for the example above, one can explicitly show that this current is not conserved under the $\Phi_{1,2}$ deformation associated to the magnetization operator.

Both examples clearly show the reason of the absence of common conserved currents, related to the different structures of the null-vectors of the different deformations. It would be a major discovery in statistical mechanics if in the future one could show the possibility of a conservation law for the multiple deformations of the Ising model.

References and Further Reading

The integrable features of classical systems can be studied by consulting:

P.G. Drazin, R.S. Johnson, *Solitons: An Introduction*, Cambridge University Press, Cambridge, 1989.

S. Novikov, S. Manakov, L. Pitaevskij, V. Zakharov, *Theory of Solitons*, Consultants Bureau, New York, 1984.

R.K. Dodd, J.C. Eilbeck, J.D. Gibbons, H.C. Morris, *Solitons and Nonlinear Wave Equations*, Academic Press, New York, 1982.

A. Scott, F. Chu, D. McLaughlin, *The Soliton: A New Concept in Applied Science*, Proc. of the IEEE, 61 (1973), 1443.

An incisive text to understand the relationship between classical and quantum systems is:

R. Rajaraman, *Solitons and Instantons*, North Holland, Amsterdam, 1982.

The integrable deformations of conformal theories have been analyzed in the fundamental paper by A.B. Zamolodchikov:

A.B. Zamolodchikov, *Integrable field theory from conformal field theory*, Adv. Stud. Pure Math., 19 (1989), 641.

The relation between the integrable $\Phi_{1,3}$ deformation of the conformal minimal models and the Sine–Gordon theory has been discussed in:

T. Eguchi, S.K. Yang, *Deformations of conformal field theories and soliton equations*, Phys. Lett. B 235 (1990), 282.

A review paper on Toda field theories and integrable theories away from criticality is:

G. Mussardo, *Off-critical statistical models: Factorized scattering theories and the bootstrap program*, Phys. Rep. 218 (1992), 215.

The quantum integrability of the Sine/Sinh–Gordon and Bullogh–Dodd models using Feynman diagrams has been discussed in:

P. Dorey, *Exact S-matrices*, Proceedings of the Eotvos Summer School in Physics: Conformal Field Theories and Integrable Model, Budapest 1996, hep-th/9810026

It is important to point out that for the integrable models that admit a lagrangian formulation the semiclassical quantization provides a useful set of information on their dynamics. The reader is urged to consult the papers:

R. Dashen, B. Hasslacher, A. Neveu, *Particle spectrum in model field theories from semiclassical functional integral techniques*, Phys. Rev. D 11 (1975), 3424.

J. Goldstone, R. Jackiw, *Quantization of nonlinear waves*, Phys. Rev. D 11 (1975), 1486.

Problems

1. Bäcklund transformations

a Write down the Bäcklund transformations for the Sine–Gordon model.

b Taking as initial solution of the equation of motion $\phi = 0$, determine the new solution $\hat{\phi}(z, \bar{z})$ and show that it coincides with the solitonic solution of the model.

c Iterate the procedure to determine the other classical solutions of the Sine–Gordon model.

2. Scattering processes of the solitons

Analyze the solution (16.3.15) with topological charge $Q_t = 2$ of the Sine–Gordon model in the limits $t \to \pm\infty$. Determine the time delay Δ_{ss}. Based on the positive sign of this quantity and the negative sign of the analogous quantity $\Delta_{s\bar{s}}$ for the scattering of the soliton and antisoliton, argue about the nature of the interactions between soliton–soliton and soliton–antisoliton.

3. Lax pair

Consider the pair of first-order differential operators (called a Lax pair)

$$L(x,t \mid \theta) = \frac{d}{dx} + i\left(\frac{\beta}{4}\partial_t\phi\,\sigma_3 + m\sinh\theta, \cos\frac{\beta\phi}{2}\,\sigma_1 + m\cosh\theta\,\sin\frac{\beta\phi}{2}\,\sigma_2\right)$$

$$M(x,t \mid \theta) = \frac{d}{dt} + i\left(\frac{\beta}{4}\partial_x\phi\,\sigma_3 + m\cosh\theta\,\cos\frac{\beta\phi}{2}\,\sigma_1 + m\sinh\theta\,\sin\frac{\beta\phi}{2}\,\sigma_2\right)$$

where σ_i are the Pauli matrices and θ the rapidity variable. If $[L, M] = 0$:

a Show that the field ϕ satisfies the equation of motion of the Sine–Gordon model

$$\Box\phi + \frac{m^2}{\beta}\sin\beta\phi \;=\; 0.$$

b Take a rectangular domain $-L \le x \le L; 0 \le t \le T$ and assume a periodic boundary condition $\phi(-L) = \phi(L)$. With the notation

$$\mathcal{T}_L(\theta, t) = \overrightarrow{\exp}\left(\int_{-L}^{L} U(x, t \mid \theta)\, dx\right), \qquad \mathcal{S}_L(\theta) = \overrightarrow{\exp}\left(\int_0^T V(x, t \mid \theta)\, dt\right)$$

for the ordered integrals, show that

$$\mathcal{T}_L(\theta, T) = \mathcal{S}_L^{-1}(\theta)\, \mathcal{T}_L(\theta, 0)\, \mathcal{S}_L(\theta)$$

so that $\operatorname{Tr} \mathcal{T}_L(\theta, t)$ is independent of t; conclude that, θ being arbitrary, there is an infinite number of conserved quantities.

4. Derrick theorem

The aim of this exercise is to show that the static solitonic solution of finite energy can only exist for $1 + 1$ dimensional theories. Consider, in $(d + 1)$-dimensional Minkowski space, the lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - U(\phi),$$

where $U(\phi)$ is a non-negative function that vanishes at the vacua of the theory. The static energy E can be written as $E = W_1 + W_2$, where

$$W_1 = \frac{1}{2}\int d^d x\, (\nabla \phi)^2, \qquad W_2 = \int d^d x\, U(\phi).$$

Let $\phi(x)$ be a static solution of the equation of motion of the theory.

a Determine the variation of W_1 and W_2 under the transformation $\phi(x) \to \phi(\lambda x)$.

b Using the condition that $\phi(x)$ is a solution of the equation of motion, show that the energy $E[\lambda]$ is stationary for $\lambda = 1$.

c Since $W_1 \ge 0$ and $W_2 \ge 0$, show that one can have non-vanishing solutions only for $d \le 2$.

5. Liouville theory and minimal models

a In the quantization scheme of the Sine–Gordon model in terms of the Liouville theory, determine the quantized values of the coupling constant g that reproduce the central charges of the minimal models. Prove that the conformal weight of the vertex operator that perturbes the Liouville action is equal to $\Delta_{1,3}$.

b Repeat the same exercise for the two Liouville theories, with complex exponentials, associated to the Bullogh–Dodd model. Show that the perturbations correspond to the operators $\Phi_{1,2}$ and $\Phi_{2,1}$ of the minimal models respectively.

6. Conserved currents

Using the algebra of the operators D_n and the null-vector equation at the level 2 satisfied by $\Phi_{1,2}$ and $\Phi_{2,1}$, find the linear combination T_6 of the basis vectors $T_6^{(1)}$ and $T_6^{(2)}$ that satisfies

$$\partial_{\bar{z}} T_6 = \partial_z\, \Theta_4.$$

Determine the density Θ_4.

17
S-Matrix Theory

All men are equal, just that some are more equal than others.

George Orwell

In this chapter we present the S-matrix theory of two-dimensional integrable models. This leads, in particular, to the exact spectrum of the massive excitations away from the critical point. From a mathematical point of view, the two-dimensional nature of the systems and their integrability are the crucial features that lead to important simplifications of the formalism and its successful application. It is worth mentioning that, initially developed to overcome the obstacles encountered by quantum field theory in dealing with the strong interactions of the hadronic particles[1] (such as protons, neutrons, or pions), the S-matrix has achieved its most beautiful intellectual triumph in the analysis of the two-dimensional statistical models away from criticality, particularly when they are described by integrable theories. These significant developments have been pioneered by A.B. Zamolodochikov.

The key point of this formalism is the self-consistent dynamical method for computing the exact expressions of all scattering amplitudes and the mass of the particles. This is the so-called *boostrap approach*,[2] where all particles are democratically on the same footing: there is no distinction between the particles of the asymptotic states and the bound states, and any massive excitation of the theory can equivalently be regarded as an asymptotic state or a bound state of a pair of particles of the same theory. From this point of view, all particles are composite states and no one is more elementary than another. The only difference between them (apart some internal quantum number) consists of the value of their masses, which may provide a hint about the number of interactions they are involved with. Quoting Orwell, we can then say that the lightest particle of the theory is the one more equal than the others.

In this chapter we firstly address the general principles of S-matrix theory and secondly we discuss their application to the two-dimensional cases. In the next chapter we present some significant examples of this remarkable formalism, in particular the exact solution of the Ising model in an external magnetic field at $T = T_c$.

[1] We refer the reader interested in these developments to the appendix of this chapter.

[2] In addition to the conformal bootstrap, this is another example of a theory whose solution is based on its own mathematical and physical self-consistency.

17.1 Analytic Scattering Theory

In a relativistic context, S-matrix theory is a generalization of the scattering process theory of quantum mechanics, briefly discussed in Appendix B of this chapter. Its aim is to derive general conditions on the transition amplitudes of the scattering processes involving the multiparticle asymptotic states, with the aim of arriving at their computation without relying on an underlying lagrangian formalism.

17.1.1 General Properties

The main properties at the root of S-matrix theory are the following:

1. the short range of the interactions;
2. the superposition principle of quantum mechanics;
3. the conservation of probability;
4. the invariance under the Lorentz transformations of special relativity;
5. the causality principle;
6. the analyticity principle.

Let's discuss in more detail each point and work out their consequences for a generic scattering theory in $d \geq 2$. The two-dimensional case will be analyzed separately later on.

To adopt the S-matrix formalism to describe the scattering processes it is necessary to assume that the interactions are short range, so that the initial and final states, in which the particles are well separated one from another, consist of free particle states. These multiparticle states can be specified assigning the momenta[3] and other possible quantum numbers. For simplicity we focus our attention only on the scattering processes of the scalar particles. Since the scattering processes involve the physical particle states instead of virtual ones, the components of their momenta satisfy the d-dimensional *on-shell* condition

$$p_\mu \, p^\mu \, = \, m^2,$$

where m is the mass of the particle. This equation gives rise to the dispersion relation

$$E^2 - | \, \vec{p} \, |^2 = \, m^2 \tag{17.1.1}$$

that links together the energy $E = p_0$ and the space component $\vec{p}$ of the momentum. The spectrum of the eigenvalues of the spatial momentum is obviously a continuum but, to simplify the discussion below, it is useful to use momentarily the compact notation $| \, n \rangle$ to denote the states of the system. They are made of free particles, and they form a basis of the Hilbert space that satisfy the orthogonal and completeness relations

$$\langle m \mid n \rangle \, = \, \delta_{nm}, \quad \sum_n \, | \, n \rangle \langle n \, | = \, 1.$$

We will specialize later the Lorentz invariant normalization condition of the states.

[3] In the following by momentum we mean the d-dimensional relativistic momentum of the particles, alias the set of all its components $(p^0, \vec{p})$. However, using the on-shell condition (17.1.1), it is obvious that the multiparticle states are identified just by the space components of their momenta.

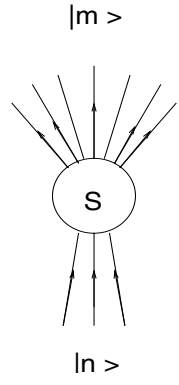


Fig. 17.1 *Quantum transition from an initial n-particle state to a final m-particle state.*

At $t = -\infty$, let $\mid i\rangle$ be the initial state of the system, given by a certain number of free particles. At $t = +\infty$, i.e. after they have interacted, the final state $\mid \tilde{f}\rangle$ of the system also consists of free particles, although not in the same number or with the same momenta as the initial state, as shown in Fig. 17.1. For the superposition principle of quantum mechanics, the final state can be written as $\mid \tilde{f}\rangle = S \mid i\rangle$, where S is a linear operator.[4] Hence, the probability that a measure on the final state produces as a result the state $\mid f\rangle$ is expressed by the modulus squared of the matrix element

$$S_{fi} = \langle f \mid S \mid i\rangle. \tag{17.1.2}$$

Consider now an initial normalizable state $\mid \psi\rangle$, given by a linear superposition of the basis vectors

$$\mid \psi\rangle = \sum_n a_n \mid n\rangle\,, \quad \sum_n \mid a_n \mid^2 = 1.$$

The total probability that this state evolves as a final state in any basis vectors is obviously equal to 1 and we have therefore

$$\begin{aligned} 1 &= \sum_m \mid \langle m \mid S \mid \psi\rangle \mid^2 = \sum_m \langle \psi \mid S^\dagger \mid m\rangle\langle m \mid S \mid \psi\rangle \\ &= \langle \psi \mid S^\dagger S \mid \psi\rangle = \sum_{n,m} a_n^\star a_m \langle n \mid S^\dagger S \mid m\rangle. \end{aligned}$$

Since this identity should hold for arbitrary values of the coefficients a_n, necessarily

$$\langle n \mid S^\dagger S \mid m\rangle = \delta_{nm},$$

[4] S is the time evolution operator from $t = -\infty$ to $t = +\infty$. If the system admits a quantum field theory formulation, it is expressed as $S = \mathcal{T} \exp[-i \int_{-\infty}^{+\infty} d^d x \mathcal{H}_i(x)]$, where $\mathcal{H}_I$ is the hamiltonian density and $\mathcal{T}$ denotes the time-ordering of the expressions obtained by the series expansion of the exponential term.

or, in operator form,

$$S^\dagger S \,=\, 1. \tag{17.1.3}$$

Similarly, imposing equal to 1 the total probability that an arbitrary final state comes from some initial state is, one obtains the condition

$$S\, S^\dagger \,=\, 1. \tag{17.1.4}$$

In conclusion, probability conservation requires S to be a unitary operator.

Let's now analyze the Lorentz invariance of the scattering theory. Let L be an arbitrary proper Lorentz transformation and $L \mid m\rangle = \mid m'\rangle$. The relativistic invariance of the theory, which ensures the independence of the physical observables from the reference frames, is expressed by the identity

$$\mid \langle m' \mid S \mid n' \rangle \mid^2 = \mid \langle m \mid S \mid n \rangle \mid^2 .$$

This relation cannot fix the relative phase between the two matrix elements but, given the intrinsic arbitrariness of the overall phase of the S-matrix, we can impose the more stringent condition

$$\langle m' \mid S \mid n' \rangle \,=\, \langle m \mid S \mid n \rangle. \tag{17.1.5}$$

As we will see later, this equation implies that the S-matrix, once we factorize a delta-function for the conservation of the total momenta, depends on the momenta of the particles only through their Lorentz invariant combinations of their scalar products.

Without interactions, the state of a system does not change and in this case the S-matrix is simply the identity operator. It is a common procedure to separate the free time evolution, given by the identity operator, and write the S-matrix as

$$S_{fi} \,=\, \delta_{fi} + i(2\pi)^d \delta^d(P_f - P_i)\, T_{fi}. \tag{17.1.6}$$

The matrix elements T_{fi} define the *scattering amplitudes*. In the second term of this expression we have also explicitly written the factor $\delta^d(P_f - P_i)$ that expresses the conservation law of the total momentum, where P_i and P_f are the sum of the momenta of the initial and final particles, respectively. For the non-diagonal matrix elements $i \to f$ the matrix elements of the identity operator vanish and we have

$$S_{fi} \,=\, i(2\pi)^d \delta^d(P_f - P_i)\, T_{fi}. \tag{17.1.7}$$

The relative probability is obtained by the modulus squared of this amplitude. In computing such a modulus squared there is however a problem, whose origin is the interpretation to assign to the square of the delta function. This problem can be solved by using initially the following representation of $\delta(x)$

$$\delta^d(P_f - P_i) \,=\, \frac{1}{(2\pi)^d} \int e^{i(P_f - P_i)x}\, d^d x.$$

Computing now another integral of this kind at $P_f = P_i$ (just for the presence of the delta-function) and taking the integral over a finite time interval t and on a

$(d-1)$-dimensional volume V, sufficiently large but finite, the result is $Vt/(2\pi)^d$. For the modulus squared of the matrix element we have then

$$\mid S_{fi} \mid^2 = (2\pi)^d \, \delta^d(P_f - P_i) \mid T_{fi} \mid^2 \, Vt.$$

Dividing now for the factor Vt, we get the transition probability per unit volume and unit time

$$P_{i\to f} = (2\pi)^d \, \delta^d(P_f - P_i) \mid T_{fi} \mid^2 .$$

The most important cases, both from a theoretical and experimental point of view, are those in which the initial state is made either of one particle or two particles. The first case concerns the *decay processes*, i.e. when a heavy particle decays in a set of lightest ones, whereas the second case is relative to the *scattering* of two particles, which can result in an elastic diffusion or in a production process.

It is now useful to specify more precisely the normalization of the states. The more convenient choice is related to the covariant normalization of the one-particle state

$$\langle p' \mid p \rangle = 2E \, (2\pi)^{d-1} \, \delta^{d-1}(\vec{p}' - \vec{p}). \tag{17.1.8}$$

This is a Lorentz invariant normalization and it is equivalent to integrating over the mass-shell state of a particle as

$$\int \frac{d^{d-1}p}{(2\pi)^{d-1}2E} \mid p\rangle \langle p \mid p'\rangle = \int \frac{d^d p}{(2\pi)^{d-1}} \, \delta(p^2 - m^2) \mid p\rangle \langle p \mid p'\rangle = \mid p'\rangle,$$

with $E > 0$. Hence, the density of states associated to a *on–shell* particle with momentum in the interval $(p, p + dp)$ is given by

$$\frac{d^{d-1}p}{(2\pi)^{d-1}2E}.$$

Decay process. Taking into account the proper normalization of the states, the probability of a decay of a particle of energy E into an n-particle state is expressed by

$$d\Gamma = (2\pi)^d \, \delta^d(P - p_1 - \cdots - p_n) \mid T_{fi} \mid^2 \frac{1}{2E} \prod_{i=1}^{n} \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}. \tag{17.1.9}$$

Scattering process $2 \to n$**.** The probability that a collision of two particles of momenta $p_1 = (E_1, \vec{p}_1)$ and $p_2 = (E_2, \vec{p}_2)$ produces their transformation in an arbitary number of other particles with momenta p_i is given by

$$dP = (2\pi)^d \, \delta^d(P - p_1 - \cdots - p_n) \mid T_{fi} \mid^2 \frac{1}{4E_1 E_2} \prod_{i=1}^{n} \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}. \tag{17.1.10}$$

In the last case, rather than the probability, it is often more interesting to compute the Lorentz invariant *cross-section* $d\sigma$ of the collision. This is obtained by dividing the probability dP by

$$j \,=\, \frac{I}{E_1 E_2},$$

where I is the scalar quantity

$$I \,=\, \sqrt{(p_1 \cdot p_2)^2 - (m_1 m_2)^2}.$$

It is easy to see that j is the flux density of the colliding particles. In fact, in the reference frame of the center of mass of the system ($\vec{p}_1 = -\vec{p}_2 = \vec{p}$), one has $I = \mid \vec{p} \mid (E_1 + E_2)$ and then

$$j \,= \mid \vec{p} \mid \left(\frac{1}{E_1} + \frac{1}{E_2} \right) \,=\, v_1 + v_2,$$

where v_1 and v_2 are the velocities of the two colliding particles. Hence the cross-section is the transition probability per unit of the flux of the scattering particles.

Note that in the probability of both decay or scattering processes there is the quantity

$$d\Phi_n \,=\, \frac{d^{d-1}p_1}{(2\pi)^{d-1} 2E_1} \cdots \frac{d^{d-1}p_1}{(2\pi)^{d-1} 2E_1} \, (2\pi)^d \, \delta^d(P - p_1 - p_2 - \cdots - p_n). \qquad (17.1.11)$$

This is the differential n-particle *phase space*. It expresses the density of states for an n-particle system with total momentum P. This quantity also enters the spectral density of the correlation functions, which will be discussed in Chapter 20. Given its relevance in many aspects of the theory, its detailed study is carried on in Appendix 17C.

Let's now investigate the consequences of the unitarity condition of the S-matrix. Substituting eqn (17.1.6), in (17.1.4) we get

$$T_{fi} - T^{\star}_{if} \,=\, i \, (2\pi)^d \sum_n \delta^d(P_f - P_i) \, T_{fn} \, T^{\star}_{in}, \qquad (17.1.12)$$

where the sum over the index n here denotes, in compact notation, both a sum and an integral over all intermediate states allowed by the conservation of the total momentum of the process. Note that the left-hand side of this equation is linear with respect to the matrix elements of T, whereas the right-hand side is quadratic. If the theory under investigation has a coupling constant g that can be regarded as a perturbative parameter, the first consequence of eqn (17.1.12) is the hermiticity of the matrix T at the first perturbative order

$$T_{fi} \,\simeq\, T^{\star}_{if}. \qquad (17.1.13)$$

In fact, the left-hand side of (17.1.12) is of first order in g, whereas the right-hand side is of second order.

Optical theorem. Another important consequence of eqn (17.1.12) is the *optical theorem* relative to the scattering process of two particles. To prove it, let's initially sandwich eqn (17.1.12) with the states $\mid p_1, p_2\rangle$ and $\mid p_3, p_4\rangle$

$$2\,\mathrm{Im}\,\langle p_3, p_4 \mid T \mid p_1, p_2\rangle \;=\; (2\pi)^d \sum_n \delta^d(P_f - P_i)\,\langle p_3, p_4 \mid T \mid n\rangle\langle p_1, p_2 \mid T^\star \mid n\rangle. \tag{17.1.14}$$

If the scattering process is purely elastic, the final state coincides with the initial state and in this case we have

$$2\,\mathrm{Im}\,T_{ii} \;=\; (2\pi)^d \sum_n \delta^d(P_f - P_i) \;\mid T_{in} \mid^2 .$$

Note that the right-hand side of this expression differs only for a multiplicative factor from the total cross-section σ_t of all possible scattering processes obtained by a given initial state i

$$\sigma_t \;=\; \left(\frac{(2\pi)^d}{j}\right) \sum_n \mid T_{in} \mid^2 \;\delta^d(P_i - P_n).$$

Therefore we have the *optical theorem*, stated by the relation

$$\sigma_t \;=\; \frac{2}{j}\,\mathrm{Im}\,T_{ii}.$$

This theorem allows us to compute the total cross-section of the theory (which also includes all the inelastic processes) in terms of the imaginary part of the purely elastic scattering amplitude of two particles.

Finally, let's comment on the final principles on which S-matrix theory is based, namely the causality and the analyticity principles. One expects that these two aspects should be deeply related to each other, on the basis of the well-known example of the dispersion relations satisfied by the Green functions of an ordinary quantum system (see Problem 1). However, in the context of relativistic quantum mechanics, it is in general a difficult problem to pin down the precise analytic structure of the S-matrix in terms of the causality principle. Quite often, in fact, the analytic properties of the S-matrix elements are conjectured on the basis of those derived in the non-relativistic scattering processes or encountered in the perturbative diagrams of the associated quantum field theory, when this is known. In short, the basic assumption on which we rely is encoded in the following statement: *the transition amplitudes coincide with the real boundary values of analytic functions of several complex variables having a minimum number of singularities dictated by specific physical processes*. The study of the two-particle scattering process will help us in clarifying some important aspects of this topic.

17.1.2 Two-body Scattering Process

Let's consider in more detail the diffusive scattering process of two initial scalar particles (with momenta p_1 and p_2) going into two scalar particles (with momenta p_3 and p_4), as shown in Fig. 17.2,

$$A_1 + A_2 \;\rightarrow A_3 + A_4. \tag{17.1.15}$$

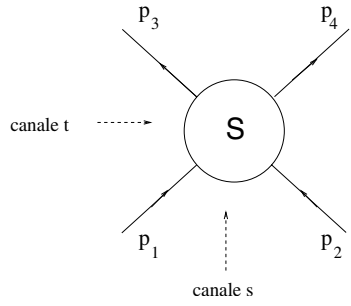


Fig. 17.2 *Two-particle scattering process.*

Once we factorize the delta-function of the conservation of the total momentum

$$\langle p_3, p_4 \mid T \mid p_1, p_2 \rangle \;=\; i(2\pi)^d\, \delta^d(p_1 + p_2 - p_3 - p_4)\; \mathcal{T}, \tag{17.1.16}$$

the remaining quantity $\mathcal{T}$ is an analytic function of the relativistic invariants of this process. They can be expressed in terms of the Mandelstam variables s, t, and u, given by

$$s = (p_1 + p_2)^2, \quad t = (p_1 - p_3)^2, \quad u = (p_1 - p_4)^2. \tag{17.1.17}$$

These quantities are not all independent, since from the conservation law

$$p_1 + p_2 \;=\; p_3 + p_4, \quad p_i^2 \;=\; m_i^2 \;\; (i = 1, 2, 3, 4)$$

one has

$$s + t + u = \sum_{i=1}^{4} m_i^2. \tag{17.1.18}$$

It is easy to understand the meaning of s, going in the reference frame of the center of mass of the process (17.1.15), defined by $\vec{p}_1 + \vec{p}_2 = 0$. In this frame $s = E^2$, where $E = E_1 + E_2$ is the total energy in the center-of-mass frame. The variable t is instead the square of the energy in the center of mass of the crossed channel

$$A_1 \;+\; \overline{A}_3 \;\rightarrow\; \overline{A}_2 \;+\; A_4, \tag{17.1.19}$$

and the same is true for the variable u, with respect to the crossed channel

$$A_1 \;+\; \overline{A}_4 \;\rightarrow\; \overline{A}_2 \;+\; A_3. \tag{17.1.20}$$

In the equations above, $\overline{A}_i$ denotes the antiparticle: going in a cross-channel, one has to reverse the arrow of the out-going particle that becomes then the antiparticle.

Production thresholds and branch cuts. In view of eqn (17.1.18), the amplitude $\mathcal{T}$ is a function of only two of the Mandelstam variables, say s and t. Let's study its analytic structure as a function of s at fixed t, assuming for simplicity that each of the four particles involved in this scattering process has the same mass m. The physical

values of s are given by $s \geq s_2 = (2m)^2$: this is the set of values for which there exists the physical state of the two asymptotic particles. In the interval

$$(2m)^2 \leq s \leq (3m)^2,$$

corresponding to values of the total energy in the center of mass less than the threshold of inelastic production, the two-particle states are the only intermediate states that can appear in the right-hand side of eqn (17.1.14). Therefore

$$2\,\mathrm{Im}\,\langle p_3, p_4 \mid T \mid p_1, p_2\rangle = (2\pi)^d \int \frac{d^{d-1}k_1}{(2\pi)^{d-1}2E_1}\frac{d^{d-1}k_2}{(2\pi)^{d-1}2E_2}\delta^d(p_1+p_2-k_1-k_2) \times \langle p_3, p_4 \mid T \mid k_1, k_2\rangle\langle p_1, p_2 \mid T^\star \mid k_1, k_2\rangle. \tag{17.1.21}$$

But once the threshold value is overcome, in the next interval

$$(3m)^2 < s < (4m)^2,$$

it is necessary to add other terms in the right-hand side of the equation above: these terms are those relative to the intermediate states made of three particles, compatible with the conservation law of energy. In the same way, there are other additional terms due to the N-particle intermediate states each time that s overcomes their threshold of production, $s_N = (N\,m)^2$.

The discontinuity in the imaginary part of the amplitude of the elastic scattering by varying s implies that it has certain singularities in correspondence with the threshold values of the inelastic processes. They are branch points of the amplitude $\mathcal{T}$, as it is easy to show using the Feynman diagrams of the perturbative series. In the complex plane of the variable S it is thus convenient to draw a series of cuts starting from the various thresholds to infinity, all along the real axis, as in Fig. 17.3. In this way the scattering amplitude becomes a one-value function on the corresponding Riemann surface. The *physical sheet* is obtained without crossing any cuts of Fig. 17.3 whereas the other sheets, called *non-physical sheets*, are defined specifying the crossing of one or more cuts of the amplitude $\mathcal{T}(s,t)$.

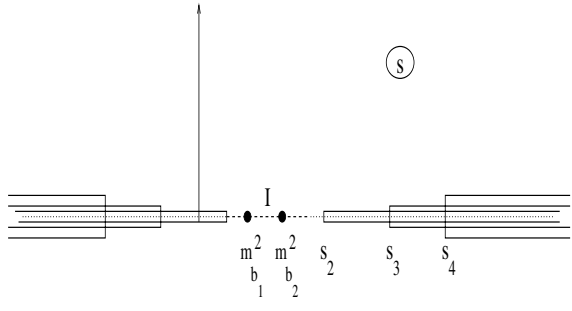


Fig. 17.3 *Analytic structure of the elastic scattering amplitude of two particles. On the right-hand side there are the branch cuts relative to the threshold values in the s-channel, while on the left-hand side there are the branch cuts relative to the threshold values of the t-channel. The circle represents the poles of the amplitude, relative to the bound states.*

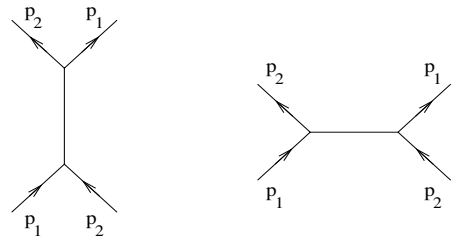


Fig. 17.4 *Feynman diagram relative to (a) the s-channel amplitude and (b) the t-channel amplitude, for the scattering process of two particles in a ϕ^3 theory.*

Bound states and poles. The lowest threshold, at $s_2 = 4m^2$, is associated to the physical state of two particles. As an analytic function of s, $\mathcal{T}(s,t)$ can also be evaluated for non-physical values of s, such as those less than s_2. The possibility to create an arbitrary number of particles starting from the two-particle state can also be considered for $s < s_2$. However, in this case, these are only the one-particle states, with mass $m_{b_i} < 2m$. These are obviously virtual processes since they are precluded by the conservation of energy that holds for the physical process. However they determine the *bound states* of the asymptotic particles and, as for the non-relativistic scattering amplitudes (see Appendix 17B), correspond to simple poles in the amplitude $\mathcal{T}(s,t)$. This analytic structure is confirmed by the perturbative theory based on the Feynman graphs. Consider, for instance, a theory in which there is a ϕ^3 interaction: in the scattering process of two particles there are the graphs shown in Fig. 17.4. The first diagram, apart from some constants, is given by

$$(a) \longrightarrow \frac{1}{(p_1+p_2)^2 - m^2 + i\epsilon}, \tag{17.1.22}$$

and gives rise to a pole in the s-channel, while the second diagram

$$(b) \longrightarrow \frac{1}{(p_1-p_2)^2 - m^2 + i\epsilon}, \tag{17.1.23}$$

gives rise to a pole in the t-channel.

Physical regions and crossing invariance. The region in which $\mathcal{T}(s,t)$ coincides with the amplitude relative to the physical scattering process (17.1.15) is that in which there are positive values of the energies of all particles and real values of their momenta. For particles of equal mass, this region is identified by the conditions[5]

$$s \geq 4m^2, \quad t \leq 0, \quad u \leq 0, \tag{17.1.24}$$

[5]If the masses are different, the conditions are slightly more complicated.

as can be seen by expressing s, t and u in terms of the momentum $\vec{q}$ and the scattering angle θ in the center-of-mass frame:

$$\begin{aligned} s &= 4(m^2 + q^2), \\ t &= -2q^2(1 - \cos\theta), \\ t &= -2q^2(1 + \cos\theta). \end{aligned}$$

Since $\mathcal{T}(s,t)$ is an analytic function of both variables, it can be analytically continued from the original domain (17.1.24) to the regions

$$t \geq 4m^2, \quad s \leq 0, \quad u \leq 0, \tag{17.1.25}$$

and

$$u \geq 4m^2, \quad s \leq 0, \quad t \leq 0. \tag{17.1.26}$$

The first region corresponds to the physical domain relative to the channel (17.1.19) while the second region to the physical domain of the channel (17.1.20). This implies that the *same analytic function* can be used to describe the three different physical processes given in (17.1.15) (the *s-channel*), in (17.1.19) (the *t-channel*), and in (17.1.20) (the *u-channel*). This fundamental property of the ampitude $\mathcal{T}(s,t)$ expresses the *crossing invariance* of the scattering processes.

***t*-channel**. As we have identified the threshold singularities of $\mathcal{T}(s,t)$ by varying s at fixed t, we can similarly identify the singularities of this amplitude by varying t and u, using the crossing invariance. In the t-channel the threshold singularities are placed at

$$t = 4m^2, \quad 9m^2, \quad 16m^2, \ldots \tag{17.1.27}$$

and analogously in the u-channel

$$u = 4m^2, \quad 9m^2, \quad 16m^2, \ldots \tag{17.1.28}$$

From the relation (17.1.18), fixing the value u_0 of the variable u, the branch points (17.1.27) then appear in the complex plane of the variable s at the points

$$s = -u_0, \quad -u_0 - 5m^2, \quad -u_0 - 12m^2, \ldots \tag{17.1.29}$$

whereas the pole at $t = m_{b_i}^2$ appears in the position

$$s = -u_0 + 3m_{b_i}^2. \tag{17.1.30}$$

The analytic structure (at $u = u_0$, fixed) is shown in Fig. 17.3.

Physical amplitude. Since we are in the presence of branch cuts along the real axis of the s-plane, it is necessary to establish the limit of the function $\mathcal{T}(s,t)$ associated to the physical amplitude in the s-channel. The physical region of this process is identified by $u_0 < 0$ and by the real values of s, with $s > 4m^2 - u_0$. Perturbative theory shows that

the physical amplitude is recovered by taking the limit from the upper half complex plane on the first cut of the function $\mathcal{T}(s,t)$, namely

$$\mathcal{T}_{phys} \;=\; \lim_{\epsilon\to 0^+} \mathcal{T}(s+i\epsilon, u_0). \tag{17.1.31}$$

Note that this result is equivalent to the Feynman prescription $i\epsilon$ in the propagators of the particles

$$\frac{1}{p^2-m^2+i\epsilon}.$$

In fact, adopting this prescription, any integral on the momenta of the intermediate particles can be computed with real external momenta, i.e. corresponding to a real value of the variable s. Moreover, eqn (17.1.31), together with hermitian analyticity, implies that the amplitude $\mathcal{T}$ is a real function in the real interval I between the two branch cuts (Fig. 17.3), as can also be proved directly from the Schwartz reflection principle in complex analysis.

17.2 General Properties of Purely Elastic Scattering Matrices

Let's now specialize the general conditions discussed in the previous section to the case of $(1+1)$ scattering theories when there is an infinite number of conserved charges $\mathcal{Q}_s$ in involution. These two circumstances give rise to a drastic simplification of the analytic structure of the S-matrix and will lead to an exact expression of the scattering amplitudes in many interesting cases.

17.2.1 Rapidity Variable and Asymptotic States

The momenta of the particles involved in scattering processes are *on-shell*. In $(1+1)$ dimensions there exists an efficient parameterization of the dispersion relation $E^2 - p^2 = m^2$ in terms of the rapidity variable θ. For a particle of mass m_i we have in fact

$$p_i^{(0)} \;=\; m_i \cosh\theta_i, \qquad p_i^{(1)} \;=\; m_i \sinh\theta_i. \tag{17.2.1}$$

Note that the Lorentz transformations can be regarded as a rotation of a hyperbolic angle α and therefore implemented as $\theta \to \theta + \alpha$. Furthermore, both components of the momentum can be changed by sign with the transformation $\theta_i \to i\pi - \theta_i$. In this way, the momentum of the original particle becomes that of its own antiparticle. Later it will also be useful to consider the light-cone components, defined by

$$\begin{aligned} p \;&=\; p^{(0)} + p^{(1)} \;=\; m_i\, e^{\theta}, \\ \overline{p} \;&=\; p^{(0)} - p^{(1)} \;=\; m_i\, e^{-\theta}. \end{aligned} \tag{17.2.2}$$

They satisfy $p\,\overline{p} = m_i^2$.

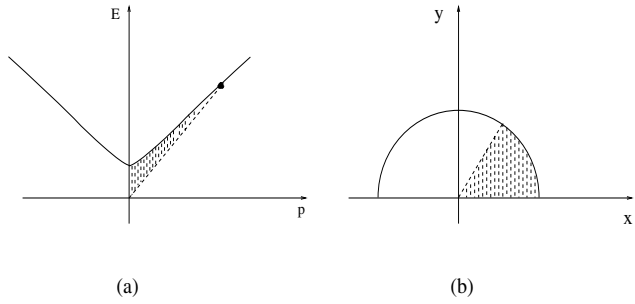


Fig. 17.5 *Geometrical interpretation of the rapidity variable* θ.

The rapidity variable has an interesting geometric intepretation, due to the Italian mathematician Riccati. In a plane with axes given by E and p, the dispersion relation $E^2 = p^2 + m^2$ represents a hyperbola, as shown in Fig. 17.5a. The rapidity is proportional to the area $\mathcal{A}$ that is encompassed between the hyperbola and the straight line that joins the origin to the point of the hyperbola identified by the variable θ. The relation is $\mathcal{A} = m^2\theta/2$. An analogous result is obtained for the angle α that parameterizes the points of a circle $x^2 + y^2 = m^2$, shown in Fig. 17.5b. Imposing $x = m\cos\alpha$ and $y = m\sin\alpha$, the area $\mathcal{A}$ between the horizontal axis and the segment that identifies the point on the circles is in fact $\mathcal{A} = m^2\alpha/2$. The two geometrical situations are related by the analytic continuation $\alpha \to i\theta$.

The n-particle states of this theory can be written as

$$| A_{a_1}(\theta_1)\, A_{a_2}(\theta_2) \ldots A_{a_n}(\theta_n)\rangle, \tag{17.2.3}$$

where by the symbol $A_{a_i}(\theta_i)$ we denote the particle of type a_i that is moving with rapidity θ_i. By means of a linear superposition of these states, we can construct wave packets that are localized both in momentum and coordinate space. In this way, we can imagine assigning a well-defined position to the particles above. In the massive theories, the interactions are short range and consequently a state like (17.2.3) represents a collection of free particles except in the time instants in which the wavepackets overlap. Let's discuss in more detail how to represent the initial and final states.

An initial asymptotic state is given by a set of free particles at $t \to -\infty$. Since in the (1+1) dimensional theories the actual motion takes place on a line, this means that the fastest particle must be on the farthest left-hand side of all the others, while the slowest must be on the right-hand side of all the others, with the remaining particles are ordered according to the value of their rapidities between those two. To express this situation in a formal way, it is appropriate to consider the symbols $A_{a_i}(\theta_i)$ as non-commuting variables, whose order is associated to the space ordering of the particles that they represent. In this way, an initial asymptotic state can be written as

$$| A_{a_1}(\theta_1)\, A_{a_2}(\theta_2) \ldots A_{a_n}(\theta_n)\rangle, \tag{17.2.4}$$

where the rapidities are ordered in a *decreasing* way

$$\theta_1 \geq \theta_2 \geq \theta_3 \cdots \geq \theta_n. \tag{17.2.5}$$

Similarly, a final asymptotic state is made of free particles at $t \to +\infty$. Hence each particle must be on the left-hand side of all the others that move faster than it. The final asymptotic states can then be represented by

$$|\, A_{a_1}(\theta_1)\, A_{a_2}(\theta_2)\, \ldots A_{a_n}(\theta_n)\rangle, \tag{17.2.6}$$

but this time with an *increasing* order of the rapidities, i.e.

$$\theta_1 \leq \theta_2 \leq \theta_3 \cdots \leq \theta_n. \tag{17.2.7}$$

Obviously each product (17.2.3) can always be ordered in the way we like by means of a certain number of commutations of the symbols $A_i(\theta_i)$ between neighbor particles. As we will see below, each commutation can be interpreted as a scattering process of two particles. It is custom any to normalize the states as

$$\langle A_i(\theta_1) \mid A_j(\theta_2)\rangle \;=\; 2\pi\delta_{ij}\,\delta(\theta_1 - \theta_2). \tag{17.2.8}$$

Consequently, the density of states with rapidities $(\theta, \theta + d\theta)$ is given by $d\theta/2\pi$.

17.2.2 Conserved Charges

The existence of an infinite number of conserved charges $\mathcal{Q}_{\pm s}$ in involution has a series of significant consequences on the scattering processes. As discussed in the previous chapter, the charges can be identified by their *spin* index s and the local ones[6] can be expressed by the integral of their current densities

$$\mathcal{Q}_s \;=\; \int \left[T_{s+1}(z,\bar z)\,dz + \Theta_{s-1}(z,\bar z)\,d\bar z\,\right], \;\; s \geq 1,$$

where $T_{s+1}(z,\bar z)$ and $\Theta(z,\bar z)$ are local fields that satisfy the conservation law

$$\partial_{\bar z}\, T_{s+1} \;=\; \partial_z\, \Theta_{s-1}.$$

Analogously, for the charges with negative spins, hereafter denoted also by $\bar{\mathcal{Q}}_s$, we have

$$\bar{\mathcal{Q}}_s \;=\; \int \left[\bar T_{s+1}(z,\bar z)\,dz + \bar\Theta_{s-1}(z,\bar z)\,d\bar z\,\right],$$

with

$$\partial_z\, \bar T_{s+1} \;=\; \partial_{\bar z}\, \bar\Theta_{s-1}.$$

Note that $\mathcal{Q}_{\pm 1}$ coincide with the light-cone components of the momentum

$$\begin{aligned} \mathcal{Q}_1 \;&=\; P \;=\; P^{(0)} + P^{(1)}, \\ \mathcal{Q}_{-1} \;&=\; \bar P \;=\; P^{(0)} - P^{(1)}. \end{aligned} \tag{17.2.9}$$

[6] In some integrable theories, such as the Sine–Gordon model or the nonlinear $O(3)$ sigma model, there are also non-local conserved charges, often associated to operators with fractional spin.

Since, by hypothesis, these charges commute among themselves

$$[\mathcal{Q}_s, \mathcal{Q}_{s'}] \,=\, 0,$$

they can be diagonalized simultaneously. The spectrum of the values s of the conserved charges varies by varying the theory and, as we shall see, it is deeply related to the structure of the bound states. Their action of the one-particle states leads to

$$\mathcal{Q}_s \;|\; A_a(\theta)\rangle \,=\, \omega_s^{(a)}(\theta) \;|\; A_a(\theta)\rangle, \tag{17.2.10}$$

where the functional dependence of the functions $\omega_s^{(a)}(\theta)$ is determined by the tensor properties of $\mathcal{Q}_s$: under the Lorentz group, $\mathcal{Q}_{|s|}$ trasforms as s copies of P while $\mathcal{Q}_{-s}$ as s copies of $\bar{P}$, and it is then natural to regard $\mathcal{Q}_{\pm s}$ as tensors of rank s. Hence we can impose

$$\omega_s^{(a)}(\theta) \,=\, \chi_s^{(a)}\, e^{s\,\theta}, \tag{17.2.11}$$

where $\chi_s^{(a)}$ is called the *eigenvalue* of the charge $\mathcal{Q}_s$ on the particle a. The spectrum of these eigenvalues is a problem interesting in itself that will be faced in some examples discussed in Section 17.5.1.

Further restrictions may come from the discrete symmetries of the model. For instance, if the theory is invariant under charge conjugation C, the conserved charges can be classified as even or odd operators $\mathcal{Q}_s^{(\pm)}$ with respect to C. Furthermore, assuming that the parity P is also a symmetry of the system, one can show the validity of the commutation relations

$$\begin{aligned} C\,\mathcal{Q}_s^{(+)}\,C \,&=\, \mathcal{Q}_s^{(+)} \,=\, (-1)^{s+1}\,\mathcal{Q}_s^{(+)} \\ C\,\mathcal{Q}_s^{(-)}\,C \,&=\, -\mathcal{Q}_s^{(-)} \,=\, (-1)^{s+1}\,\mathcal{Q}_s^{(-)}. \end{aligned} \tag{17.2.12}$$

They imply that the values of s for the C-even charges are only odd numbers, while those of the C-odd charges are even integers.

Let's now analyze how the infinite conserved charges constrain the scattering processes. S. Coleman and J. Mandula, in their famous paper, have shown that in $(3+1)$-dimensional theories the existence of only one conserved charge of tensor rank larger than 2 implies a trivial S-matrix, i.e. $S = 1$. This result does not apply to the $(1+1)$-dimensional theories but, in this case, there is a series of severe constraints that are listed below.

1. The number of particles with mass m_a remains the same before and after the collision has taken place.
2. The set of the final momenta of the particles is the same of the initial momenta, namely the scattering processes are purely elastic.
3. The scattering amplitude for the process in which n particles are involved can be completely factorized in terms of the $n(n-1)/2$ elastic scattering two-particle amplitudes.

Let's now prove these properties.

17.2.3 Elasticity in the Scattering Processes

In order to prove the elasticity of the scattering processes note that the conserved charges act on the multiparticle states as

$$\mathcal{Q}_s \mid A_{a_1}(\theta_1) \dots A_{a_n}(\theta_n)\rangle = \sum_{i=1}^{n} \chi_s^{(a_i)} e^{s\theta_i} \mid A_{a_1}(\theta_1) \dots A_{a_n}(\theta_n)\rangle.$$

Since

$$\frac{d\mathcal{Q}_s}{dt} = 0,$$

there is an infinite sequence of constraints that involve the sum of the higher powers of the momenta of the initial and final particles

$$\sum_{i\in in} \chi_s^{(a_i)} e^{s\theta_i} = \sum_{j\in fin} \chi_s^{(a_j)} e^{s\theta_j}. \tag{17.2.13}$$

The only solution to these infinite numbers of equations (apart from the permutations of particles with the same quantum numbers) corresponds to the case in which the final and the initial sets of rapidity are equal. Hence, in theories having an infinite number of conserved charges, the annihilation and production processes are absent: all scattering processes are therefore elastic.

17.2.4 Factorization of the Scattering Processes

In addition to being elastic, the scattering processes in these theories are also factorized. For a heuristic explanation of this feature, it is necessary to understand the action of the conserved charges $\mathcal{Q}_s$ on a localized wavepacket. If Q_s is the space component of the two charges $\mathcal{Q}_{\pm s}$, assuming for simplicity that $\chi_s^{(a)} = 1$ we have

$$e^{icQ_s} \mid A_a(p)\rangle = e^{icp^s} \mid A_a(p)\rangle.$$

Now, let

$$\psi(x) = \int_{-\infty}^{+\infty} dp\, e^{-a(p-p_0)^2} e^{ip(x-x_0)},$$

be the wavefunction of a state that is well localized both in momentum space (around $p = p_0$) and in coordinate space (around $x = x_0$). Acting by e^{icQ_s} on this state we have

$$\tilde{\psi}(x) = \int_{-\infty}^{+\infty} dp\, e^{-a(p-p_0)^2} e^{ip(x-x_0)} e^{icp^s}.$$

This new function is now localized at $x = x_0 - scp_0^{s-1}$, as can be seen by a saddle point computation. Hence, for $s > 1$, the center of the wavepacket is translated by a quantity that depends on the $(s-1)$th power of its momentum (for $s = 1$, Q_s coincides with the ordinary momentum operator that shifts equally all wavepackets by the same

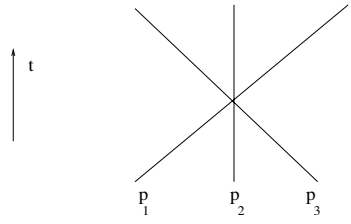


Fig. 17.6 *A simultaneous collision of three particles.*

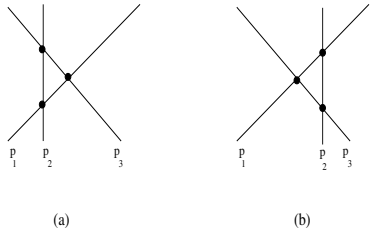


Fig. 17.7 *A three-particle collision realized by a sequence of two-particle collisions. These two cases and the one drawn in the previous figure are related by a symmetry transformation and therefore have the same amplitude.*

amount). The above result shows that wavepackets with different momenta can be shifted differently acting on them with the conserved charges e^{icQ_s} of higher spin.[7]

Consider now the collision of three particles of momenta $p_1 < p_2 < p_3$, associated to wavepackets well-localized both in momentum and coordinate space. Depending on the initial positions of the three packets, we can have three types of collisions, as shown in Figs 17.6 and 17.7, respectively. The first type consists of the simultaneous collision of the three particles. The other two types are drawn in Fig. 17.7, in which the scattering process is made of three distinct *two-particle* collisions, well separated in space and time. Obviously the chronological sequence of these collisions is different in the two graphs of Fig. 17.7.

In a generic scattering theory, the processes relative to Figs 17.6 and 17.7 have different amplitudes. However, for integrable theories, the three different situations can be obtained one from the other by an appropriate action of the operators e^{icQ_s}. Since these operators commute with the hamiltonian of the system (associated to $Q_{\pm 1}$), their action must lead to equivalent physical situations. Therefore, in integrable theories, there is equality of the three scattering amplitudes! We have thus achieved two extremely important results:

- Since in an integrable theory the S-matrix of a three-particle process can be factorized in two different but equivalent ways (corresponding to the different

[7]This result clarifies the Coleman–Mandula theorem. In fact, in $(d+1)$-dimensional theories, with $d > 1$, the possibility to translate differently particles of different momenta means that their trajectories can never cross: theirs is a free motion without collision and therefore $S = 1$.

sequences of two-particle collisions shown in Fig. 17.7), the two-particle scattering amplitudes $S^2(p_a, p_b)$ must satisfy the so-called *Yang–Baxter* equation[8]

$$S^2(p_2,p_3)\,S^2(p_3,p_1)\,S^2(p_1,p_2) \;=\; S^2(p_1,p_2)\,S^2(p_1,p_3)\,S^2(p_2,p_3). \qquad (17.2.14)$$

- The previous result can be easily generalized to n-particle processes. In fact, it is easy to show that the fulfilment of the Yang–Baxter equations (17.2.14) are sufficient and necessary conditions for the factorization of this amplitude in terms of the $n(n-1)/2$ two-particle elastic amplitudes. As before, in these collisions a possible exchange of the momenta can occur only between particles with the same mass and the same quantum numbers.

For the properties of elasticity and factorization, the S-matrix theory of a two-dimensional system is drastically simplified and the explicit expression for S can be found for many important physical models. It is in fact sufficient to find the two-particle scattering amplitudes to have full control over any other scattering processes. In turn, the two-particle scattering amplitudes can be found as solutions of the Yang–Baxter equation, together with the general requirements of unitarity and crossing symmetry.

17.3 Unitarity and Crossing Invariance Equations

In this section we discuss the unitary and crossing symmetry equations that hold for the two-particle elastic scattering amplitudes of a (1+1)-dimensional integrable theory.

Let p_1 and p_2 be the initial and final momenta of the incoming particles A_i and A_j and the outgoing ones A_l and A_k, as shown in Fig. 17.8. In addition to the delta function $\delta^{(2)}(p_1+p_2-p_3-p_4)$ of the conservation of the total energy and momentum, the Lorentz invariance equires that the scattering amplitude depends on the particle momenta only by their invariant combinations, given by the Mandelstam variables s, t, and u defined in eqn (17.1.17). Note that for the $(1+1)$-dimensional systems and for the elasticity of the scattering process u vanishes identically, $u = 0$, while s and t can both be expressed in terms of the difference of the rapidites of the particles.[9] In fact, using the parameterization (17.2.1), the Mandelstam variable s of the process

$$A_i\,A_j \;\to\; A_k\,A_l,$$

is given by

$$s(\theta_{ij}) \;=\; (p_1+p_2)^2 \;=\; m_i^2 + m_j^2 + 2\,m_i\,m_j\,\cosh\theta_{ij}, \qquad \theta_{ij} = \theta_i - \theta_j. \qquad (17.3.1)$$

For the physical processes θ_{ij} assumes real values and consequently also s is real and takes values $s \geq (m_i + m_j)^2$. The Mandelstam variable t is instead given by

$$t(\theta_{ij}) \;=\; (p_1-p_2)^2 \;=\; m_i^2 + m_j^2 - 2\,m_i\,m_j\,\cosh\theta_{ij}. \qquad (17.3.2)$$

[8] The detailed matrix structure of this equation will be specified later.

[9] For the elastic processes in the $(1+1)$-dimensional system there is only one independent Mandelstam variable for the equalities $p_3 = p_2$ and $p_1 = p_4$ of the momenta.

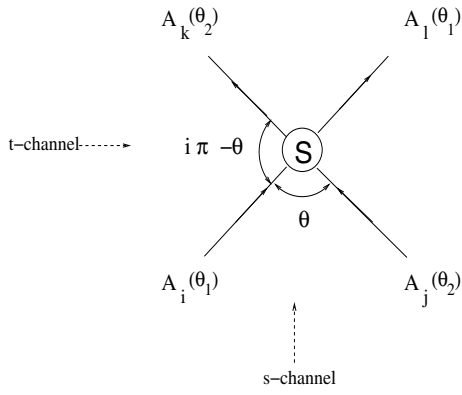


Fig. 17.8 *Elastic scattering process of two particles.*

Consequently, we can switch between the s and the t-channels by the analytic continuation

$$t(\theta) \;=\; s(i\pi - \theta), \tag{17.3.3}$$

which admits the natural geometrical interpretation shown in Fig. 17.8, if we regard θ as the (imaginary) angle between the lines of the incoming particles.

In (1+1)-dimensional systems, the two-particle S-matrix elements are defined by[10]

$$\mid A_i(\theta_1)\, A_j(\theta_2) \rangle \;=\; S_{ij}^{kl}(\theta) \;\mid\; A_k(\theta_2)\, A_l(\theta_1) \rangle, \tag{17.3.4}$$

with $\theta = \theta_{12}$ and $\theta_1 > \theta_2$, consistently with the definition of the initial and final asymptotic states previously discussed. In this equation a sum over the indices k and l is implicit; this occurs if the particles with $k \neq i$ and $l \neq j$ are not distinguished by any eigenvalues of the conserved charges. Note that the dependence of the S-matrix on the difference of the rapidities is a consequence of the relativistic invariance of the theory, since a Lorentz transformation changes the value of the rapidity of each particle by a constant. There is a relation between the S-matrix given above and the one written in terms of the original Mandelstam variable s, here denoted by $\mathcal{S}$: this relation is given by the jacobian of the transformation $s(\theta)$

$$\mathcal{S}_{ij}^{kl}(s) \;=\; 4m_i m_j \,\sinh\theta\, S_{ij}^{kl}(\theta). \tag{17.3.5}$$

Constraints from discrete symmetries. In an elastic scattering theory with r types of particles, the set of r^4 functions $S_{ij}^{kl}(\theta)$ completely determines the full S-matrix of the problem. However these functions are not all independent. First of all, the matrix elements $S_{ij}^{kl}(\theta)$ are non-zero only when the particles A_i and A_k (as well as A_j and A_l) have the same quantum numbers with respect to the conserved charges. This implies, in particular, the equality of their masses $m_i = m_k$ and $m_j = m_l$. Moreover, assuming

[10] In these theories it is customary to define S as the unitary operator that maps the *initial states* onto the *final states*, i.e. $\mid \text{in}\rangle = S \mid \text{fin}\rangle$. This is the definition that we will use hereafter. Strictly speaking, this definition corresponds to the operator S^{-1} previously introduced.

the invariance of the theory under the charge conjugation C, the parity P and the time reversal T, there are the further relations

$$\begin{aligned} S_{i\,j}^{k\,l}(\theta) &= S_{j\,i}^{l\,k}(\theta), \quad P \\ S_{i\,j}^{k\,l}(\theta) &= S_{\bar{i}\,\bar{j}}^{\bar{k}\,\bar{l}}(\theta), \quad C \\ S_{i\,j}^{k\,l}(\theta) &= S_{l\,k}^{j\,i}(\theta), \quad T \end{aligned} \tag{17.3.6}$$

where $\bar{a} = Ca$ denotes the antiparticle state.

Yang–Baxter equations. The Yang–Baxter equations impose additional equations on these amplitudes: the explicit form of these equations is (there is a sum over all the repeated indices)

$$S_{ij}^{ab}(\theta_{12})\, S_{bk}^{cl}(\theta_{13})\, S_{ac}^{nm}(\theta_{23}) \;=\; S_{jk}^{ab}(\theta_{23})\, S_{ia}^{nc}(\theta_{13})\, S_{cb}^{ml}(\theta_{12}). \tag{17.3.7}$$

These correspond to r^6 equations, in correspondence with the values of the six external indices i, j, k, l, m, n. This is an overdetermined set of equations because their number is larger that the r^4 amplitudes to be determined. Hence, solutions of these equations can only be found for special functional forms of the functions $S_{ij}^{kl}(\theta)$. Note that, from their homogeneity, the Yang–Baxter equations (17.3.7) can only fix the ratios of the scattering amplitudes. Some explicit examples of solutions will be considered in later sections.

Let's now focus our attention on the analytic properties of the scattering amplitudes. They can be derived by specializing the general considerations presented in the first section of this chapter. We will initially consider the analytic properties with respect to the Mandelstam variable s, to translate them later in terms of the rapidity θ. We have the following properties:

- $S(s)$ is a one-value analytic function in the complex plane of s with two elastic branch cuts, the first for $s \leq (m_i - m_j)^2$ and the second for $s \geq (m_i + m_j)^2$. The physical domains of this function are for values just above the branch cut on the right, i.e. $s^+ = s + i0$ and $s > (m_i + m_j)^2$. The first sheet of the Riemann surface of this function is called the *physical sheet*.
- S is a real analytic function, namely it assumes complex conjugate values at complex conjugate points

  $$S_{ij}^{kl}(s^*) \;=\; \left[S_{ij}^{kl}(s)\right]^* .$$

  In particular this implies that $S(s)$ assumes real values when s is itself real, with $(m_i - m_j)^2 \leq s \leq (m_i + m_j)^2$.

The unitarity equation is expressed by $S(s^+)S^\dagger(s^+) = 1$. This is a matrix relation, with a sum over all intermediate states between S and $S^\dagger$. When s^+ increases, it is energetically possible that states with a higher number of particles enter this sum, giving rise to production processes and consequently to additional branch cuts of $S(s)$. However this circumstance does not occur in integrable theories and, in this case, the unitarity conditions involve only the two-particle states

$$S_{ij}^{kl}(s^+)\, \left[S_{kl}^{nm}(s^+)\right]^* \;=\; \delta_i^n\, \delta_j^m .$$

Using the real analyticity of these functions, this equation can be written as

$$S_{ij}^{kl}(s^+)\, S_{kl}^{nm}(s^-) \;=\; \delta_i^n\, \delta_j^m,$$

with $s^- = s - i0$. This equation shows the necessity to introduce a branch cut at $s = (m_i + m_j)^2$ and, furthermore, that this branch cut is of the square root type. To prove this, let $S_\gamma(s)$ be the function obtained by the analytic continuation of $S(s)$ after an anticlockwise path around that point. The unitarity condition imposes the validity of $S(s^+)S_\gamma(s^+) = 1$ for all physical values of s^+. This relation can be analytically continued for all values of s, with the result

$$S_\gamma(s) \;=\; S^{-1}(s).$$

In particular, if s^- is a point below the cut, we have

$$S_\gamma(s^-) \;=\; S^{-1}(s^-) \;=\; S(s^+),$$

where the second equality follows from applying the unitarity equation twice. Since $S_\gamma(s^-)$ is just the analytic continuation of $S(s^+)$ obtained with a double twist around the point $s = (m_i + m_j)^2$, it follows that at this point there is a square-root singularity.

Concerning the second cut, the one that goes from $s = (m_i - m_j)^2$ to $s = -\infty$, it can be discussed using the fundamental invariance of the relativistic scattering theories under the crossing transformations. In fact, if one of the incoming particles, say the one with index j, inverts its motion so that it becomes an outgoing particle and the same operation is done with the outgoing particle of index l to transform it into an incoming particle, the original amplitude becomes the amplitude of another scattering channel. For this new amplitude we have then i and $\bar{l}$ as incoming particles, and k and $\bar{j}$ as outgoing particles, where the symbols $\bar{a}$ denotes the antiparticles. This ends up looking at Fig. 17.8 from left to right, instead of from bottom to top, so that now the direct channel is described by the Mandelstam variable t instead of the original variable s. Since in this new process $p_2 = p_3$, the relation between s and t is simply

$$t \;=\; (p_1 - p_2)^2 \;=\; 2p_1^2 + 2p_2^2 - (p_1 + p_2)^2 \;=\; 2m_i^2 + 2m_j^2 - s.$$

The crossing invariance permits us to recover the amplitude relative to this scattering process by means of the analytic continuation of the original amplitude in the region of the s plane where the variable t assumes physical values, i.e. $t \in \Re\,\mathrm{e}\; t \geq (m_i + m_j)^2$. The physical amplitudes are then related by

$$S_{ij}^{kl}(s^+) \;=\; S_{i\bar{l}}^{k\bar{j}}(2m_i^2 + 2m_j^2 - s^+). \tag{17.3.8}$$

Also here it is easy to prove that the point $s = (m_i - m_j)^2$ has a square-root branch singularity. This does not imply though that the Riemann surface associated to the function $S(s)$ is only made of two sheets. In fact, by an analytic continuation of this function along a path that crosses the branch cut on the left we may reach a different sheet than the one obtained by an analytic continuation through the cut on the right. Hence, moving up or down these sheets and crossing the left and right cuts, we can

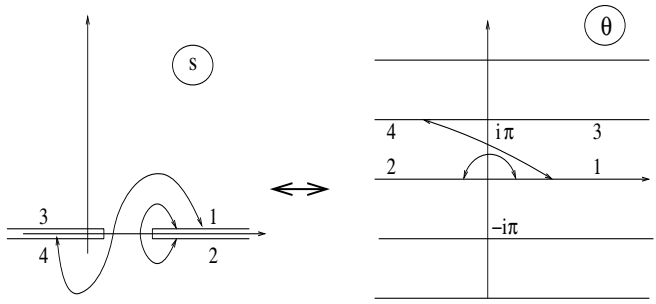


Fig. 17.9 *Map between the s-plane and the θ-plane, together with the unitarity and crossing symmetry conditions.*

span the Riemann surface of the S-matrix, made in general of several sheets, possibly infinite.

Let's now translate the considerations above in terms of the rapidity variable. Note that the inverse transformation of (17.3.1)

$$\theta_{ij} \;=\; \log\left[\frac{s - m_i^2 - m_j^2 + \sqrt{[(s-(m_i+m_j)^2)(s-(m_i-m_j)^2]}}{2m_i m_j}\right],$$

maps the physical sheet of the s-plane in the strip $0 \leq \mathrm{Im}\,\theta_{ij} \leq \pi$. The second sheet is instead mapped in the strip $-\pi \leq \mathrm{Im}\,\theta_{ij} \leq 0$. This structure repeats with period $2\pi i$, as shown in Fig. 17.9. Moreover, as shown in eqn (17.3.2), the Mandelstam variable t is obtained by substituting $\theta_{ij} \to i\pi - \theta_{ij}$ in eqn (17.3.1). Hence, the map (17.3.1) realizes a uniformization of the original analytic structure, since in the plane of the variable θ there are no longer branch cuts. This implies that the S-matrix, considered as a function of θ, is an analytic function at the image points of the original cuts, i.e. at 0 and $i\pi$, as well as at all other points $in\pi$ of the other sheets. Since the integrability of the theory guarantees that these are the only branch points of the original amplitude, we arrive to the important result that $S(\theta)$ is a meromorphic function of θ. Since $S(s)$ is a real analytic function, $S(\theta)$ assumes real values on the imaginary axis of θ.

Expressed in terms of θ, the unitarity condition becomes

$$\sum_{n,m} S_{ij}^{nm}(\theta)\, S_{nm}^{kl}(-\theta) \;=\; \delta_i^k\, \delta_j^l, \tag{17.3.9}$$

with the crossing invariance condition

$$S_{i\,j}^{k\,l}(\theta) \;=\; S_{i\,\bar{l}}^{k\,\bar{j}}(i\pi - \theta). \tag{17.3.10}$$

It is interesting to stress some important aspects of the formulation of the S-matrix theory in terms of the rapidity variable. The first aspect is that the unitarity and crossing symmetry equations can be analytically continued for arbitrary values of θ and therefore they hold in all the complex plane of this variable. The second aspect

concerns the definition itself of the S-matrix that, as a function of θ, can be written in an operator form as

$$A_i(\theta_1)\, A_j(\theta_2) \;=\; S_{ij}^{kl}(\theta)\, A_k(\theta_2)\, A_l(\theta_1). \tag{17.3.11}$$

This equation defines an algebra for the symbols $A_a(\theta)$, the so-called *Faddev–Zamolodchikov algebra*. Therefore the scattering processes can be equivalently interpreted as commutation relations among the operators that create the particles. In this respect, the unitarity equation (17.3.9) can be seen as a simple consequence of this algebra. Analogously, the Yang–Baxter equations simply derive by the associativity condition of the Faddev–Zamolodchikov algebra, as shown in Problem 4.

17.4 Analytic Structure and Bootstrap Equations

The elastic S-matrices are meromorphic analytic functions in the complex plane of θ. The bound states, originally associated to the simple poles of these amplitudes in the interval of s between $(m_i - m_j)^2$ and $(m_i + m_j)^2$, correspond now to simple poles with positive residue[11] along the imaginary segment $(0, i\pi)$ of the θ variable. Consider an S-matrix with incoming particles A_i and A_j that has a simple pole in the s-channel at $\theta = i\, u_{ij}^n$. Corresponding to of this pole, the amplitude can be expressed as

$$S_{ij}^{kl} \;\simeq\; i\, \frac{R^{(n)}}{\theta - i u_{ij}^n}, \tag{17.4.1}$$

with the residue $R^{(n)}$ related to the *on-shell* vertex functions of the incoming particles and the bound state A_n, as shown in Fig. 17.10

$$R^{(n)} \;=\; f_{ij}^n\, f_{kl}^n. \tag{17.4.2}$$

A non-zero value of f_{ij}^n obviously implies a pole singularity in the other two amplitudes S_{in} and S_{jn} as well, where the poles are now due to the bound states A_j and A_i. Since in the bootstrap approach the bound states are on the same footing as the asymptotic states, there is an important relation among the masses of the system: if $\theta = i u_{ij}^n$ is the position of the pole in the scattering of the particles A_i and A_j, the mass of the bound state is given by

$$m_n^2 \;=\; m_i^2 + m_j^2 + 2 m_i m_j\, \cos u_{ij}^n. \tag{17.4.3}$$

This relation is simply obtained by substituting in the Mandelstam variable s given in eqn (17.3.1) the resonance condition $\theta = i u_{ij}^n$. Notice that this formula expresses a well-known geometrical relation, known as Carnot's theorem, among the sides of a triangle (here equal to the values of the masses), where u_{ij}^n is one of the external angles as shown in Fig. 17.11). This figure clearly highlights the symmetric role played by the three particles.

[11] In the next chapter we will see that this concept can be generalized both to the cases of poles with negative residues and higher order poles.

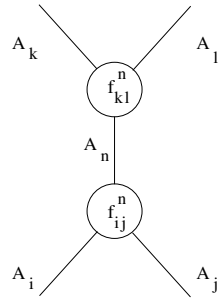


Fig. 17.10 *Residue of the pole and its expression in terms of the* on-shell *coupling constants.*

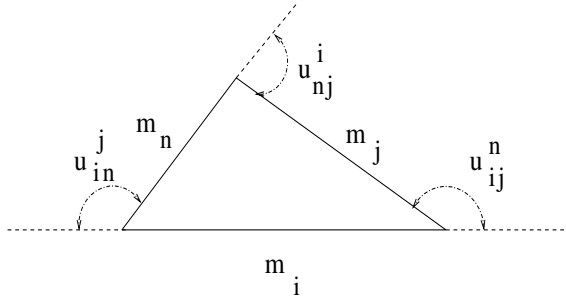


Fig. 17.11 *Mass triangle.*

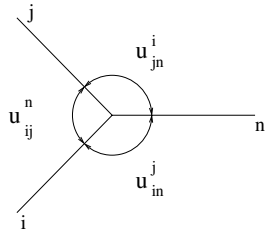


Fig. 17.12 *Relation among the positions of the poles.*

From the deep geometrical nature of the quantities involved in this formulation and as a consequence of (17.4.3), it is easy to show that the positions of the poles in the three channels satisfy

$$u_{ij}^{n} + u_{in}^{j} + u_{jn}^{i} \;=\; 2\pi. \tag{17.4.4}$$

This relation, shown in Fig. 17.12, expresses a well-known properties of the external angles of a triangle.

As we are going to see later, the elastic S-matrix of (1+1)-dimensional systems may also have higher order poles, whose interpretation stays in the singularities coming from multiple scattering processes. Instead of an abstract discussion, we prefer to illustrate their features later by means of some explicit examples.

Diagonal S-matrices. To proceed further in the discussion of the analytic structure of the elastic S-matrices, it is convenient to make an additional simplification in the theory so far presented. This simplification occurs in two cases: (i) when the system has a non-degenerate mass spectrum and (ii) when the system has a degenerate spectrum but with all particles uniquely identified thanks to the different eigenvalues with respect to the conserved charges. In both cases, the elasticity of the scattering processes enforces the vanishing of the reflection amplitude (see Problem 5): the corresponding S-matrix is then completely diagonal and the Yang–Baxter equations are then identically satisfied. The unitarity and crossing symmetry conditions simplify as follows

$$S_{ab}(\theta)\, S_{ab}(-\theta) \,=\, 1, \quad S_{ab}(\theta) \,=\, S_{a\bar{b}}(i\pi - \theta), \tag{17.4.5}$$

where $\bar{b}$ is the antiparticle of b. These two equations imply that the amplitudes $S_{ab}(\theta)$ are periodic functions of θ with period $2\pi i$: in this case the Riemann surface of the S-matrix consists of a double covering of the complex plane s. Remarkably enough, there is a general solution of eqn (17.4.5) that can be expressed in terms of products of the meromorphic functions

$$s_x(\theta) \,=\, \frac{\sinh \frac{1}{2}(\theta + i\pi x)}{\sinh \frac{1}{2}(\theta - i\pi x)}. \tag{17.4.6}$$

From their periodicity, the parameter x can always be chosen as $-1 \le x \le 1$. In the double covering of the original variable s, i.e. in the strip $-\pi \le \operatorname{Im}\theta < \pi$, these functions have a simple pole at $\theta = i\pi x$ and a simple zero at $\theta = -i\pi x$. Moreover, they have the properties

$$\begin{aligned} &s_x(\theta)\, s_x(-\theta) = s_x(\theta) s_{-x}(\theta) = 1, \\ &s_x(\theta) = s_{x+2}(\theta) = s_{-x}(-\theta), \\ &s_0(\theta) = -s_1(\theta) \,=\, 1, \\ &s_x(i\pi - \theta) \,=\, -s_{1-x}(\theta). \end{aligned} \tag{17.4.7}$$

A suggestive interpretation of these functions is proposed in Problem 8.

When the particles involved in the scattering are instead neutral, i.e. when the particles coincide with their antiparticles, the solution of eqns (17.4.5) can be expressed in terms of the functions

$$f_x(\theta) \,=\, s_x(\theta)\, s_x(i\pi - \theta) \,=\, \frac{\tanh \frac{1}{2}(\theta + i\pi x)}{\tanh \frac{1}{2}(\theta - i\pi x)}. \tag{17.4.8}$$

The simple poles of these functions are at $\theta = i\pi x$ and $\theta = i\pi(1 - x)$ and they are related by the crossing transformation. They also have simple zeros at $-i\pi x$ and $-i\pi(1 - x)$. Important properties of these functions are

$$f_x(\theta) \,=\, f_x(i\pi - \theta) \,=\, f_{1-x}(\theta), \quad f_x(-\theta) \,=\, f_{-x}(\theta) \,=\, 1/f_x(\theta). \tag{17.4.9}$$

In summary, as a consequence of the unitarity and crossing symmetry equations, any amplitude $S_{ab}(\theta)$ of a diagonal S-matrix can be expressed as

$$S_{ab}(\theta) \;=\; \prod_{x\in\mathcal{A}_{ab}} s_x(\theta), \tag{17.4.10}$$

if there are charged particles, or by

$$S_{ab}(\theta) \;=\; \prod_{x\in\mathcal{A}_{ab}} f_x(\theta), \tag{17.4.11}$$

if the particles are neutral.

Bootstrap principle. The unitarity and crossing symmetry equations alone are not, however, able to fix the position of the poles of these amplitudes, namely to determine the sets $\mathcal{A}_{ab}$. To achieve this aim it is necessary to make use of a dynamical condition. This is provided by the bootstrap principle that posits that the bound states are on the same footing as the asymptotic states. As a consequence, the amplitudes that involve the bound states can be obtained in terms of the amplitudes of the external particles and vice versa. This translates into an additional equation satisfied by the scattering amplitudes

$$S_{i\bar{l}}(\theta) \;=\; S_{ij}(\theta + i\bar{u}^k_{jl})\, S_{ik}(\theta - i\bar{u}^j_{lk}), \tag{17.4.12}$$

where

$$\bar{u}^c_{ab} \;\equiv\; \pi - u^c_{ab}. \tag{17.4.13}$$

This equation comes from the commutativity of the two processes shown in Fig. 17.13, obtained one from the other by the translation of the world-line of the asymptotic particle A_i (see Problem 6).

Rules of the game. To summarize, in order to determine the S-matrix by the bootstrap approach one has to find a set of poles relative to all amplitudes S_{ab} that are compatible with the bootstrap equation (17.4.12) and that can be interpreted in terms of bound states or multiparticle scattering processes of the asymptotic particles themselves. The masses of the particles are determined by the relation (17.4.3). In practice this means starting from the amplitude that involves the lighest particle, therefore with the simplest pole structure, and then iteratively applying the bootstrap equations (17.4.12) in order to get the scattering amplitudes involving the bound states of higher mass.

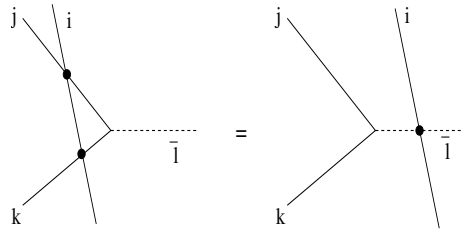


Fig. 17.13 *Bootstrap equation that links the S-matrix amplitudes, where $A_{\bar{l}}$ is the bound state in the scattering process of the particles A_j and A_k.*

It should be stressed, though, that not all the choices of the initial amplitude give rise to consistent bootstrap systems. Presently, the theoretical problem of classifying in their full generality the integrable models in the bootstrap interaction is still open. Valuable information is gained by the spectrum of the conserved charges, as discussed in the next section. Important examples of consistent S-matrices will be given in the next chapter and they are extremely helpful to clarify several aspects of the iterative bootstrap procedure.

To simplify the repeated applications of the bootstrap equations (17.4.12), it is useful to define the operator $\mathcal{R}^y$, whose application to a function $G(\theta)$ is given by

$$\mathcal{R}^y(G(\theta)) \,=\, G(\theta + i\pi y)\, G(\theta - i\pi y).$$

Applying $\mathcal{R}^y$ to the functions $s_x(\theta)$ and $f_x(\theta)$ and using their properties, one has

$$\begin{aligned} \mathcal{R}^y(s_x(\theta)) \,&=\, s_{x+y}(\theta)\, s_{x-y}(\theta), \\ \mathcal{R}^y(f_x(\theta)) \,&=\, f_{x+y}(\theta)\, f_{x-y}(\theta). \end{aligned}$$

They also have the commutative and distributive properties

$$\mathcal{R}^y(\mathcal{R}^z(G)) \,=\, \mathcal{R}^z(\mathcal{R}^y(G)), \quad \mathcal{R}^y(G_1)\, \mathcal{R}^y(G_2) \,=\, \mathcal{R}^y(G_1\, G_2).$$

Finally, if a function $G(\theta)$ satisfies the equation

$$G(\theta) \,=\, G(i\pi - \theta) \,=\, 1/G(-\theta), \tag{17.4.14}$$

the same holds for the function transformed by $\mathcal{R}^y$.

17.5 Conserved Charges and Consistency Equations

In this section we study the relation between the spins of the conserved charges and the bound states of a scattering theory. The integrals of motion $\mathcal{Q}_s$ are a set of dynamical data relative to each scattering theory. If the lagrangian of the model was known, it would be possible in principle to determine them explicitly. Knowledge of the S-matrix alone leads only to some constraints on the values of the spins s. It also leads to the determination of the ratios of the eigenvalues of $\mathcal{Q}_s$. As shown below, these results derive from the bootstrap principle and the locality properties of the conserved charges.

Let $\mathcal{Q}_s$ be the set of all conserved charges. Since they commute with each other, they can be simultaneously diagonalized together with the hamiltonian, and the asymptotic states $A_a(\theta)$ are also eigenvectors of $\mathcal{Q}_s$

$$\mathcal{Q}_s \;|\; A_a(\theta)\rangle \,=\, \chi_s^{(a)}\, e^{s\theta} \;|\; A_a(\theta)\rangle. \tag{17.5.1}$$

For a conserved charge of spin s, there exists at least an eigenvalue $\chi_s^{(a)}$ different from zero. Note that $\chi_1^{(a)}$ is simply the mass of the particle a, $\chi_1^{(a)} = m_a$. The locality of the conserved charges implies that their action on the multiparticle states is given by

$$\mathcal{Q}_s \;|\; A_{a_1}(\theta_1) \cdots A_{a_n}(\theta_n)\rangle \,=\, (\omega_s^{(a_1)}(\theta_1) + \cdots \omega_s^{(a_n)}(\theta_n)) \;|\; A_{a_1}(\theta_1) \cdots A_{a_n}(\theta_n)\rangle. \tag{17.5.2}$$

Suppose that the amplitude S_{ab} presents a pole at $\theta = iu^c_{ab}$ corresponding to the bound state $A_{\bar{c}}$. Correspondingly, this can be defined as

$$\lim_{\epsilon\to 0} \epsilon \mid A_a(\theta + i\bar{u}^b_{ac} + \epsilon)\, A_b(\theta - i\bar{u}^a_{bc})\rangle = \mid A_{\bar{c}}(\theta)\rangle.$$

Now applying $\mathcal{Q}_s$ to both terms of this equation and using eqns (17.5.1) and (17.5.2), one obtains an infinite-dimensional homogeneous system of linear equations for the eigenvalues $\chi^{(a)}_s$:

$$\chi^{(a)}_s\, e^{is\bar{u}^b_{ac}} + \chi^{(b)}_s\, e^{-is\bar{u}^a_{bc}} = \chi^{(\bar{c})}_s. \tag{17.5.3}$$

A solution of this system is obviously $\chi^{(i)}_s = 0$ $(\forall s, i)$. However, this is not an interesting solution because it implies the absence of all conserved charges. Non-trivial solutions can be found only for particular values of the resonance angles u^c_{ab} of the S-matrix, corresponding to the vanishing of the determinant of the homogeneous linear system (17.5.3).

Consider, for instance, the case in which $a = b$, with $\chi^{(a)}_s \neq 0$. Equation (17.5.3) can be written in this case as

$$2\,\cos(s\,\bar{u}^a_{ac}) = \frac{\chi^{(c)}_s}{\chi^{(a)}_s}. \tag{17.5.4}$$

If the bound state c corresponds to the same initial particle a, this equation admits the solutions

$$\bar{u}^a_{aa} = \frac{\pi}{3}, \quad s = 1, 5 \pmod 6. \tag{17.5.5}$$

Note that the exact value of the resonance angle $\bar{u}^a_{aa} = \frac{\pi}{3}$ comes directly from the geometry of the mass triangle, in this case an equilateral triangle. The S-matrix of this example presents the so-called Φ^3 property, since the particle A_a is simultaneously a bound state of itself. Read in reverse, this result hints that each time that the spectrum of the conserved spins consists of integer numbers that are not divisible by both 2 and 3, the particle mass spectrum may present the Φ^3 property.

To proceed in our analysis, it is useful to introduce the notion of *bootstrap fusion rules*. Let A_a be the operator that creates a particle a in the bootstrap interaction. The bound state structure can be encoded in this relation

$$A_i \times A_j = \sum_k n^k_{ij}\, A_k. \tag{17.5.6}$$

where n^k_{ij} are boolean variables, with values 0 and 1, different from zero only when A_k is the bound state of the scattering process of the particles A_i and A_j. Even though there is a strong analogy of this relation with the Verlinde algebra of the conformal field theory, it should be stressed that the bootstrap fusion rules do not form an associative algebra.

As mentioned above, the full classification of all bootstrap systems is still an open problem, even though there are strong indications that the only consistent systems are those related to Toda field theories or reductions thereof. Below we present only some simple but instructive examples of consistent bootstrap systems.

17.5.1 Non-degenerate Bootstrap Systems

Let's assume the existence of a non-trivial solution of the set of equations (17.5.3). From their homogeneous form, we can always choose to normalize to 1 all the non-zero eigenvalues of the lightest particle. For a neutral particle, it is easy to show by induction that all remaining eigenvalues are real. Equations (17.5.3) then split into two different sets

$$(\chi_s^{(c)})^2 \;=\; (\chi_s^{(a)})^2 + (\chi_s^{(b)})^2 + 2\chi_s^{(a)}\,\chi_s^{(b)}\,\cos(s\,u_{ab}^c),$$

$$\chi_s^{(a)}\,\sin(s\bar{u}_{ac}^b) \;=\; \chi_s^{(b)}\,\sin(s\bar{u}_{bc}^a).$$

The first provides a generalization of the mass triangle equation (17.4.3), while the second generalizes a simple geometrical property of this triangle. It should be stressed that the second equation is particularly useful from a computational point of view: to have non-zero values of $\chi_s^{(a)}$ and $\chi_s^{(b)}$, the ratio of the two trigonometric functions $\sin(s\bar{u}_{ac}^b)/\sin(s\bar{u}_{bc}^a)$ must in fact be *independent* of any bound state A_c in the channel $|\, A_a\, A_b\rangle$. Hence, knowing the resonance angle of any of the bound states in this channel, one can use this equation either to correctly identify the value of the others or, alternatively, to prove that it is impossible to have conserved charges of higher spins compatible with the structure of the bootstrap fusions.

Let's consider some significant examples of bootstrap systems that involve N particles, starting from the simplest case $N = 1$.

- **N=1**. In this case, assuming the existence of only one bound state, the only fusion process is the one that sees the particle as a bound state of itself:

$$A \,\times\, A \to A. \tag{17.5.7}$$

  The resonance angle is $u_{aa}^a = \frac{\pi}{3}$ and the only possible values of the spins of the conserved charges are

$$s \;=\; 1, 5 \pmod 6. \tag{17.5.8}$$

  A physical realization of this system is provided by the off-critical Yang–Lee model or by the Bullogh–Dodd lagrangian, as we will see in the next chapter.
- **N=2**. In addition to the reducible fusion rules $A_a \times A_a \to A_a$, $A_b \times A_b \to A_b$, consider the examples

$$\begin{array}{lll} (i) & A_a \times A_a \to A_b & , \;\; A_b \times A_b \to A_a \\ (ii) & A_a \times A_a \to A_a + A_b & , \;\; A_b \times A_b \to A_a. \end{array}$$

  The consistency equations of the processes (i) are

$$2\,\chi_s^{(a)}\,\cos(s\bar{u}_{ab}^a) \;=\; \chi_s^{(b)}, \qquad 2\,\chi_s^{(b)}\,\cos(s\bar{u}_{ab}^b) \;=\; \chi_s^{(a)}.$$

  For $\chi_s^{(a,b)} \neq 0$ they become

$$4\,\cos(s\bar{u}_{ab}^a)\,\cos(s\bar{u}_{ab}^b) \;=\; 1.$$

This equation admits two types of solutions

$$\bar{u}^a_{ab} = \frac{\pi}{12}, \quad \bar{u}^b_{ab} = \frac{5\pi}{12}, \quad s = 1, 4, 5, 7, 8, 11 \ (\text{mod } 12) \tag{17.5.9}$$

$$\bar{u}^a_{ab} = \frac{\pi}{5}, \quad \bar{u}^b_{ab} = \frac{2\pi}{5}, \quad s = 1, 3, 7, 9, \ (\text{mod } 10). \tag{17.5.10}$$

Note that the spectrum of s of the first solution coincides with the Coxeter exponents of the Toda field theory on $E_6^{(1)}$.
If we restrict our attention to neutral particles, there are no conserved spins with $s = 2k$. In this case the spectrum of the conserved spins of the first solution becomes

$$s = 1, 5, 7, 11 \ (\text{mod } 12). \tag{17.5.11}$$

It coincides with the Coxeter exponents of the Toda field theory based on $\tilde{F}_4 = E_6^{(2)}$, obtained by folding the original E_6 Dynkin diagram with respect to its Z_2 automorphism.
For the process (ii), as possible values of the spins it is necessary to take those compatible with the one-particle subprocess. For instance, for the solution (17.5.10), we have

$$s = 1, 7, 11, 13, 17, 19, 23, 29 \ (\text{mod } 30). \tag{17.5.12}$$

This spectrum coincides with the Coxeter exponents of the Toda field theory based on $E_8^{(1)}$.

- **Bootstrap chains**. For a generic bootstrap system of N neutral particles it is easy to analyze the case in which there is a bootstrap chain of bound states

$$A_k \times A_k \to A_{k+1} \ \ k = 1, 2, \ldots N, \qquad A_{N+1} = A_1.$$

The consistency equation is

$$\prod_{k=1}^{N} 2 \cos(s\, \bar{u}^k_{k,k+1}) = 1,$$

whose solution is given by

$$\bar{u}^k_{k,k+1} = \frac{k\pi}{2N+1} \tag{17.5.13}$$
$$s = 1, 3, \ldots, 2N-1, 2N+3, \ldots, 4N+1 \ (\text{mod } 4N+2).$$

In this case the spectrum of conserved spins coincides with the Coxeter exponents of the Toda field theories based on $A_{2N}^{(2)}$.

Appendix 17A. Historical Development of S-Matrix Theory

S-matrix theory is an interesting chapter in elementary particle physics and it is worth mentioning its basic developments. The reader can also consult the references at the end of the chapter for a broader perspective on the subject.

Proposed originally by W. Heisenberg to overcome the difficulties of quantum field theory in dealing with the divergences of the perturbative series, S-matrix theory received considerable attention during the 1950s and the 1960s, in particular in the study of strong interactions of hadronic particles, such as protons, neutrons, and pions. The enormous number of particles and hadronic resonances discovered during those decades made clear the difficulty of calling all of them *elementary particles*. Furthermore, it was discovered that the hadronic resonances present high values of their spin J, related to the square of their mass by a linear relation, $J = \alpha' m^2$, where the constant $\alpha' \sim 1$ (Gev)^2 is the *Regge slope*. The first attempts to use quantum field theory to describe the hadronic phenomena were very disastrous. There was in fact the difficulty of incorporating both the unstable particles (the resonances) and the particles with spin higher than 1: the only known consistent quantum field theories, i.e. renormalizable, are those limited to stable particles with spin 0, 1/2, and 1. The large values of the effective coupling constants coming from experiments also led to doubt about the efficiency and validity of the possible perturbativie theories for such processes.

From all these drawbacks, it was necessary to look for an alternative theory of the hadronic processes, to eventually extend to other interactions too. The new approach, based on a set of principles and on the analytic properties of the quantum amplitudes, was boosted under the name of *The analytic theory of the S-matrix*. Proposed and studied in great detail by the group of physicists in Berkeley, in particular by Chew and Mandelstam, the theory developed further with the important contributions by Weisskopf, Frautschi, Regge, and many others. Since the analysis of the scattering processes is the common and closest point between theory and experiment, the expectations were that the results derived by S-matrix theory should not depend on the existence or the absence of an underlying quantum field theory of the interactions. A fundamental theory based on the S-matrix should be able to answer a series of questions, such as the following:

1. What is the difference between stable and unstable particles? Does where exist a theoretical framework for both? As is well known, the lagrangian formulation of quantum field theory only makes use of the stable asymptotic particles and therefore it does not allow an equal footing for both cases.
2. Is it possible to determine the mass spectrum and the coupling constants of the theory? One should recall that, in a lagrangian theory, on the contrary, both masses and coupling constants are free parameters of the model.

The initial studies of the S-matrix as a function of the energy, momentum, angular momentum, etc., showed the suggestive circumstance that the analytic structure of the S-matrix appeared to be the simplest possible. This was assumed then as a principle

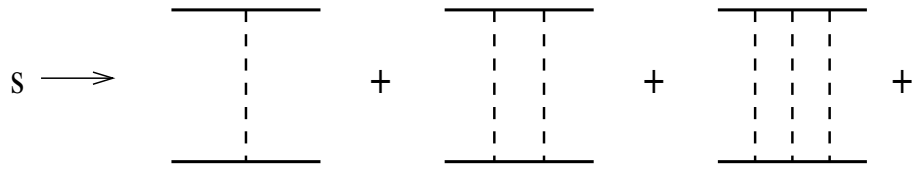


Fig. 17.14 *Amplitudes that determine the high-energy behavior of the scattering process.*

and formalized under the heading of *the principle of maximum analyticity of the S-matrix*. If this hypothesis were correct, the physics of the strong interactions should not have arbitrary constants, except for the fundamental constants of nature, such as the speed of light c, the Planck constant h, and one parameter scale. Consequently, all the strong interaction particles would be composite particles and could be considered on the same footing. This was the basis of the *bootstrap principle*.

All these theoretical developments were deeply influenced by the formalism proposed by Regge to analyze the scattering amplitudes as functions in the complex plane of the angular momentum. In particular, using Regge's theory, it was possible to study elegantly the asymptotic behavior of the amplitudes for large values of s and to give an estimate of the high-energy limit of the cross-sections. Among the results obtained thanks to Regge's theory it is worth mentioning:

1. The prediction of the high-energy asymptotic behavior of the scattering processes dominated by the exchange of particles (with the relative associated poles) in the t-channel, as shown in Fig. 17.14
$$\sigma_{tot} \simeq s^{\alpha_0 - 1}.$$
2. The prediction of the relation between the total cross-section of a process with incoming particles $A + B$ and the cross-sections relative to the incoming particles $A + A$ and $B + B$:
$$\sigma_{tot}^{(A+B)} = \left[\sigma_{tot}^{(A+A)}\, \sigma_{tot}^{(B+B)}\right]^{1/2}.$$
This prediction was based on the close relation between the Regge poles and the resonances, with the factorized expression of the amplitude near a Regge pole
$$f_{nm}(l, s) \simeq \frac{\gamma_n\, \gamma_m}{l - \alpha(s)}.$$

However, the most important result obtained by analytic S-matrix theory was the scattering amplitude discovered by Gabriele Veneziano, which exactly implements the duality between the s- and t-channels. Let's discuss this in more detail. In the presence of particles exchanged in the t-channel, having an increasing values of mass and spin, the amplitude in this channel assumes the form

$$A(s, t) = -\sum_J \frac{g_J^2\, (-s)^J}{t - m_j^2}. \tag{17.A.1}$$

If there is only a *finite* number of these terms, their sum defines an amplitude that does not have poles in the s-channel, since at any fixed value of t, $A(s, t)$ is manifestly

an integer function[12] of s. However, one arrives at a different conclusion if the series is infinite, for it could diverge at different values of s, giving rise then to poles also in the s-channel. In this case, it would not be obvious that to implement the crossing symmetry one shoud also include the corresponding terms of the s-channel, for they could already be present in the series (17.A.1).

Obviously the same conclusion could be reached starting with the s-channel, arriving in this case to an analogous formula

$$\tilde{A}(s,t) \;=\; -\sum_J \frac{g_J^2\,(-t)^J}{s-m_J^2}. \tag{17.A.2}$$

It is now possible to imagine that, with an appropriate choice of the coupling constants g_J and the masses m_J, the two amplitudes $A(s,t)$ and $\tilde{A}(s,t)$ define the same function: if this is the case, the scattering amplitude could be equivalently written as a series on the infinite poles of the t-channel or the s-channel, with an explicit duality between the two pictures. This was explicitly shown by Veneziano with the amplitude

$$A(s,t) \;=\; \frac{\Gamma[-\alpha(s)]\,\Gamma[-\alpha(t)]}{\Gamma[-\alpha(s)-\alpha(t)]}, \quad \alpha(x)=\alpha_0+\alpha'\,x. \tag{17.A.3}$$

From the linear behavior of $\alpha(x)$, it is easy to show that the singularities of the amplitude (17.A.3) are simple poles, corresponding to the exchange of particles of mass $m^2=(n-\alpha_0)/\alpha'$, $n=0,1,2,\ldots$ both in the s- and t-channels. Moreover, the residue at the pole $\alpha(t)=n$ is a polynomial of order n in s, corresponding to a particle of spin n. The same happens for the poles of the s-channel. Using the asymptotic behavior of the function $\Gamma(z)$, it is easy to see that the Veneziano amplitude presents a Regge behavior in both variables

$$\begin{aligned} A(s,t) &\simeq s^{\alpha(t)}, \quad s\to\infty, \quad t \text{ fixed} \\ A(s,t) &\simeq t^{\alpha(s)}, \quad t\to\infty, \quad s \text{ fixed}. \end{aligned}$$

The discovery of the Veneziano amplitude has had an enormous influence on the development of strong interaction studies. Moreover, it has been the starting point for string theory.

The Regge theory and analytic S-matrix theory have dominated theoretical studies for a long time, becoming an extremely sophisticated field, with many subtleties and adjustments, introduced to incorporate in the formalism new phenomena in the strong interaction domain discovered over the years. It was also in fierce competition and often in open opposition with the formulation given of the fundamental interactions by quantum field theory. There were violent polemics among the supporters of the two different formulations, as it was in the past among those who supported the wave or the corpuscular theory of light. The scientific atmosphere of those years is condensed in this humorous story.

[12]We recall that, in Feynman perturbation theory, in order to implement the crossing symmetry one has to include both the diagrams of the s- and t-channels.

A student was curious to know whether the Mandelstam dispersion relation of the scattering amplitude could be derived by quantum field theory. He addressed the question to Weisskopf who answered: "Field theory? What is a field theory?". He went on then to ask the same question to Wigner, who said: "Mandelstam? Who is Mandelstam?". Finally, quite discouraged, the student thought to address the question directly to Chew who, having heard the question, pronounced: "Proof? What is a proof?"

However, despite the initial triumphs, S-matrix theory sank into oblivion, not because it was proved wrong but simply because it was too complicated to handle and many years of study have produced only modest advances. Finally it was supplanted by quantum field theory, which came back into vogue because of the suggestive hints of *deep inelastic scattering processes*. The new quantum chromodynamics theory, a quantum field theory based on a non-abelian gauge group, had the important feature of asymptotic freedom that, in addition to being compatible with all the experimental data, also permits us to make new quantitative predictions.

In light of these historical developments, it is fair to say that the vindication of the basic principles of S-matrix theory comes from the study of two-dimensional statistical models, with the solution of important systems, such as the Ising model in an external magnetic field, which resisted theoretical attempts for many years.

Appendix 17B. Scattering Processes in Quantum Mechanics

In this appendix we recall the main formulas of scattering theory in quantum mechanics. We examine, in particular, one-dimensional systems, i.e. those closer to the S-matrix theory of the $(1+1)$-dimensional systems studied in the text. In the following we impose $\hbar = 1$. Consider initially a particle of mass m and momentum p that moves freely along the real axis, with hamiltonian

$$H_0 \,=\, \frac{p^2}{2m},$$

Since p commutes with H_0, we can simultaneously diagonalize both operators. The common eigenfunctions are the plane waves

$$\begin{aligned} \psi_k(x) \,&=\, e^{ikx} \\ p\,\psi_k(x) \,&=\, k\psi_k(x) \\ H_0\,\psi_k(x) \,&=\, \tfrac{k^2}{2m}\,\psi_k(x). \end{aligned}$$

The time evolution of these eigenfunctions is

$$\psi_k(x,t) \,=\, e^{-iE_k t}\,\psi_k(x) \,=\, e^{-it\,k^2/2m}\,\psi_k(x). \tag{17.B.1}$$

The energy spectrum is continuous and doubly degenerate, since it depends on the square of the momentum. Hence any linear combination of ψ_k and ψ_{-k} is also an

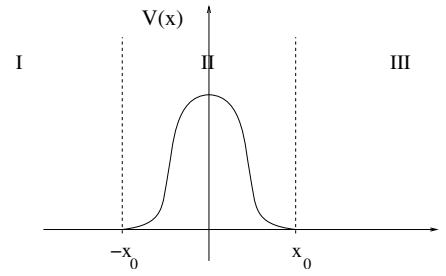


Fig. 17.15 *Potential of the scattering process. In regions I and III the particle moves freely.*

eigenfunction of H_0. H_0 also commutes with the parity operator P and therefore we can choose a basis with functions of a given parity

$$\begin{aligned} \psi_{k0}(x) &= \cos kx, \quad P\,\psi_{k0}(x) = \psi_{k0}(x) \\ \psi_{k1}(x) &= \sin kx, \quad P\,\psi_{k1}(x) = -\psi_{k1}(x). \end{aligned} \tag{17.B.2}$$

Let's imagine now adding to the hamiltonian a potential $V(x)$, finite and different from zero, only inside a region $\mid x \mid < x_0$, as in Fig. 17.15. For simplicity, let's assume that V is an even function, $V(x) = V(-x)$:

$$\begin{aligned} H &= \tfrac{p^2}{2m} + V(x) \\ V(x) &= 0 \quad \text{for} \mid x \mid > x_0. \end{aligned} \tag{17.B.3}$$

The spectrum of the eigenvalues with $E \geq 0$ remains invariant, as well as the eigenfunctions in the external regions I and III

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx}, & x < -x_0 \\ Ce^{ikx} + De^{-ikx}, & x > x_0. \end{cases} \tag{17.B.4}$$

The linear relation that links A and B to the coefficients C and D depends on the shape of the potential $V(x)$.

Consider now the *scattering solutions* $\psi_+(x)$ of the Schrödinger problem, i.e. those with $D = 0$

$$\psi_+(x) = \begin{cases} Ae^{ikx} + Be^{-ikx}, & x < -x_0 \\ Ce^{ikx}, & x > x_0. \end{cases} \tag{17.B.5}$$

In this case, A is the coefficient of the incoming wave, B is the amplitude of the reflected wave, while C is the amplitude of the transmitted wave. The reflection and transmission coefficients are given by

$$\begin{aligned} \mathcal{R} &= \frac{B}{A}, \\ \mathcal{T} &= \frac{C}{A}. \end{aligned} \tag{17.B.6}$$

Since the sum of the densities of the reflected and transmitted waves must be equal to the density of the incoming wave, we have

$$|\mathcal{R}|^2 + |\mathcal{T}|^2 \,=\, 1. \tag{17.B.7}$$

The reflection and transmission coefficients can be expressed in terms of the *phase shifts* δ_0 and δ_1, defined by the stationary eigenfunctions of the hamiltonian

$$\begin{array}{llll} \psi_0 \,=\, \cos(kx+\delta_0) & (x>x_0) & ;\psi_0 \,=\, \cos(kx-\delta_0) & (x<-x_0) \\ \psi_1 \,=\, \sin(kx+\delta_1) & (x>x_0) & ;\psi_1 \,=\, \sin(kx-\delta_1) & (x<-x_0). \end{array} \tag{17.B.8}$$

The S-matrix in the channels of a given parity is given by

$$S_a \,=\, e^{2i\delta_a}, \quad a=0,1. \tag{17.B.9}$$

The linear combination of eigenstates of given parity (17.B.8) that gives rise to the scattering eigenfunction ψ_+ is

$$\psi_+ \,=\, \begin{cases} e^{i\delta_0}\,\psi_0 + i\,e^{i\delta_1}\,\psi_1 \,=\, \frac{1}{2}(e^{2i\delta_0}+e^{2i\delta_1})\,e^{ikx} & (x>x_0) \\ e^{ikx} + \frac{1}{2}(e^{2i\delta_0}-e^{2i\delta_1})\,e^{-ikx} & (x<-x_0) \end{cases}. \tag{17.B.10}$$

Hence

$$\begin{aligned} \mathcal{R} &= \tfrac{1}{2}(e^{2i\delta_0}-e^{2i\delta_1}) = \tfrac{1}{2}[(e^{2i\delta_0}-1)-(e^{2i\delta_1})-1)] \\ &= \textstyle\sum_{l=0}^{1} i(-1)^l\,e^{i\delta_l}\,\sin\delta_l \\ \mathcal{T} &= \tfrac{1}{2}(e^{2i\delta_0}+e^{2i\delta_1}) = \tfrac{1}{2}[(e^{2i\delta_0}-1)+(e^{2i\delta_1})-1)]+1 \\ &= 1+\textstyle\sum_{l=0}^{1} ie^{i\delta_l}\,\sin\delta_l \end{aligned} \tag{17.B.11}$$

and the reflection and transmission coefficients are completely determined by the phase shifts of the even and odd eigenfunctions.

Consider now the case in which the potential is given by

$$V(x) \,=\, -2\,g\,\delta(x). \tag{17.B.12}$$

Let's start from the even eigenfunctions. Imposing the continuity of the wavefunction at the origin and the discontinuity of its derivative, ruled by the $\delta(x)$ function

$$\begin{aligned} &\psi_0(0^+) \,=\, \psi(0^-) \\ &\frac{d\psi_0(0^+)}{dx} - \frac{d\psi_0(0^-)}{dx} \,=\, -2\,k\,\sin\delta_0 \,=\, -2\,g\,\psi_0(0) \,=\, -g\,\cos\delta_0 \end{aligned}$$

we can determine the even phase shift

$$\tan\,\delta_0 \,=\, \frac{g}{k}. \tag{17.B.13}$$

The S-matrix in this channel is then

$$e^{2i\delta_0} \,=\, \frac{1+i\tan\delta_0}{1-i\tan\delta_0} \,=\, \frac{k+ig}{k-ig}. \tag{17.B.14}$$

The variation of the phase is then

$$\delta_0(+\infty)-\delta_0(-\infty) \,=\, -2\pi g/|g|,$$

and depends on the sign of g.

The odd solution vanishes at the origin, hence the odd phase shift is identically zero. The corresponding S-matrix is then equal to 1:

$$\begin{aligned} \delta_1 &= 0 \\ e^{2i\delta_1} &= 1 \end{aligned} \tag{17.B.15}$$

The expressions δ_0 and δ_1 permit us to obtain the ratios (17.B.7) and to define a solution of the Schrödinger equation for all values of k.

It is interesting to analyze the nature of this solution for complex values of the momentum $k = k_1 + ik_2$. The real part can always be considered positive or zero since it corresponds to the physical momentum of the incoming particle. Substituting k in (17.B.5) one sees that the imaginary part k_2 enters the real part of the exponentials. Choosing now k as the value of the pole of the S-matrix, i.e. $k = ig$, one can have a normalizable eigenfunction by imposing $A = 0$. This solution corresponds to a *bound state* of the system, whose energy is $E_b = -g^2/(2m)$. Obviously in this case we should have $g > 0$.

More generally, one can show the following properties of the non-relativistic S-matrix:

1. The poles of the S-matrix with positive imaginary values of the momentum, $k_n = ia_n$ $(a_n > 0)$ correspond to the energies $E_n = -a_n^2/(2m)$ of the bound states of the system.
2. There are no poles in the complex plane of the variable $k = k_1 + ik_2$ with a non-vanishing real part k_1 in the half-plane $k_2 > 0$.
3. The poles in the complex plane with negative imaginary part, $k_2 < 0$, correspond instead to the resonances.

The proof of property (1) follows that given for the potential $\delta(x)$. For point (2), let's suppose that the S-matrix has a pole at $k = k_1 + ik_2$, with $k_2 > 0$. Substituting in (17.B.5) and putting to zero the coefficient A, we have also in this case a normalizable eigenfunction. The problem, though, is in the time evolution of this eigenfunction: using eqn (17.B.1) one gets

$$\psi_+(x,t) = e^{-it(k_1^2-k_2^2)/2m}\, e^{tk_1\,k_2/m}\, \psi_+(x) \tag{17.B.16}$$

and, if $k_1 > 0$, the eigenfunction grows exponentially when $t \to +\infty$, leading to a violation of the conservation of probability.

A pole in the complex plane but with negative imaginary part is however perfectly plausible. It corresponds to a solution whose probability decreases in a given channel. This means that it will grow in another channel so that there is a global conservation of the probability. Poles with negative imaginary part correspond to *resonances*. The situation in the plane of the complex variables k and E is shown in Fig. 17.16.

Since the S-matrix in any channel of a given parity is a unitary operator, in the vicinity of a pole $\bar{k}$ it can be parameterized as

$$S = e^{2i\delta} = \frac{k - \bar{k}^*}{k - \bar{k}}, \tag{17.B.17}$$

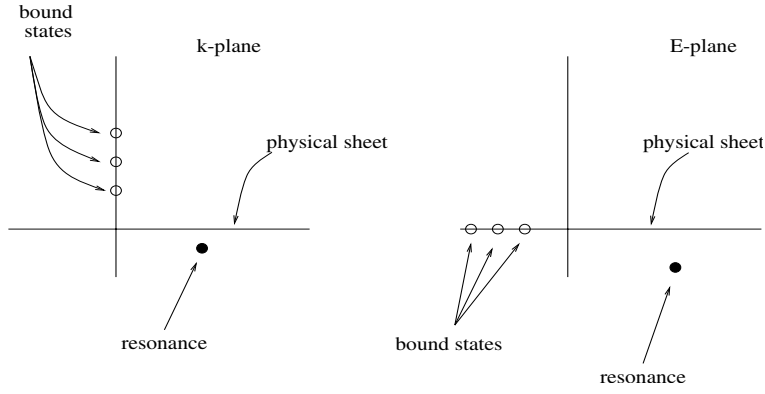


Fig. 17.16 *Analytic structure of the S-matrix in the planes of the complex variables k and E.*

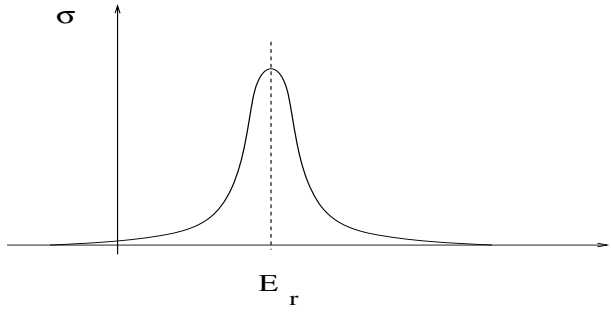


Fig. 17.17 *Cross-section relative to an S-matrix with a resonance pole.*

where $\bar{k}^*$ is the complex conjugate of $\bar{k}$. Changing to the energy $E = E_r - i\Gamma/2$ (with $\Gamma > 0$, since there could be no poles in the upper half-plane), we have

$$S = \frac{E - E_r - i\Gamma/2}{E - E_r + i\Gamma/2}. \tag{17.B.18}$$

Note that, close to the energy of the resonance, the phase $\delta(E)$ of the S-matrix has an abrupt jump of 2π. We can now compute the diffusion amplitude T, defined by $S = 1 + iT$

$$T = -\frac{\Gamma}{E - E_r + i\Gamma/2}. \tag{17.B.19}$$

and the cross-section

$$\sigma \sim |T|^2 = \frac{\Gamma^2}{(E - E_r)^2 + \Gamma^2/4}. \tag{17.B.20}$$

As shown in Fig. 17.17, the cross-section has the typical bell shape of a resonance phenomenon, with the width determined by the parameter Γ. It is easy to see that this is related to the life-time τ of the resonance state given by $\tau = 1/\Gamma$.

Appendix 17C. n-particle Phase Space

An important quantity that enters the probability computation of the scattering and decay processes is the differential n-particle *phase space*

$$d\Phi_n \;=\; \frac{d^{d-1}p_1}{(2\pi)^{d-1}2E_1}\cdots\frac{d^{d-1}p_1}{(2\pi)^{d-1}2E_1}\,(2\pi)^d\,\delta^d(P-p_1-p_2-\cdots-p_n). \qquad (17.C.1)$$

The integral of this expression is a relativistic invariant quantity that depends only on the modulus of the total momentum, i.e. P^2:

$$\Phi_n(P^2) \;=\; \int\frac{d^{d-1}p_1}{(2\pi)^{d-1}2E_1}\cdots\int\frac{d^{d-1}p_n}{(2\pi)^{d-1}2E_n}\,(2\pi)^d\,\delta^d(P-p_1-p_2-\cdots-p_n). \quad (17.C.2)$$

This quantity has an analog in statistical mechanics. In fact, its definition recalls the partition function of a statistical model in the microcanonical ensemble, the role of the total energy being played here by P^2. For the delta function that involves all momenta, its exact computation can be done only in a few cases or for particular values of P^2.

Two-particle phase space. Let's study in more detail the properties of $\Phi_n(P^2)$, starting with the computation of the two-particle phase space when the momentum P is time-like (which is the more relavant case). This is the only case in which the phase case can be computed exactly. Since Φ_2 is a relativistic invariant quantity, we can choose a reference frame where $P=(E,0)$ and

$$\begin{aligned}\Phi_2(E) \;&=\; \int\frac{d^{d-1}p_1}{(2\pi)^{d-1}2E_1}\int\frac{d^{d-1}p_2}{(2\pi)^{d-1}2E_2}\,(2\pi)^d\,\delta^{d-1}(\vec{p}_1-\vec{p}_2)\,\delta(E-E_1-E_2)\\ &=\; \frac{\Omega(d-1)}{4(2\pi)^{d-2}}\int_0^\infty dp\,\frac{p^{d-2}}{\sqrt{(p^2+m_1^2)(p^2+m_2^2)}}\,\delta(E-\sqrt{p^2+m_1^2}-\sqrt{p^2+m_2^2})\\ &=\; \frac{1}{2^{d-1}\,\pi^{\frac{d-3}{2}}\Gamma(\frac{d-1}{2})}\,\frac{|\,p_{cm}\,|^{d-3}}{E_{cm}}\,\Theta(E-(m_1+m_2)), \qquad (17.C.3)\end{aligned}$$

where

$$\Theta(x) \;=\; \begin{cases}1\,,\ \text{if}\ \ x>0\\ 0\,,\ \text{if}\ \ x<0\end{cases}$$

and $|\,p_{cm}\,|$ is the modulus of the space component of the momentum in the reference frame of the center of mass, corresponding to the energy E

$$|\,p_{cm}\,| = \frac{1}{2E_{cm}}\,\sqrt{[E^2-(m_1+m_2)^2][E^2-(m_1-m_2)^2]}. \qquad (17.C.4)$$

To arrive at (17.C.3), we used the expression (2.6.3) of the solid angle in $(d-1)$ dimensions.

Recursive equation. The explicit computation of the phase space with a higher number of particles cannot be done exactly. However, its numerical determination can be reached by means of the recursive equation

$$\Phi_n(P^2) \;=\; \int\frac{d^{d-1}p_n}{(2\pi)^{d-1}2E_n}\,\Phi_{n-1}(P-p_n). \qquad (17.C.5)$$

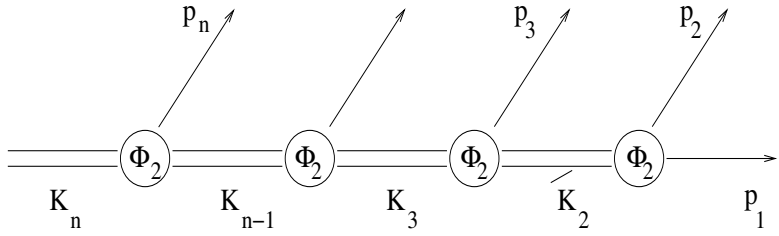


Fig. 17.18 *Iteration of the recursive equation for n-particle phase space.*

By iteration, this formula leads to integrals that involve the two-particle phase space, as shown in Fig. 17.18. The proof of (17.C.5) is immediate. From its definition we have

$$\begin{aligned}\Phi_n(P^2) &= (2\pi)^d \int \prod_{i=1}^{n} \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}\,\delta(P-p_1-\cdots-p_n)\\ &= \int \frac{d^{d-1}p_n}{(2\pi)^{d-1}2E_n}\left[(2\pi)^d \int \prod_{i=1}^{n-1} \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}\,\delta((P-p_n)-p_1\cdots-p_{n-1})\right]\\ &= \int \frac{d^{d-1}p_n}{(2\pi)^{d-1}2E_n}\,\Phi_{n-1}(P-p_n), \qquad (17.C.6)\end{aligned}$$

where $\Phi_n(P)$ is a function of $P^2 \equiv M_n^2$. Analogously $\Phi_{n-1}(P-p_n)$ is function of

$$(P-p_n)^2 = (p_1+\cdots p_{n-1})^2 \equiv K_{n-1}^2 \equiv M_{n-1}^2,$$

where M_{n-1}^2 is the square of the invariant mass of the system of particles $1, 2, \ldots,$ $(n-1)$. Since Φ_{n-1} is a function only of this last variable, it is convenient to write eqn (17.C.5) using the identity

$$\begin{aligned}1 &= \int dM_{n-1}^2 \delta(M_{n-1}^2 - K_{n-1}^2),\\ 1 &= \int d^d K_{n-1}\delta^d(P-p_n-K_{n-1}).\end{aligned}$$

Hence

$$\begin{aligned}\Phi_n(M_n^2) &= \int dM_{n-1}^2 \delta(K_{n-1}^2 - M_{n-1}^2) \int d^d K_{n-1}\delta^d(P-p_n-K_{n-1})\\ &\quad\times \int \frac{d^d p_n}{(2\pi)^{d-1}}\,\delta(p_n^2-m_n^2)\,\Phi_{n-1}(M_{n-1}^2) \qquad (17.C.7)\\ &= \frac{1}{2\pi}\int_{\mu_{n-1}^2}^{(M_n-m_n)^2} dM_{n-1}^2\,\Phi_2(M_n^2;K_{n-1},p_n)\,\Phi_{n-1}(M_{n-1}^2),\end{aligned}$$

where

$$\mu_i \equiv m_1 + m_2 + \cdots m_i.$$

$\Phi_2(M_n^2; K_{n-1}, p_n)$ is the two-particle phase space of total momentum $P^2 = M_n^2$, relative to the masses of the momenta K_{n-1} and p_n given by (17.C.3).

Laplace transform. It is useful to make use of the Laplace transform to solve the constraint on the momenta given by the delta function. Define

$$\Phi_n(\alpha) \,=\, \int d^dP\, e^{-\alpha\cdot P}\, \Phi_n(P^2), \tag{17.C.8}$$

where α is a Lorentz time-like vector $\alpha = (\alpha_0, \vec{\alpha})$, with $\alpha_\mu \alpha^\mu > 0$. Thanks to this transformation we have

$$\begin{aligned}\Phi_n(\alpha) &= (2\pi)^d \int d^dP \prod_{i=1}^n \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}\, \delta^d(P - p_1 - \cdots - p_n)\, e^{-\alpha\cdot P} \\ &= (2\pi)^d \prod_{i=1}^n \int \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}\, e^{-\alpha\cdot p_i} \,=\, (2\pi)^d \prod_{i=1}^n \phi_i(\alpha).\end{aligned} \tag{17.C.9}$$

The functions $\phi_i(\alpha)$ can be easily computed choosing the reference frame where $\alpha = (\beta, 0)$ and computing the integral using spherical coordinates. In fact we have

$$\begin{aligned}\phi_i(\alpha) &= \int \frac{d^{d-1}p_i}{(2\pi)^{d-1}2E_i}\, e^{-\alpha\cdot p_i} \,=\, \frac{\Omega(d-1)}{(2\pi)^{d-1}} \int_0^\infty dp\, \frac{p^{d-2}}{2E}\, e^{-\beta E} \\ &= \frac{\Omega(d-1)}{(2\pi)^{d-1}} \int_{m_i^2}^\infty dE\, (E^2 - m_i^2)^{\frac{d-3}{2}}\, e^{-\beta E},\end{aligned}$$

where $\Omega(d-1)$ is the solid angle in $(d-1)$ dimensions, given by eqn (2.6.3). The last integral can be expressed in terms of the Bessel function $K_\nu(z)$, whose integral representation is

$$K_\nu(z) \,=\, \frac{\left(\frac{z}{2}\right)^\nu \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\nu + \frac{1}{2}\right)} \int_1^\infty e^{-zt}\, (t^2-1)^{\nu-\frac{1}{2}}\, dt. \tag{17.C.10}$$

So, we have

$$\phi_1(\beta) \,=\, \frac{2}{(2\pi)^{\frac{d}{2}}} \left(\frac{m}{\beta}\right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(\beta m). \tag{17.C.11}$$

On the other hand, in the reference frame where $\alpha = (\beta, 0)$, eqn (17.C.8) can also be expressed as

$$\begin{aligned}\Phi_n(\beta) &= \int d^dp\, e^{-\beta E}\, \Phi_n(p^2) \\ &= \int ds \int d^dp\, \delta(p^2 - s)\, e^{-\beta E}\, \Phi_n(s) \\ &= \int ds \int \frac{d^{d-1}p}{2E}\, e^{-\beta E}\, \Phi_n(s) \\ &= \Omega(d-1) \int ds \int_{\sqrt{s}}^\infty dE\, (E^2 - s)^{\frac{d-3}{2}}\, e^{-\beta E}\, \Phi_n(s).\end{aligned} \tag{17.C.12}$$

Using also in this case the integral representation (17.C.10) and the $(d-1)$-dimensional solid angle, the last expression can be written as

$$\Phi_n(\beta) \;=\; \frac{(2\pi)^{\frac{d}{2}}}{\pi}\,\frac{1}{\beta^{\frac{d-2}{2}}}\,\int_0^\infty ds\, s^{\frac{d-2}{4}}\, K_{\frac{d-2}{2}}(\beta\sqrt{s})\,\Phi_n(s). \tag{17.C.13}$$

Hence we have the identity

$$\frac{(2\pi)^{\frac{d}{2}}}{\pi}\,\frac{1}{\beta^{\frac{d-2}{2}}}\,\int_0^\infty ds\, s^{\frac{d-2}{4}}\, K_{\frac{d-2}{2}}(\beta\sqrt{s})\,\Phi_n(s) \;=\; (2\pi)^d \prod_{i=1}^n \frac{2}{(2\pi)^{\frac{d}{2}}}\,\left(\frac{m_i}{\beta}\right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(\beta m_i). \tag{17.C.14}$$

Phase space at the threshold. Let's use eqn (17.C.14) in the limit $\beta \to \infty$ to estimate the behavior of $\Phi_n(s)$ near the threshold energy $\sqrt{s} = \sum_i^n m_i$. Using the asymptotic behavior of the Bessel function

$$K_\nu(z) \simeq \left(\frac{\pi}{2z}\right)^{\frac{1}{2}}\, e^{-z} \qquad z \to \infty$$

substituting this expression in (17.C.14) and simplifying, we have

$$\frac{1}{(2\pi)^{\frac{d+1}{2}}}\,\frac{1}{\beta^{\frac{d-1}{2}}}\,\int_0^\infty ds\, s^{\frac{d-3}{4}}\; e^{-\beta\sqrt{s}}\,\Phi_n(s) \;=\; \prod_{i=1}^n \frac{1}{(2\pi)^{\frac{d-1}{2}}}\,\frac{m_i^{\frac{d-3}{2}}}{\beta^{\frac{d-2}{2}}}\, e^{-\beta m_i}. \tag{17.C.15}$$

With the change of variable $E = \sqrt{s}$, eqn (17.C.15) becomes

$$\int_0^\infty dE\, E^{\frac{d-1}{2}}\; e^{-\beta E}\,\Phi_n(E) \;=\; A_n\,\beta^{\frac{1-n}{2}(d-1)}\, e^{-\beta\sum_i m_i},$$

with

$$A_n \;=\; \frac{1}{2}(2\pi)^{\frac{1}{2}[d+1-n(d-1)]}\,\left(\prod_{i=1}^n m_i\right)^{\frac{d-3}{2}}.$$

Using the general properties of the Laplace transform $\mathcal{L}$

$$\mathcal{L}[F(s-a)] \;=\; e^{-a\beta}\mathcal{F}(\beta), \qquad \mathcal{L}[x^\nu] \;=\; \frac{\Gamma(\nu+1)}{\beta^{\nu+1}},$$

where $\mathcal{F}$ denotes the Laplace transform $\mathcal{L}$ of the function $F(s)$, it is easy to see that for $E \to \sum_i m_i$, the n-particle phase space goes to zero as

$$\Phi_n(E) \;\simeq\; \frac{A_n}{\left(\sum_i m_i\right)^{\frac{d+1}{2}}}\,\frac{1}{\Gamma\left(\frac{1}{2}(d-1)(n-1)\right)}\,(E-\sum_i m_i)^{\frac{(n-1)(d-1)-2)}{2}}. \tag{17.C.16}$$

Phase space at high energy. Let's use now the formula (17.C.14) to study the behavior of the n-particle phase space for $m_i \to 0$, i.e. in the massless limit or equivalently at high energy. In this case it is necessary to distinguish two cases: (a) $d \neq 2$

and (b) $d = 2$. Let's consider the first case. Using the series expansion of the Bessel function $K_\nu(x)$ for $\nu \neq 0$

$$K_\nu(x) \simeq \frac{1}{x^\nu}, \quad x \to 0$$

the mass terms in the right-hand side of eqn (17.C.14) simplify and we have

$$\frac{(2\pi)^{\frac{d}{2}}}{\pi} \frac{1}{\beta^{\frac{d-2}{2}}} \int_0^\infty ds\, s^{\frac{d-2}{4}} K_{\frac{d-2}{2}}(\beta\sqrt{s})\, \Phi_n(s) = (2\pi)^d \frac{2^n}{(2\pi)^{\frac{nd}{2}}} \frac{1}{\beta^{n(d-2)}}. \tag{17.C.17}$$

With the change of variable $E = \sqrt{s}$ in the integral on the left and collecting terms we have

$$\int_0^\infty dE\, E^{\frac{d}{2}} K_{\frac{d-2}{2}}(\beta E)\, \Phi_n(E) = \pi \frac{2^{n-1}}{(2\pi)^{\frac{(n-1)d}{2}}} \left(\frac{1}{\beta}\right)^{(2n-1)(\frac{d-2}{2})}. \tag{17.C.18}$$

Since

$$\int_0^\infty x^\mu K_\nu(ax)\, dx = 2^{\mu-1} a^{-\mu-1} \Gamma\left(\frac{1+\mu+\nu}{2}\right) \Gamma\left(\frac{1+\mu-\nu}{2}\right),$$

the n-particle phase space behaves for $E \to \infty$ and $d \neq 2$ as

$$\Phi_n(E) \simeq B_n\, E^{n(d-2)-d}, \tag{17.C.19}$$

where

$$B_n = \pi \frac{2^{n(3-d)+\frac{d}{2}-1}}{(2\pi)^{(n-1)\frac{d}{2}}} \frac{1}{\Gamma\left(\frac{n(d-2)}{2}\right) \Gamma\left(\frac{(n-1)(d-2)}{2}\right)}.$$

Let's now consider the behavior of the n-particle phase space for large values of the energy when $d = 2$. For dimensional reasons we expect that it scales as

$$\Phi_n(s) \simeq \frac{1}{s}, \quad s \to \infty$$

but there could be logarithmic corrections. On the basis of the cases $n = 2$ and $n = 3$, let's impose the ansatz

$$\Phi_n(s) \simeq \alpha_n \frac{1}{s} \left(\ln \frac{s}{m^2}\right)^{n-2}, \tag{17.C.20}$$

where m is a mass scale whereas α_n is a constant to be determined. The presence of the logarithms does not allow us to follow the previous computation, where we set to zero all the masses. Consider now the recursive equation (17.C.7) in the limit $M_n^2 \to \infty$

$$\Phi_n(M_n^2) \simeq \frac{1}{2\pi} \int_\epsilon^{M_n^2} \Phi_2(M_n^2; K_{n-1}, p_n)\, \Phi_{n-1}(M_{n-1}^2),$$

where ϵ is a small but non-zero quantity. Substituting in this formula the expression of the two-particle phase space and the ansatz (17.C.20), we have

$$\Phi_n(M_n^2) \simeq \frac{1}{2\pi}\,\alpha_{n-1}\int_\epsilon^{M_n^2}\frac{1}{(M_n^2-M_{n-1}^2)}\,\frac{1}{M_{n-1}^2}\,\left(\ln\frac{M_{n-1}^2}{m^2}\right)^{n-3}$$
$$= \frac{1}{2\pi M_n^2}\,\alpha_{n-1}\int_\epsilon^{M_n^2} dM_{n-1}^2\left[\frac{1}{M_{n-1}^2}-\frac{1}{M_n^2-M_{n-1}^2}\right]\left(\ln\frac{M_{n-1}^2}{m^2}\right)^{n-3}.$$

The first term of this equation is responsible for the most singular part and keeping only this, one has

$$\Phi_n(M_n^2) \simeq \frac{\alpha_{n-1}}{2\pi}\,\frac{1}{n-2}\,\frac{1}{M_n^2}\left(\ln\frac{M_n^2}{m^2}\right)^{n-2}. \tag{17.C.21}$$

Comparing this expression with the ansatz (17.C.20), we obtain the recursive equation for the constants α_n

$$\alpha_n \;=\; \frac{1}{2\pi(n-2)}\,\alpha_{n-1},$$

whose solution is

$$\alpha_n \;=\; \frac{1}{(2\pi)^{n-2}}\,\frac{1}{(n-2)!}.$$

Hence, in $d=2$, the asymptotic expression of the n–particle phase space for $s\to\infty$ is

$$\Phi_n(s) \;\simeq\; \frac{1}{(2\pi)^{n-2}}\,\frac{1}{(n-2)!}\,\frac{1}{s}\,\left(\ln\frac{s}{m^2}\right)^{n-2}. \tag{17.C.22}$$

References and Further Reading

For a discussion of the general principles of S-matrix theory, it is useful to consult:

L.D. Landau, E.M. Lifshitz, *Quantum Mechanics*, Pergamon Press, Oxford, 1977.

L.D. Landau, E.M. Lifshitz, L.P. Pitaevskii, *Relativistic Quantum Mechanics*, Pergamon Press, Oxford, 1974.

R.J. Eden, P.V. Landshoff, D.I. Olive, J.C. Polkinghorne, *The Analytic S-Matrix*, Cambridge University Press, Cambridge, 1966.

J.R. Taylor, *Scattering Theory*, John Wiley, New York, 1980.

H. Lipkin, *Quantum Mechanics. New Approaches to Selected Topics*, North Holland, Amsterdam, 1978.

The famous article by Coleman and Mandula quoted in the text about the triviality of the S-matrix in $d>2$ in the presence of conserved charges is:

S. Coleman, J. Mandula, *All possible symmetries of the S-matrix*, Phys. Rev. 159 (1967), 1251.

Fundamental papers for the development of S-matrix theory in $(1+1)$-dimensional systems are:

A.B. Zamolodchikov, Al.B. Zamolodchikov, *Factorized S-matrices in two dimensions as the exact solutions of certain relativistic quantum field theories*, Ann. Phys. 120 (1979), 253.

A.B. Zamolodchikov, *Integrable field theory from conformal field theory*, Adv. Stud. Pure Math., 19 (1989), 641.

For non-local conserved currents in two-dimensional quantum field theories see:

D. Bernard and A. Leclair, *Quantum group symmetries and non-local currents in 2-D QFT*, Comm. Math. Phys. 142 (1991), 99.

For further details on the historical developments of S-matrix theory and the crossing of paths with quantum field theory, it is useful to read the papers:

S. Weinberg, *Particle physics: Past and future*, Int. J. Mod. Phys. A 1 (1986), 135.

A.R. White, *The Past and Future of S-Matrix Theory*, in *Scattering*, Academic Press, San Diego, 2002 editors E.R. Pike and P. Sabatier.

A book on bootstrap interactions that has become a classic of popular scientific books is:

F. Capra, *The Tau of Physics*, Bantam Books, Torento, 1984.

String theory has been brought to the attention of the public by the book:

B. Green, *The Elegant Universe*, Norton, New York, 1999.

Professional books in which the reader can deepen the knowledge on string theory are:

M. Green, J.H. Schwarz, E. Witten, *String Theory*, Cambridge University Press, Cambridge, 2000.

J. Polchinski, *String Theory*, Cambridge University Press, Cambridge, 2002.

Problems

1. Causality and analyticity

Consider a linear system in which the output $b(t)$ depends on the input $a(t)$ as

$$b(t) \;=\; \int_{-\infty}^{t} G(t-t')\; a(t')\, dt'.$$

If the system is causal, the Green function $G(t-t')$ vanishes when $t < t'$. Let

$$\hat{G}(\omega) \;=\; \int_{-\infty}^{\infty} e^{i\omega\tau}\, G(\tau)\, d\tau \;=\; \int_{0}^{\infty} e^{i\omega\tau}\, G(\tau)\, d\tau$$

be its Fourier transform. If $a(t)$ and $b(t)$ are both real, also $G(\tau)$ is a real function and

$$\hat{G}^*(\omega) \;=\; \hat{G}(-\omega^*).$$

a Show that, if $G(\tau)$ is a square integrable function, then $\hat{G}(\omega)$ is an analytic function in the upper half-plane $\operatorname{Im}\omega > 0$. This implies that $\hat{G}(\omega)$, for real ω, is a function obtained as a boundary value of an analytic function.

b Letting $\hat{G}(\omega) = \hat{G}_1(\omega) + i\hat{G}_2(\omega)$, use the Cauchy theorem to prove that these functions are related one to the other by the dispersion relations

$$\hat{G}_1(\omega) = \frac{1}{\pi}\,\mathcal{P}\int_{-\infty}^{+\infty} \frac{1}{\nu-\omega}\,\hat{G}_2(\nu)\, d\nu$$

$$\hat{G}_2(\omega) = -\frac{1}{\pi}\,\mathcal{P}\int_{-\infty}^{+\infty} \frac{1}{\nu-\omega}\,\hat{G}_1(\nu)\, d\nu$$

where $\mathcal{P}$ denotes the principal part of the integral.

2. Decay process

A particle of mass M and four-dimensional momentum P decays into two particles of masses m_1 and m_2.

a Use the conservation of energy and momentum to prove that the total energy of the first particle in the center-of-mass reference frame is

$$E_1 \;=\; \frac{M^2 + m_1^2 - m_2^2}{2M}$$

and that E_2 is obtained from the previous expression exchanging m_1 with m_2.

b Show that the kinetic energy T_i of the particle i, in the same reference frame, is given by

$$T_i \;=\; \Delta M\left(1 - \frac{m_i}{M} - \frac{\Delta M}{2M}\right)$$

where $\Delta M = M - m_1 - m_2$.

3. Physical region of the amplitudes

Determine the physical region of the s-channel process when the masses of the particles are different.

4. Yang–Baxter equations

Prove that the Yang–Baxter equations given in eqn (17.3.7) of the text can be obtained as a consequence of the associativity condition of the Faddev–Zamolodchikov algebra.

5. Reflection amplitude

Consider the following scattering amplitudes of a particle A and its antiparticle $\overline{A}$:

$$
\begin{aligned}
&| A(\theta_1)\, A(\theta_2)\rangle \;=\; S(\theta)\;\, | A(\theta_2)\, A(\theta_1)\rangle, \\
&| A(\theta_1)\, \overline{A}(\theta_2)\rangle \;=\; t(\theta)\;\, | \overline{A}(\theta_2)\, A(\theta_1)\rangle + r(\theta)\;\, | A(\theta_2)\, \overline{A}(\theta_1)\rangle.
\end{aligned}
$$

a Prove that

$$
\begin{aligned}
&S(\theta)\, S(-\theta) = t(\theta)\, t(-\theta) + r(\theta)\, r(-\theta) \;=\; 1 \\
&t(\theta)\, r(-\theta) + r(\theta)\, t(-\theta) \;=\; 0 \\
&t(\theta) \;=\; S(i\pi - \theta), \quad r(\theta) \;=\; r(i\pi - \theta).
\end{aligned}
$$

b Prove that if the particles A and $\overline{A}$ are uniquely distinguishable by their different eigenvalues of the conserved charges, then the reflection amplitude vanishes, i.e. $r(\theta) = 0$.

6. Bootstrap equations

Derive the bootstrap equations (17.4.12) imposing the commutativity of the processes shown in Fig. 17.13.

Hint. Note that the line of the particle A_i in the second graph is parallel to the same line of the first graph. Identify the angles in the two figures and use the resonance condition.

7. Scattering in a potential with two delta functions

Consider a one-dimensional system of quantum mechanics with hamiltonian given by

$$
\mathcal{H} \;=\; \frac{p^2}{2m} + V(x)
$$

with

$$
V(x) \;=\; -g_1\, \delta(x + a) - g_2\, \delta(x + a)
$$

(g_1 and g_2 positive).

a Compute the phase shifts δ_0 and δ_1 and the corresponding S-matrix elements.

b Analyze the analytic structure of the S-matrix by varying the momentum k.

c Determine the wavefunction of the bound states.

8. Interpretation of the two-dimensional S-matrix

The non-relativistic S-matrix of a particle of mass $m = 1$ relative to the potential $V(x) = -2a\pi\delta(x)$ is given by

$$
\tilde{S}(k) \;=\; \frac{k + i\pi a}{k - i\pi a}.
$$

If we would like to generalize this result to the relativistic case, we must use the rapidity variable θ. Note that for small values of the momentum, $\theta \simeq k$. Substituting in the expression of S, we have

$$\tilde{S}(\theta) = \frac{\theta + i\pi a}{\theta - i\pi a}.$$

This expression, however, does not fulfill the important property $S(\theta) = S(\theta \pm 2\pi i)$ of the relativistic S-matrix.

a Discuss how the periodicity of the relativistic S-matrix can be iteratively implemented starting from $\tilde{S}(\theta)$.

b Use the infinite product representation of the hyperbolic function $\sinh x$

$$\sinh x = x \prod_{k=1}^{\infty} \left[1 + \left(\frac{x}{k\pi}\right)^2\right],$$

to show that the final result can be expressed as

$$S(\theta) = \frac{\sinh \frac{1}{2}(\theta + i\pi a)}{\sinh \frac{1}{2}(\theta - i\pi a)} = s_a(\theta).$$

18
Exact S-Matrices

The particles are nothing else than lumps of energy, they come and go, their own identity is all in this dance of creation and annihilation processes.

Kenneth Ford

In this chapter we present the exact S-matrix associated to several two-dimensional statistical models away from their critical point. Closing the bootstrap procedure, one is able to find at the same time the set of all amplitudes and the mass spectrum of the theory. A crucial step in the determination of the scattering amplitudes is the knowledge of the spectrum of the spins relative to the conserved currents. In the first sections we address the minimal S-matrix of several off-critical statistical systems. Many of these examples are related to the Toda field theories previously discussed. Later we use the minimal S-matrices of the statistical models to determine the exact S-matrices of the lagrangian Toda field theories. The scattering amplitudes of the Toda field theories shows an important symmetry of these models under the weak–strong duality transformation $g \to 8\pi/g$ of their coupling constant. At the end of the chapter we discuss the exact S-matrix of the Sine–Gordon model, with a series of comments on its interesting features, and the quantum group reductions which lead to the general S-matrices of integrable deformations of conformal field theories.

18.1 Yang–Lee and Bullogh–Dodd Models

The conformal field theory associated to the Yang–Lee edge singularity is non-unitary, with central charge $c = -22/5$ and only one relevant field ϕ with conformal weight $\Delta = -1/5$. As discussed in Section 14.5, this theory describes the critical behavior of an Ising model in a purely imaginary magnetic field ih. The Landau–Ginzburg lagrangian is given by

$$\mathcal{L} = \int \left[\frac{1}{2}(\partial\phi)^2 - i(h - h_c)\phi - ig\phi^3 \right] d^2x \tag{18.1.1}$$

and the scaling region near the critical point is a one-dimensional space, spanned by the (purely imaginary) coupling constant of the relevant field ϕ. We can use the characters of the identity family and of the field ϕ to count the dimensions of the quasi-primary fields at level n, as shown for the first representatives in Table 18.1. We can then apply

Table 18.1: *Dimensions of the spaces $\hat{\Lambda}_{n+1}$ and $\hat{\phi}_n$. For each value of n for which the former is larger than the latter there must exist a conserved current.*

n	1	2	3	4	**5**	6	**7**	8	9	10	**11**	12	**13**	14	15	16	**17**
$\hat{\Lambda}_{n+1}$	1	0	0	0	1	0	1	0	1	0	2	0	2	1	2	1	3
$\hat{\phi}_n$	0	0	0	1	0	1	0	1	1	1	1	2	1	2	2	3	2

the counting method (see Section 16.8.3) to establish that the off-critical system has conserved charges with spin

$$s = 1, 5, 7, 11, 13, 17, 19, 23. \tag{18.1.2}$$

The sequence of these spins is made of odd numbers not divisible by 3 and is therefore compatible with the existence of a massive excitation associated to a particle A that is the bound state of itself. Hence, its exact S-matrix must have a pole at $\theta = 2i\pi/3$. The crossing symmetry helps in fixing the position of the pole in the t-channel at $\theta = i\pi/3$. Assuming that there are no additional poles, the only solution of the bootstrap equation

$$S_{AA}(\theta) \;=\; S_{AA}\left(\theta - \frac{i\pi}{3}\right) S_{AA}\left(\theta + \frac{i\pi}{3}\right) \tag{18.1.3}$$

is given by

$$S_{AA} \;=\; \frac{\tanh\frac{1}{2}(\theta + i\frac{2\pi}{3})}{\tanh\frac{1}{2}(\theta - i\frac{2\pi}{3})} = f_{\frac{2}{3}}. \tag{18.1.4}$$

One can extract the value of the *on-shell* renormalized coupling constant[1] comparing with the Feynman diagrams coming from the lagrangian (18.1.1), as shown in Fig. 18.1

$$-ig^2 \;=\; 3m^4 \sinh(2i\pi/3) \;=\; i\frac{3\sqrt{3}}{2}m^4. \tag{18.1.5}$$

Unitarity paradox and its solution. Notice that the residue has *opposite sign* with respect to what is expected in a unitary theory. On the other hand, the S-matrix (18.1.4) satisfies by construction the unitarity equation $S(\theta)\,S(-\theta) = 1$. Hence, it seems we are in the presence of an apparent contradiction. The solution of this paradox and, consequently, the compatibility of the two definitions of the unitarity condition is the following. For this theory, it is possible to define a charge conjugate operator C ($C^2 = 1$) through the position

$$C\,\phi\,C \;=\; -\phi.$$

The hamiltonian associated to the lagrangian (18.1.1) is not hermitian but satisfies $H^\dagger = CHC$. The multiparticle states of the Fock space are created by the iterate action of the field ϕ on the vacuum state. They are eigenvectors of C with eigenvalues $(-1)^N$, where N is the number of particles. Since H is not a hermitian operator, its left

[1]This is defined as i times the residue at the pole of the S-matrix. The factor m^4 is introduced for dimensional reason.

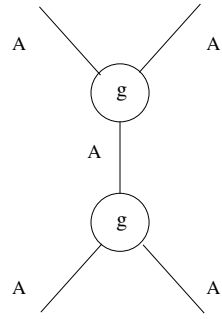


Fig. 18.1 *Residue at the pole expressed in terms of the* on-shell *coupling constant.*

eigenvectors $\langle n_l|$ do not coincide with the adjoint right eigenvectors, but are related to them by the relation $\langle n_l| = \langle n_r|C$. The completeness relation of this theory is then

$$\sum_n |\, n_r\rangle\langle n_l \,| = \sum_n |\, n_r\rangle\langle n_r \,|\; C.$$

The unitarity condition of the S-matrix

$$SS^\dagger = 1 \tag{18.1.6}$$

simply expresses that the initial and final states form a basis of the Hilbert space and it is not sensitive to whether the hamiltonian is hermitian or not. However, if we insert the completeness relation in (18.1.6), each of the intermediate states is weighted by $(-1)^N$. This is the reason for the negative sign of the residue, for it comes from the one-particle intermediate state. In conclusion, the S-matrix is unitary since it conserves the probability but it has a negative sign of the residue for the negative eigenvalue of C on the one-particle state. Because of its simplicity, the Yang–Lee model has proved to be the ideal theoretical playground for the analysis of integrable deformations of conformal models. A successful check of this S-matrix can be performed by the thermodynamic Bethe ansatz, as discussed in the next chapter.

Bullogh–Dodd model. The S-matrix of the Yang–Lee model is the so-called *minimal part* of the S-matrix of the Bullogh–Dodd model, defined by the lagrangian

$$\mathcal{L} \;=\; \frac{1}{2}(\partial_\mu\phi)^2 - \frac{\mu^2}{6\lambda^2}\left(2e^{\lambda\phi} + e^{-2\lambda\phi}\right). \tag{18.1.7}$$

To determine the S of this theory, notice that both models share the same spectrum of the spins of the conserved charges and have only one particle exitation. The S-matrix of the lagrangian model can then be obtained by multiplying the minimal S-matrix of the Yang–Lee model with some additional terms, called the Z *factors*, satisfying the following requirements: (i) they must be solutions of the bootstrap equation; (ii) they must not introduce additional poles; and, finally, (iii) they must depend on the coupling constant.

As discussed in Problem 1, another solution of the bootstrap equation (18.1.3) is given by

$$S(\theta) \;=\; f_{\frac{2}{3}}(\theta)\, f_{-\frac{B}{3}}(\theta)\, f_{\frac{B-2}{3}}(\theta), \tag{18.1.8}$$

and the quantity B can be determined by comparing the perturbative expansion of the S-matrix with the Feynman diagrams coming from the Bullogh–Dodd lagrangian. Notice that the two additional Z-factors introduce a set of zeros in the physical sheet of the scattering amplitude and no additional poles. From the perturbative comparison at lowest orders, one can conjecture that the exact result is expressed by the relation

$$B(\lambda) \;=\; \frac{\lambda^2}{2\pi}\frac{1}{1+\frac{\lambda^2}{4\pi}}. \tag{18.1.9}$$

Note that, assuming the validity of (18.1.9), the exact S-matrix of the Bullogh–Dodd model is invariant under the transformation $B(\lambda) \;\to\; 2 - B(\lambda)$, namely, under the weak–strong duality transformation of the coupling constant

$$\lambda \;\to\; \frac{4\pi}{\lambda}. \tag{18.1.10}$$

For all values of λ, except $\lambda = 0, \infty$ and the self-dual point $\lambda = \sqrt{4\pi}$, the S-matrix presents a simple pole at $\theta = 2\pi i/3$, which corresponds to the bound state given by the particle itself. The residue allows us to find the on-shell three-particle vertex of this theory

$$\Gamma^2(B) \;=\; 2\sqrt{3}\frac{\tan\left(\frac{\pi B}{6}\right)}{\tan\left(\frac{\pi B}{6}-\frac{2\pi}{3}\right)}\frac{\tan\left(\frac{\pi}{3}-\frac{\pi B}{6}\right)}{\tan\left(\frac{\pi B}{6}+\frac{\pi}{3}\right)}. \tag{18.1.11}$$

Notice that $\Gamma(B)$ vanishes at $B = 0$ and $B = 2$ (both points correspond to the free lagrangian model) but it also vanishes at the self-dual point $B = 1$, where the S-matrix becomes

$$S(\theta, 1) \;=\; f_{-\frac{2}{3}}(\theta). \tag{18.1.12}$$

The vanishing of Γ at the self-dual point is essentially due to the non-simply laced nature of this Toda field theory.[2]

18.2 $\Phi_{1,3}$ Integrable Deformation of the Conformal Minimal Models $\mathcal{M}_{2,2n+3}$

The Yang–Lee model belongs to the series of non-unitary minimal models $\mathcal{M}_{2,2n+3}$, whose Kac table consists of only one column. In these theories, in addition to the identity operator, there are n conformal fields with negative conformal weights

$$\Delta_{1,r} = \Delta_{1,2n+3-r} = -\frac{(r-1)(2n+2-r)}{2(2n+3)}, \qquad r = 0, 1, \ldots n. \tag{18.2.1}$$

The central charge c and the effective central charge c_{eff} are given by

$$c = -\frac{2n(6n+5)}{2n+3}, \qquad \tilde{c} = \frac{2n}{2n+3}.$$

[2]This does not happen for all the other Toda field theories based on simply laced algebras.

The scattering theory defined by the $\Phi_{1,3}$ integrable deformation is supported by the spectrum of the spins of the conserved charges given by

$$s = 1, 3, \ldots, 2n-1, 2n+3, \ldots 4n+1 \pmod{4n+2}.$$

This spectrum is compatible with a set of n massive particles with the bootstrap fusions

$$\begin{array}{c} A_1 \times A_1 \to A_2 \\ A_2 \times A_2 \to A_3 \\ \cdots \\ A_n \times A_n \to A_1. \end{array} \tag{18.2.2}$$

Using the results of the previous chapter, a solution of the consistency equations for the resonance angles is

$$\overline{u}^k_{k,k+1} = \frac{k\pi}{2n+1}, \qquad k = 1, 2, \ldots n,$$

and, consequently, the exact mass spectrum is

$$m_a \;=\; \sin\left(\frac{a\pi}{2n+1}\right), \qquad a = 1, 2, \ldots n. \tag{18.2.3}$$

The minimal scattering amplitude of the lowest mass particle A_1 is

$$S_{11}(\theta) \;=\; f_{\frac{2}{2n+1}}(\theta), \tag{18.2.4}$$

whereas all other amplitudes can be obtained by applying recursively the bootstrap equations

$$S_{ab}(\theta) \;=\; f_{\frac{|a-b|}{2n+1}}\, f_{\frac{a+b}{2n+1}} \prod_{k=1}^{min(a,b)-1} \left(f_{\frac{|a-b|+2k}{2n+1}}\right)^2. \tag{18.2.5}$$

The simple pole of the first term (for $a \neq b$)

$$\theta = i\, u_{ab}^{|a-b|} = i\, \left(1 - \frac{|\, a-b\,|}{2n+1}\right)\pi \tag{18.2.6}$$

corresponds to the particle $A_{|a-b|}$ which appears as a bound state in this scattering process. The simple pole of the second factor

$$\theta \;=\; i\, u_{ab}^{n(a,b)} = i\,(a+b)\frac{\pi}{2n+1} \tag{18.2.7}$$

is due to the particle of type $n(a,b) = \min\,(a+b, 2n+1-a-b)$. The double poles of the remaining functions derive from the bootstrap procedure and are associated to the multiple intermediate scattering processes, such as those shown in Fig. 18.2. In these

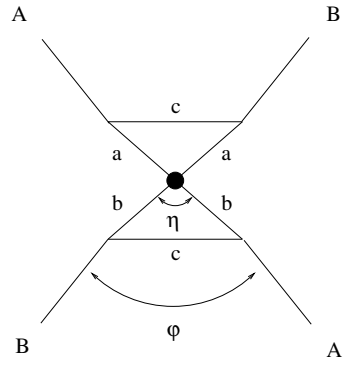


Fig. 18.2 *Multiscattering process that gives rise to a double pole in the S-matrix* $S_{AB}(\theta)$ *at* $\theta = i(u^a_{Ac} + u^b_{Bc} - \pi)$.

processes, two initial particles A and B "break" in the intermediate particles a, b, and c, with the relative angles dictated by scattering theory

$$\begin{aligned} \varphi &= u^a_{Ac} + u^b_{Bc} - \pi \\ \eta &= \pi - u^a_{ac} - u^B_{bc}. \end{aligned} \tag{18.2.8}$$

In order to actually draw these graphs and to have correspondingly a double pole, it is necessary that the resonance angles satisfy the geometrical condition

$$u^A_{ac} + u^B_{bc} \leq \pi \tag{18.2.9}$$

This condition puts a dynamical constraint on the set of resonance angles by having a double pole in the scattering amplitudes.

As we will see at the end of the chapter, these S-matrices can be obtained as RSOS reduction of the Sine–Gordon model, when the solitons disappear from the spectrum and only breathers remain. The amplitudes above are also the minimal S-matrices of the Toda field theories based on A^2_{2n}. In order to obtain the full S-matrix of these lagrangian models it is necessary to multiply the minimal S-matrix for the Z-factors that do not contain additional poles, a solution of the bootstrap equations, and functions of the coupling constant. For these theories this scheme can be implemented starting from the scattering amplitude involving the particle with the lowest mass

$$S_{11}(\theta) = f_{\frac{2}{2n+1}}(\theta)\, f_{-B}(\theta)\, f_{-\frac{2}{2n+1}+B}(\theta), \tag{18.2.10}$$

where the function $B(\lambda)$ is given by

$$B(\lambda) = \frac{1}{4\pi(2n+1)} \frac{\lambda^2}{1 + \frac{\lambda^2}{4\pi}}. \tag{18.2.11}$$

All other scattering amplitudes of the Toda field theories are obtained by applying the bootstrap equations.

18.3 Multiple Poles

The S-matrix discussed in the previous section shows the presence of double poles. In the S-matrices that we will meet in the following sections there are also higher order poles. This analytic structure is a necessary consequence of the bootstrap equations. However, a consistent interpretation of scattering theory requires an explanation of these higher order poles in terms of the elementary processes that take place in the system. Let's then briefly discuss the origin of these singularities in order to better understand the scattering theory of two-dimensional systems.

The simple poles of an S-matrix are associated to the bound states. This identification holds in any dimension and it is one of the key points of the analytic theory of the S-matrix. The higher order poles, on the other hand, only occur in the two-dimensional S-matrices. In four-dimensional theories, for instance, the same diagrams that produce the multiple poles of the two-dimensional theories give rise instead to branch cut singularities in the Mandelstam variable s. It is only for the dimensionality of the space-time that these singularities become double, triple, and higher order poles instead of branch cuts. In this respect, it is important to notice that two-dimensional scattering processes have the peculiar feature of being in one-to-one correspondence with the geometrical figures that one can draw on a page, i.e. the angles between the world-lines of the particles A_i, A_j, and A_k are precisely those associated to the resonance angles u_{ij}^k.

Assuming that the scattering theory corresponds to a set of Feynman rules (which for simplicity we assume to be of the $g\phi^3$), it is possible to prove that there is a very simple rule to determine the order of the pole: an S-matrix has a higher pole of order p

$$S_{ab}(\theta) \simeq \frac{g^{2p} R_p}{(\theta - \theta_0)^p}, \tag{18.3.1}$$

if one can actually draw the Feynman diagrams associated to this scattering process, starting from the resonance angles u_{ij}^k, in which there are P internal propagators and L loops, with the condition

$$p = P - 2L. \tag{18.3.2}$$

Applying this rule for instance to the diagram in Fig. 18.2, we see that there are six internal propagators and two loops, and therefore this diagram corresponds to a double order pole. However, as stressed in the previous section, it should actually be possible to draw such a diagram, i.e. the resonance angles u_{ij}^k should be those that permit the existence of such a geometrical figure. Analogously, if in the scattering theory there are resonance angles that allow us to draw a diagram such as the one shown in Fig. 18.3a, then there is a third-order pole in the amplitude, whereas the possibility of drawing a diagram as in Fig. 18.3b provides the explanation of a fourth-order pole in the scattering process of the asymptotic particles A and B.

Another general rule concerning the higher order poles is the following: those of odd orders can be generally associated to bound states, such as shown by the diagram in Fig. 18.3a, in which there is in the middle the propagator of a one-particle state, whereas those with even order generally describe multiscattering processes which do not lead to creation of a bound state, as is the case for instance of the diagram of Fig. 18.3b.

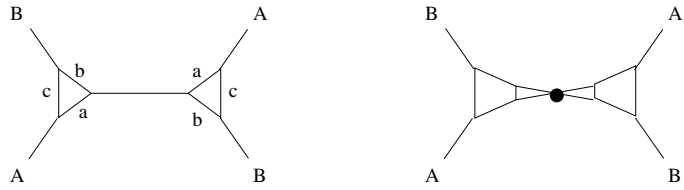


Fig. 18.3 *(a) Multiple scattering process responsible for a third order pole in $S_{AB}(\theta)$; (b) multiple scattering process that gives rise to a fourth order pole in S_{AB}.*

18.4 S-Matrices of the Ising Model

The Ising model has two integrable deformations. The first is the thermal deformation that moves the system away from its critical temperature at zero magnetic field. The second is the magnetic deformation, obtained by coupling the system to an external magnetic field but keeping the temperature of the system at its critical value. The S-matrices of the two deformations have a completely different structure: the first is the simplest possible S-matrix, while the second is the richest one! Moreover, the first is the minimal S-matrix of the lagrangian Sinh–Gordon model, whereas the second is the minimal S-matrix of the Toda field theory based on the exceptional algebra E_8. Let's discuss each of them in more detail.

18.4.1 Thermal Deformation of the Ising Model

At zero magnetic field, the Ising model away from the critical temperature is a theory of free Majorana fermions, with a lagrangian given by

$$\mathcal{L} \;=\; \psi\frac{\partial}{\partial\bar{z}}\psi + \bar{\psi}\frac{\partial}{\partial z}\bar{\psi} + i\,m\,\bar{\psi}\,\psi. \tag{18.4.1}$$

The mass parameter measures the displacement of the temperature from the critical value

$$m = T - T_c.$$

The low-temperature phase is related to the high-temperature phase by duality. At low temperature there are two degenerate vacua: the Z_2 symmetry of the model is spontaneously broken and the massive excitations consist of the soliton and the antisoliton that interpolate between the two vacua. These are neutral particles, here denoted by $A(\theta)$, associated to the fermionic field. Since the S-matrix can be regarded as the operator that implements the commutation relation between the operators that create the particles, by the fermionic nature of the problem we have

$$A(\theta_1)\,A(\theta_2) \;=\; -A(\theta_2)\,A(\theta_1),$$

namely

$$S = -1. \tag{18.4.2}$$

In the high-temperature phase the system has a unique vacuum. The massive excitation $A(\theta)$ of this phase is odd under the Z_2 spin symmetry and can be regarded as a bosonic particle created by the operator $\sigma(x)$, since we have

$$\langle 0|\sigma(0)|A(\theta)\rangle \neq 0.$$

From the self-duality of the model, the S-matrix is, as before, $S = -1$. Notice, however, that in this phase the particle A appears as an interacting particle, otherwise its S-matrix should be that of a free theory given, for a bosonic theory, by $S = 1$. In both phases the model does not have additional bound states.

As in the models analyzed in the previous section, the S-matrix of the thermal deformation of the Ising model can be regarded as the minimal S-matrix of a lagrangian system, in this case the Sinh–Godon model with lagrangian

$$\mathcal{L} \;=\; \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{g^2}(\cosh g\phi - 1). \tag{18.4.3}$$

To determine the exact S-matrix of this integrable model, we have to identify the appropriate Z-factor as a function of the coupling constant. The simplest choice leads to the following expression for the exact S of the Sinh–Gordon model

$$S(\beta) \;=\; f_{-B}(\theta), \tag{18.4.4}$$

where $B(g)$ is a function of the coupling constant which can be determined by comparing the perturbative expansion of (18.4.4) with the Feynman diagrams coming from the lagrangian (18.4.3). The final result is

$$B(g) \;=\; \frac{g^2}{8\pi}\frac{1}{1+\frac{g^2}{8\pi}}. \tag{18.4.5}$$

As we will see in Section 18.8, this expression can be obtained as the analytic continuation of an analogous formula established for the Sine–Gordon model. Notice the invariance of the S-matrix of the Sinh–Gordon model under the weak–strong duality

$$g \;\to\; \frac{8\pi}{g}. \tag{18.4.6}$$

This symmetry is not evident in the lagrangian of the model and it is only shown up in its exact S-matrix. Presently it is still an open problem to find the proper lagrangian formulation (if any) that explicitly shows this dynamical invariance of the Sinh–Gordon model.

Coming back to the Ising model, it is interesting to observe that its S-matrix can be obtained as a limiting case of the exact S-matrix of a generic Z_n model, when $T > T_c$, obtained by Koberle and Swieca. In this theory there are $n-1$ particles, with mass spectrum given by

$$m_a \;=\; \sin\left(\frac{\pi a}{n}\right), \quad a = 1, 2, \ldots n-1. \tag{18.4.7}$$

The S-matrix of the fundamental particle is

$$S_{11} \;=\; \frac{\tanh\frac{1}{2}(\theta + i\frac{2\pi}{n})}{\tanh\frac{1}{2}(\theta - i\frac{2\pi}{3})} \;=\; f_{\frac{2}{n}}\,, \tag{18.4.8}$$

and, substituting $n = 2$ in this formula, we get the S-matrix (18.4.2).

18.4.2 Magnetic Deformation of the Ising Model

The counting argument shows that the magnetic deformation of the Ising model has a spectrum of the first spins of the conserved charges given by

$$s \;=\; 1, 7, 11, 13, 17, 19. \tag{18.4.9}$$

Notice the lack of spins that have 3 or 5 as divisors. The absence of multiples of 3 can be easily explained by postulating the existence of a fundamental particle A_1 (with mass m_1) that possesses the "Φ^3" property, i.e. to be a bound state of itself. In the S-matrix of this particle there should then be a pole at $\theta = u^1_{11} = 2\pi i/3$. This feature is compatible with the explicit breaking of the Z_2 of the model. Concerning the absence of spin s divisible by 5, it can be explained by conjecturing the existence of a second particle A_2 (with mass m_2) that, together with A_1, gives rise to a subsystem of bootstrap fusions

$$\begin{array}{l} A_1 \times A_1 \;\to\; A_1 + A_2 \\ A_2 \times A_2 \;\to\; A_1. \end{array} \tag{18.4.10}$$

Let u^2_{11} be the resonance angle corresponding to the bound state A_2 that appears in the amplitude $S_{11}(\theta)$, and let u^1_{22} be the resonance angle associated to A_1 in the amplitude $S_{22}(\theta)$. Using the variables

$$y_1 \;=\; \exp\left(\frac{i}{2}u^2_{11}\right)\,, \quad y_2 \;=\; \exp\left(\frac{i}{2}u^1_{22}\right)\,,$$

the consistency equations involving the spins of conserved charges and the resonance angles become

$$\begin{array}{rcl} y_1^s + y_1^{-s} &=& \left(\frac{m_2}{m_1}\right)^s \frac{\chi^2_s}{\chi^1_s}, \\ y_2^s + y_2^{-s} &=& \left(\frac{m_1}{m_2}\right)^s \frac{\chi^1_s}{\chi^2_s}. \end{array} \tag{18.4.11}$$

For the values of s given in (18.4.9), a non-trival solution is given by

$$y_1 \;=\; \exp\left(\frac{i\pi}{5}\right)\,, \quad y_2 \;=\; \exp\left(\frac{2i\pi}{5}\right)\,.$$

with the mass ratio

$$\frac{m_2}{m_1} \;=\; 2\cos\frac{\pi}{5}.$$

In light of these results, one can conclude that in the amplitude $S_{11}(\theta)$ of the fundamental particle there are poles with positive residue at the resonance angles

$$\theta \;=\; iu_{11}^1 = \frac{2\pi i}{3}\,, \qquad \theta = iu_{11}^2 \;=\; \frac{2\pi i}{5}, \tag{18.4.12}$$

and poles with negative residue in the crossing channel at

$$\theta \;=\; i\overline{u}_{11}^1 = \frac{i\pi}{3}\,, \qquad \theta = i\overline{u}_{11}^2 \;=\; \frac{3\pi i}{5}. \tag{18.4.13}$$

However, as we have seen in Section 18.1, it is impossible to solve the bootstrap equations

$$S_{11}(\theta) \;=\; S_{11}\left(\theta - \frac{i\pi}{3}\right) S_{11}\left(\theta + \frac{i\pi}{3}\right) \tag{18.4.14}$$

in terms of a function that has only the sets of poles (18.4.12) and (18.4.13). In fact, it is necessary to include at least another set of poles, without breaking the conserved currents with spins given in (18.4.9). The minimal way to do so is to introduce an additional pole at $\theta = i\pi/15$ (with positive residue) and its companion of the crossed channel at $\theta = i14/15$ (with negative residue). In such a way, the exact S-matrix of the fundamental particle is expressed by

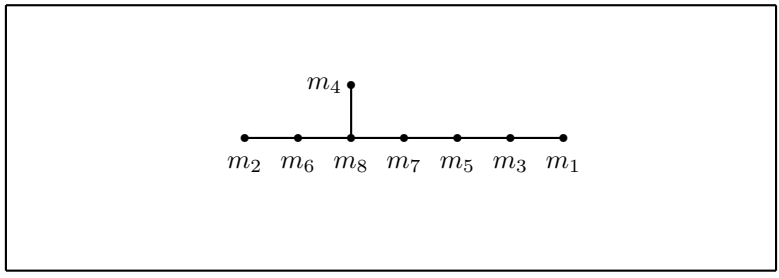
$$S_{11}(\theta) \;=\; f_{\frac{2}{3}}(\theta) f_{\frac{2}{5}}(\theta) f_{\frac{1}{15}}(\theta). \tag{18.4.15}$$

Using this expression as the initial seed of the bootstrap equation, we can complete the bootstrap procedure. The final theory has eight particles, whose mass spectrum coincides with that of the Toda field theory based on the exceptional algebra E_8:

$$\begin{aligned}
m_1 &= m \\
m_2 &= 2m_1 \cos\frac{\pi}{5} = (1.6180339887..)\, m_1 \\
m_3 &= 2m_1 \cos\frac{\pi}{30} = (1.9890437907..)\, m_1 \\
m_4 &= 2m_2 \cos\frac{7\pi}{30} = (2.4048671724..)\, m_1 \\
m_5 &= 2m_2 \cos\frac{2\pi}{15} = (2.9562952015..)\, m_1 \\
m_6 &= 2m_2 \cos\frac{\pi}{30} = (3.2183404585..)\, m_1 \\
m_7 &= 4m_2 \cos\frac{\pi}{5}\cos\frac{7\pi}{30} = (3.8911568233..)\, m_1 \\
m_8 &= 4m_2 \cos\frac{\pi}{5}\cos\frac{2\pi}{15} = (4.7833861168..)\, m_1.
\end{aligned}$$

As observed in Chapter 16, the masses can be put in correspondence with the Perron–Frobenius eigenvector of the incidence matrix of the corresponding Dynkin diagram, as

Table 18.2: *Dynkin diagram of E_8, with the association of the masses to the dots of the diagram.*



shown in Table 18.2. Notice that in this bootstrap system only the first three particles have a mass less than the lowest threshold $2m_1$. The stability of the particles with mass higher than the lowest decay threshold $2m_1$ is entirely due to the integrability of the theory.

The complete set of scattering amplitudes is given in Tables 8.3 and 8.4, where we use the notation

$$(\gamma) \equiv f_{\frac{\gamma}{30}}(\theta).$$

Several amplitudes have higher order poles that can be explained in terms of the multiscattering processes constructed in terms of the resonance angles of the theory.

E_8 **Toda theory**. The underlying E_8 structure of this scattering theory can be traced back to the coset realization $(E_8)_1 \otimes (E_8)_1/(E_8)_2$ of the critical Ising model and its Liouville quantization based on the same set of simple roots (see Section 16.7). This suggests that to obtain the exact S-matrix of the lagrangian Toda field theory based on E_8 it is sufficient to multiply the minimal S-matrix elements provided by the Ising model in a magnetic field by the appropriate Z-factors that carry the coupling constant dependence from λ, without introducing addional poles in the physical sheet. For the amplitude of the fundamental particle, the Z-factor is given by

$$Z_{11}(\theta) = f_{-B}(\theta)\, f_{-\frac{1}{15}+B}(\theta)\, f_{-\frac{2}{3}-B}(\theta)\, f_{-\frac{2}{5}+B}(\theta), \tag{18.4.16}$$

where

$$B(\lambda) = \frac{2}{h}\,\frac{\lambda^2}{8\pi}\,\frac{1}{1+\frac{\lambda^2}{8\pi}}, \tag{18.4.17}$$

with $h = 30$, the Coxeter number of this algebra. Also in this case, the S-matrix of the lagrangian model presents remarkable symmetry under weak–strong duality:

$$\lambda \to \frac{8\pi}{\lambda}.$$

Table 18.3: *S-matrix of the Ising model in a magnetic field at $T = T_c$. The factors $\left(f_{\gamma/30}(\theta)\right)^{p_\gamma}$ in $S_{ab}(\theta)$ correspond to $(\gamma)^{p_\gamma}$ ($p_\gamma = 1$ is omitted). The upper index* **c** *in (γ) denotes the particle A_c which appears as a bound state of $A_a A_b$ at $\theta = i\pi\gamma/30$ in the amplitudes $S_{ab}(\theta)$.*

a	b	S_{ab}
1	1	$(20)^{\mathbf{1}}\,(12)^{\mathbf{2}}\,(2)^{\mathbf{3}}$
1	2	$(24)^{\mathbf{1}}\,(18)^{\mathbf{2}}\,(14)^{\mathbf{3}}\,(8)^{\mathbf{4}}$
1	3	$(29)^{\mathbf{1}}\,(21)^{\mathbf{2}}\,(13)^{\mathbf{4}}\,(3)^{\mathbf{5}}\;(11)^2$
1	4	$(25)^{\mathbf{2}}\,(21)^{\mathbf{3}}\,(17)^{\mathbf{4}}\,(11)^{\mathbf{5}}\,(7)^{\mathbf{6}}\;(15)$
1	5	$(28)^{\mathbf{3}}\,(22)^{\mathbf{4}}\,(14)^{\mathbf{6}}\,(4)^{\mathbf{7}}\;(10)^2\,(12)^2$
1	6	$(25)^{\mathbf{4}}\,(19)^{\mathbf{5}}\,(9)^{\mathbf{7}}\;(7)^2\,(13)^2\,(15)$
1	7	$(27)^{\mathbf{5}}\,(23)^{\mathbf{6}}\,(5)^{\mathbf{8}}\;(9)^2\,(11)^2\,(13)^2\,(15)$
1	8	$(26)^{\mathbf{7}}\,(16)^{\mathbf{8}3}\;(6)^2\,(8)^2\,(10)^2\,(12)^2$
2	2	$(24)^{\mathbf{1}}\,(20)^{\mathbf{2}}\,(14)^{\mathbf{4}}\,(8)^{\mathbf{5}}\,(2)^{\mathbf{6}}\;(12)^2$
2	3	$(25)^{\mathbf{1}}\,(19)^{\mathbf{3}}\,(9)^{\mathbf{6}}\;(7)^2\,(13)^2\,(15)$
2	4	$(27)^{\mathbf{1}}\,(23)^{\mathbf{2}}\,(5)^{\mathbf{7}}\;(9)^2\,(11)^2\,(13)^2\,(15)$
2	5	$(26)^{\mathbf{2}}\,(16)^{\mathbf{6}3}\;(6)^2(8)^2(10)^2(12)^2$
2	6	$(29)^{\mathbf{2}}\,(25)^{\mathbf{3}}\,(19)^{\mathbf{5}3}\,(13)^{\mathbf{7}3}\,(3)^{\mathbf{8}}\;(7)^2(9)^2(15)$
2	7	$(27)^{\mathbf{4}}\,(21)^{\mathbf{6}3}\,(17)^{\mathbf{7}3}\,(11)^{\mathbf{8}3}\;(5)^2(7)^2(15)^2$
2	8	$(28)^{\mathbf{6}}\,(22)^{\mathbf{7}3}\;(4)^2(6)^2(10)^4(12)^4(16)^4$
3	3	$(22)^{\mathbf{2}}\,(20)^{\mathbf{3}3}\,(14)^{\mathbf{5}}\,(12)^{\mathbf{6}3}\,(4)^{\mathbf{7}}\;(2)^2$
3	4	$(26)^{\mathbf{1}}\,(16)^{\mathbf{5}3}\;(6)^2(8)^2(10)^2(12)^2$
3	5	$(29)^{\mathbf{1}}\,(23)^{\mathbf{3}}\,(21)^{\mathbf{4}3}\,(13)^{\mathbf{7}3}\,(5)^{\mathbf{8}}\;(3)^2(11)^4(15)$
3	6	$(26)^{\mathbf{2}}\,(24)^{\mathbf{3}3}\,(18)^{\mathbf{6}3}\,(8)^{\mathbf{8}3}\;(10)^2(16)^4$
3	7	$(28)^{\mathbf{3}}\,(22)^{\mathbf{5}3}\;(4)^2(6)^2(10)^4(12)^4(16)^4$
3	8	$(27)^{\mathbf{5}}\,(25)^{\mathbf{6}3}\,(17)^{\mathbf{8}5}\;(7)^4(9)^4(11)^2(15)^3$
4	4	$(26)^{\mathbf{1}}\,(20)^{\mathbf{4}3}\,(16)^{\mathbf{6}3}\,(12)^{\mathbf{7}3}\,(2)^{\mathbf{8}}\;(6)^2(8)^2$
4	5	$(27)^{\mathbf{1}}\,(23)^{\mathbf{3}3}\,(19)^{\mathbf{5}3}\,(9)^{\mathbf{8}3}\;(5)^2(13)^4(15)^2$
4	6	$(28)^{\mathbf{1}}\,(22)^{\mathbf{4}3}\;(4)^2(6)^2(10)^4(12)^4(16)^4$

Table 18.3: *continued.*

4	7	$\overset{\mathbf{2}}{(28)}\,\overset{\mathbf{4}}{(24)^3}\,\overset{\mathbf{7}}{(18)^5}\,\overset{\mathbf{8}}{(14)^5}\;(4)^2(8)^4(10)^4$
4	8	$\overset{\mathbf{4}}{(29)}\,\overset{\mathbf{5}}{(25)^3}\,\overset{\mathbf{7}}{(21)^5}\;(3)^2(7)^4(11)^6(13)^6(15)^3$
5	5	$\overset{\mathbf{4}}{(22)^3}\,\overset{\mathbf{5}}{(20)^5}\,\overset{\mathbf{8}}{(12)^5}\;(2)^2(4)^2(6)^2(16)^4$
5	6	$\overset{\mathbf{1}}{(27)}\,\overset{\mathbf{2}}{(25)^3}\,\overset{\mathbf{7}}{(17)^5}\;(7)^4(9)^4(11)^4(15)^3$
5	7	$\overset{\mathbf{1}}{(29)}\,\overset{\mathbf{3}}{(25)^3}\,\overset{\mathbf{6}}{(21)^5}\;(3)^2(7)^4(11)^6(13)^6(15)^3$
5	8	$\overset{\mathbf{3}}{(28)}\,\overset{\mathbf{4}}{(26)^3}\,\overset{\mathbf{5}}{(24)^5}\,\overset{\mathbf{8}}{(18)^7}\;(8)^6(10)^6(16)^8$
6	6	$\overset{\mathbf{3}}{(24)^3}\,\overset{\mathbf{6}}{(20)^5}\,\overset{\mathbf{8}}{(14)^5}\;(2)^2(4)^2(8)^4(12)^6$
6	7	$\overset{\mathbf{1}}{(28)}\,\overset{\mathbf{2}}{(26)^3}\,\overset{\mathbf{5}}{(22)^5}\,\overset{\mathbf{8}}{(16)^7}\;(6)^4(10)^6(12)^6$
6	8	$\overset{\mathbf{2}}{(29)}\,\overset{\mathbf{3}}{(27)^3}\,\overset{\mathbf{6}}{(23)^5}\,\overset{\mathbf{7}}{(21)^7}\;(5)^4(11)^8(13)^8(15)^4$
7	7	$\overset{\mathbf{2}}{(26)^3}\,\overset{\mathbf{4}}{(24)^5}\,\overset{\mathbf{7}}{(20)^7}\;(2)^2(8)^6(12)^8(16)^8$
7	8	$\overset{\mathbf{1}}{(29)}\,\overset{\mathbf{2}}{(27)^3}\,\overset{\mathbf{4}}{(25)^5}\,\overset{\mathbf{6}}{(23)^7}\,\overset{\mathbf{8}}{(19)^9}\;(9)^8(13)^{10}(15)^5$
8	8	$\overset{\mathbf{1}}{(28)^3}\,\overset{\mathbf{3}}{(26)^5}\,\overset{\mathbf{5}}{(24)^7}\,\overset{\mathbf{7}}{(22)^9}\,\overset{\mathbf{8}}{(20)^{11}}\;(12)^{12}(16)^{12}$

Table 18.4: *Mass spectrum of the thermal tricritical Ising model, together with their numerical values and the Z_2 quantum numbers.*

m_1	=	M	1	odd
m_2	=	$2M\cos(\frac{5\pi}{18})$	1.28557	even
m_3	=	$2M\cos(\frac{\pi}{9})$	1.87938	odd
m_4	=	$2M\cos(\frac{\pi}{18})$	1.96961	even
m_5	=	$4M\cos(\frac{5\pi}{18})\cos(\frac{\pi}{18})$	2.53208	even
m_6	=	$4M\cos(\frac{\pi}{9})\cos(\frac{2\pi}{9})$	2.87938	odd
m_7	=	$4M\cos(\frac{\pi}{18})\cos(\frac{\pi}{9})$	3.70166	even

Bootstrap fusion rules. The bootstrap fusion rules of both models (Ising and Toda) can be written in a general form once a proper notation is introduced. Notice that the squares of the masses $\{m_1, m_6, m_5, m_7\}$ are the roots of the fourth-order polynomial

$$P_1 = x^4 - 30x^3 + 300x^2 - 1080x + 720$$

and for these quantities, let's introduce the notation

$$(m_1, m_6, m_5, m_7) \to \quad (C_1, C_2, C_3, C_4).$$

The squares of the masses $\{m_2, m_3, m_8, m_4\}$ are instead the roots of the fourth-order polynomial

$$P_1 = x^4 - 30x^3 + 240x^2 - 720x + 720$$

and to denote them, let's introduce the notation

$$(m_2, m_3, m_8, m_4) \rightarrow \quad (B_1, B_2, B_3, B_4).$$

In this way, the bootstrap fusion rules of the bootstrap fusions of the bound states related to the E_8 algebra can be written as (with cyclic notation, i.e. $B_{i+4} \equiv B_i$ and $C_{i+4} \equiv C_i$)

$$\begin{aligned}
C_i \times C_i &= C_i + B_i + B_{i+1} \\
C_i \times C_{i+1} &= C_{i+2} + C_{i+3} + B_{i+3} \\
C_i \times C_{i+2} &= C_{i+1} + C_{i+3} + B_{i+1} + B_{i+3} \\
C_i \times C_{i+3} &= C_{i+1} + C_{i+2} + B_{i+2} \\
B_i \times B_i &= C_i + A_{i+1} + C_{i+2} + B_i + B_{i+3} \\
B_i \times B_{i+1} &= C_i + C_{i+1} + B_{i+1} \\
B_i \times B_{i+2} &= C_{i+1} + C_{i+3} \\
B_i \times B_{i+3} &= C_i + B_i + C_{i+3} \\
C_i \times B_i &= C_i + B_i + B_{i+1} + B_{i+3} \\
C_i \times B_{i+1} &= C_i + C_{i+2} + B_i + B_{i+3} \\
C_i \times B_{i+2} &= B_{i+2} + C_{i+3} \\
C_i \times B_{i+3} &= C_{i+1} + C_{i+2} + B_i + B_{i+1} + B_{i+3}.
\end{aligned} \tag{18.4.18}$$

An explicit check of the S-matrix of the Ising model in a magnetic field is provided by the thermodynamics Bethe ansatz, as discussed in more detail in Chapter 19.

18.5 The Tricritical Ising Model at $T \neq T_c$

The tricritical Ising model away from its critical temperature is described by the integrable deformation $\epsilon = \Phi_{1,2}$ with conformal weight $\Delta = 1/10$. This corresponds to the massive deformation of the Liouville action based on the E_7 algebra. Therefore we expect that the corresponding scattering theory involves seven particles. Let's see how this theory can be derived. The perturbed action is

$$\mathcal{S} = \mathcal{S}_{CFT} + \lambda \int d^2z \, \epsilon(z, \bar{z}). \tag{18.5.1}$$

For $\lambda > 0$ the system is in its Z_2 symmetric phase. Its low-temperature phase, $\lambda < 0$, is related by duality to the high-temperature one. Therefore, in the following we focus our attention only on the massive theory (18.5.1) with $\lambda > 0$.

The spins of the conserved charges coincide with the Coxeter exponents of the Toda field theory based on the E_7 algebra, whose Coxeter number is $h = 18$

$$s = 1, 5, 7, 9, 11, 13, 17 \pmod{18}.$$

In computing the mass spectrum and the scattering amplitudes, it is important to notice that the *fundamental particle* cannot be a bound state of itself for the Z_2 symmetry of the model which can be used to label the particles. We expect that the fundamental particle is odd under this symmetry and therefore cannot fulfill the Φ^3 property. However, the existence of a Z_2 even particle with the Φ^3 property is not in contradiction with the spins of the conserved charges, as long as the charge $\mathcal{Q}_9$

annihilates this state. In the light of this observation, let's assume that the lightest Z_2 even particle, here denoted by A_2, appears as a bound state in the scattering amplitude of the fundamental Z_2 odd particle A_1. Since for the eigenvalues of the conserved charges we have $\chi_9^{(1)} \neq 0$ but $\chi_9^{(2)} = 0$, using the consistency equation one obtains the resonance angle u_{11}^2 by the condition

$$\cos\left(\frac{9u_{11}^2}{2}\right) = 0. \tag{18.5.2}$$

The solution that gives rise to a consistent system is identified as $u_{11}^2 = 5\pi/9$. This fixes the mass ratio of these two particles

$$m_2 = 2\cos\left(\frac{5\pi}{18}\right) m_1.$$

The pole in S_{11} at $\theta = i5\pi/9$ with positive residue implies a pole in S_{12} at $\theta = i5\pi/9$ with negative residue, corresponding to the particle A_1 in the crossed channel. With these data, the bootstrap equations that involve S_{11} and S_{12} become

$$S_{12}(\theta) = S_{11}\left(\theta + i\frac{5\pi}{18}\right) S_{11}\left(\theta - i\frac{5\pi}{18}\right), \tag{18.5.3}$$

$$S_{11}(\theta) = S_{11}\left(\theta + i\frac{4\pi}{9}\right) S_{12}\left(\theta - i\frac{5\pi}{18}\right). \tag{18.5.4}$$

One cannot satisfy these equations only with a pole in S_{11} and S_{12}. The minimal way to satisfy them is to introduce an additional pole at $\theta = i\pi/9$ (with positive residue) in S_{11} and a pole at $\theta = i\,7\pi/18$ (with positive residue) in S_{12}. The new pole at $\theta = i\,\pi/9$ in S_{11} corresponds to a new bound state A_4, a particle that is even under the Z_2 symmetry, with mass

$$m_4 = 2\cos\left(\frac{\pi}{18}\right) m_1.$$

The pole at $\theta = i7\pi/18$ in S_{12} represents another bound state A_3, odd under the Z_2 symmetry, with mass

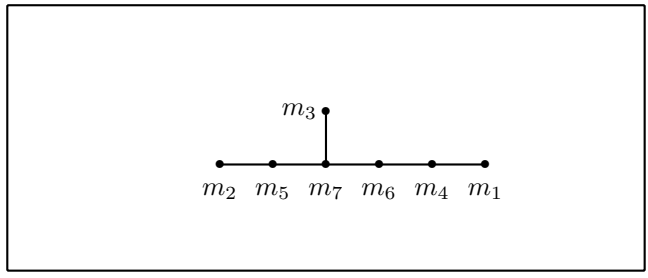
$$m_3 = 2\cos\left(\frac{\pi}{9}\right) m_1.$$

So,

$$S_{11}(\theta) = -f_{\frac{1}{9}}(\theta)\, f_{\frac{5}{9}}(\theta), \qquad S_{12}(\theta) = f_{\frac{7}{18}}(\theta)\, f_{\frac{13}{18}}(\theta). \tag{18.5.5}$$

All other amplitudes can be iteratively computed by employing the bootstrap equations (17.4.12). The bootstrap process closes with seven particles, whose masses and Z_2 quantum numbers are given in Table 18.4. Also in this case, as in the Ising model in a magnetic field, the masses can be associated to the dots of the Dynkin diagram of the E_7 algebra. They enter in fact the component of the Perron–Frobenius eigenvector of the incidence matrix of this Dynkin diagram, see Table 18.5. The complete set of scattering amplitudes is shown in Table 18.6.

Table 18.5: *Dynkin diagram of the E_7 algebra and correspondence between the masses of the particles and the dots of the diagram.*



Bootstrap fusion rules. According to the roots of the algebraic equations that determine the masses of the corresponding Toda field theory (see Section 16.6), the seven particles can be organized into two triplets and one singlet:

$$\begin{array}{rcl}(Q_1, Q_2, Q_3) & \equiv & (m_6, m_3, m_1) \\ (K_1, K_2, K_3) & \equiv & (m_2, m_4, m_7) \\ (N) & \equiv & (m_5).\end{array}$$

The first triplet consists of the odd particles under the Z_2 symmetry. The second triplet and the singlet are made of Z_2 even particles. The bootstrap fusions that involve $[N]$ and $[N, K_i]$ form a closed subsystem of these fusions:

$$\begin{array}{llll} N \cdot N = N & , & N \cdot K_A = K_1 + K_2 + K_3 \\ K_A \cdot K_{A+1} = K_A + N & , & K_A \cdot K_A = K_A + K_{A+1} + N. \end{array} \tag{18.5.6}$$

The other particles couple only to the previous ones

$$\begin{array}{llll} K_A \cdot Q_A = Q_{A+1} & , & K_A \cdot Q_{A+1} = Q_1 + Q_2 + Q_3 \\ K_A \cdot Q_{A-1} = Q_{A-1} + Q_{A+1} & , & Q_A \cdot Q_A = K_{A-1} + K_{A+1} \\ Q_A \cdot Q_{A+1} = K_A + K_{A-1} + N & , & N \cdot Q_A = Q_{A-1} + Q_{A+1}. \end{array} \tag{18.5.7}$$

A check that confirms the validity of this S-matrix description of the thermal deformation of the tricritical Ising model will be given by the thermodynamics Bethe ansatz.

E_7 **Toda theory**. As for the other S-matrices previously discussed, also the S-matrix of the thermal deformation of the tricritical Ising model can be used as a minimal S-matrix of the corresponding lagrangian model, given by the Toda field theory based on the exceptional E_7 algebra. In this case the Z-factor that enters the amplitude of the fundamental particle is

$$Z_{11}(\theta) \; = \; -f_{-B}(\theta)\, f_{-\frac{1}{9}+B}(\theta)\, f_{-\frac{4}{9}-B}(\theta), \tag{18.5.8}$$

Table 18.6: *S-matrix of the thermal deformation of the tricritical Ising model. The factors* $\big(t_{\gamma/18}(\theta)\big)^{p_\gamma}$ *in* $S_{ab}(\theta)$ *correspond to* $(\gamma)^{p_\gamma}$ *(*$p_\gamma = 1$ *is omitted). The upper index* **c** *in* (γ) *denotes the bound state* A_c *in the amplitude* $S_{ab}(\theta)$*, whose pole is at* $\theta = i\pi\gamma/18$*.*

a	b	S_{ab}
1	1	$-\,(10)^{\mathbf{2}}\,(1)^{\mathbf{4}}$
1	2	$(13)^{\mathbf{1}}\,(7)^{\mathbf{3}}$
1	3	$-\,(14)^{\mathbf{2}}\,(10)^{\mathbf{4}}\,(6)^{\mathbf{5}}$
1	4	$(17)^{\mathbf{1}}\,(11)^{\mathbf{3}}\,(3)^{\mathbf{6}}\;(9)$
1	5	$(14)^{\mathbf{3}}\,(8)^{\mathbf{6}}\;(6)^3$
1	6	$-\,(16)^{\mathbf{4}}\,(12)^{\mathbf{5}}\,(4)^{\mathbf{7}}\;(10)^2$
1	7	$(15)^{\mathbf{6}}\;(9)\,(5)^2\,(7)^2$
2	2	$(12)^{\mathbf{2}}\,(8)^{\mathbf{4}}\,(2)^{\mathbf{5}}$
2	3	$(15)^{\mathbf{1}}\,(11)^{\mathbf{3}}\,(5)^{\mathbf{6}}\;(9)$
2	4	$(14)^{\mathbf{2}}\,(8)^{\mathbf{5}}\;(6)^2$
2	5	$(17)^{\mathbf{2}}\,(13)^{3\,\mathbf{4}}\,(3)^{\mathbf{7}}\;(9)\,(7)^2$
2	6	$(15)^{\mathbf{3}}\;(9)\,(5)^2(7)^2$
2	7	$(16)^{\mathbf{5}}\,(10)^{3\,\mathbf{7}}\;(4)^2(6)^2$
3	3	$-\,(14)^{\mathbf{2}}\,(2)^{\mathbf{7}}\;(12)^2\,(8)^2$
3	4	$(15)^{\mathbf{1}}\;(9)(5)^2(7)^2$
3	5	$(16)^{\mathbf{1}}\,(10)^{3\,\mathbf{6}}\;(4)^2(6)^2$
3	6	$(16)^{\mathbf{2}}\,(8)^{3\,\mathbf{7}}\,(12)^{3\,\mathbf{5}}\;(4)^2$
3	7	$(17)^{\mathbf{3}}\,(13)^{3\,\mathbf{6}}\;(9)^2(3)^2(7)^4$
4	4	$(12)^{\mathbf{4}}\,(4)^{\mathbf{7}}\,(10)^{3\,\mathbf{5}}\;(2)^2$
4	5	$(15)^{\mathbf{2}}\,(13)^{3\,\mathbf{4}}\,(7)^{3\,\mathbf{7}}\;(9)$
4	6	$(17)^{\mathbf{1}}\,(11)^{3\,\mathbf{6}}\;(9)^2(3)^2(5)^2$
4	7	$(16)^{\mathbf{4}}\,(14)^{3\,\mathbf{5}}\;(8)^4(12)^4$
5	5	$(12)^{\mathbf{5}\,3}\;(4)^2(2)^2(8)^4$
5	6	$(16)^{\mathbf{1}}\,(14)^{3\,\mathbf{3}}\;(6)^4(8)^4$
5	7	$(17)^{\mathbf{2}}\,(15)^{3\,\mathbf{4}}\,(11)^{5\,\mathbf{7}}\;(9)^3(5)^4$
6	6	$-\,(14)^{3\,\mathbf{4}}\,(10)^{5\,\mathbf{7}}\;(16)^2(12)^4$
6	7	$(17)^{\mathbf{1}}\,(15)^{3\,\mathbf{3}}\,(13)^{5\,\mathbf{6}}\;(9)^3(5)^6$
7	7	$(16)^{3\,\mathbf{2}}\,(14)^{5\,\mathbf{5}}\,(12)^{7\,\mathbf{7}}\;(8)^8$

with

$$B(g) \;=\; \frac{2}{h}\,\frac{g^2}{8\pi}\,\frac{1}{1+\frac{g^2}{8\pi}}, \tag{18.5.9}$$

where $h = 18$ is the Coxeter number of the E_7 algebra. All the other amplitudes can be obtained by applying the bootstrap equations.

18.6 Thermal Deformation of the Three-state Potts Model

The universality class of this model is described by a subset of operators of the minimal model $\mathcal{M}_{5,6}$ with central charge $c = 4/5$. The Landau–Ginzburg theory of the critical model is

$$\mathcal{L} = (\partial_\mu \Phi)(\partial_\mu \Phi^*) + \left[(\Phi)^3 + (\Phi^*)^3\right], \tag{18.6.1}$$

where Φ is a complex scalar field. The two most relevant magnetization operators can be identified with Φ and Φ^*, whereas the other two sub-leading magnetic operators correspond to $(\Phi^*)^2\Phi$ and $\Phi^*\Phi^2$. The energy operator is associated to $\Phi^*\Phi$.

Away from the critical temperature, the action of the model can be written as

$$\mathcal{A} \;=\; \mathcal{A}_{CFT} + \lambda \int \epsilon(x)\, d^2x. \tag{18.6.2}$$

and it corresponds to an integrable theory. In the Landau–Ginzburg formalism, the thermal deformation is equivalent to adding a mass term $m^2\Phi^*\Phi$ in the lagrangian (18.6.1). The perturbed theory is still invariant under the permutation group S_3 present at the critical point and therefore the particles can be labeled by the corresponding quantum numbers. An irreducible representation of this discrete symmetry group is given by a particle–antiparticle doublet $(A, \overline{A})$ of mass m. Under the action of the generators of the group, these states transform as

$$\vartheta\, A \;=\; \omega A; \quad \vartheta\, \overline{A} \;=\; \overline{\omega}\,\overline{A}; \quad C\, A \;=\; \overline{A},$$

where $\omega = \exp(2\pi i/3)$. In this case, the most general S-matrix is given by

$$\begin{aligned} \mid A(\theta_1)A(\theta_2)\rangle_{in} \;&=\; u(\theta_{12}) \mid A(\theta_1)A(\theta_2)\rangle_{out}; \\ \mid A(\theta_1)\overline{A}(\theta_2)\rangle_{in} \;&=\; t(\theta_{12}) \mid A(\theta_1)\overline{A}(\theta_2)\rangle_{out} + r(\theta_{12}) \mid \overline{A}(\theta_1)A(\theta_2)\rangle_{out}. \end{aligned}$$

However, as a direct consequence of the infinite conserved charges of this theory, it is easy to show that the reflection amplitude vanishes. Therefore the S-matrix is completely diagonal. Furthermore, the crossing invariance implies

$$t(\theta) \;=\; u(i\pi - \theta),$$

while the unitarity condition leads to

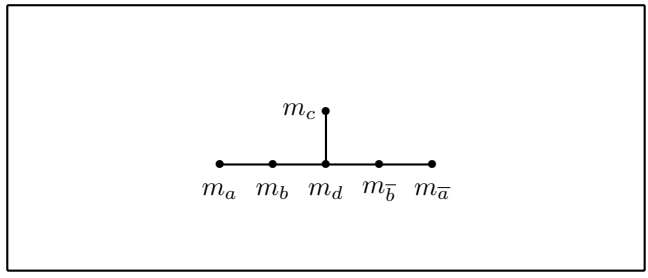
$$t(\theta)\, t(-\theta) \;=\; 1; \qquad u(\theta)\, u(-\theta) \;=\; 1.$$

The minimal solution of these equations is

$$u(\theta) \;=\; \frac{\sinh\left(\theta/2 + i\pi/3\right)}{\sinh\left(\theta/2 - i\pi/3\right)}, \qquad t(\theta) \;=\; \frac{\sinh\left(\theta/2 + i\pi/6\right)}{\sinh\left(\theta/2 - i\pi/6\right)}. \tag{18.6.3}$$

Notice that the antiparticle $\overline{A}$ appears as a bound state of the particle A and vice versa.

Table 18.7: *Dynkin diagram of E_6, with the relative association between the masses and the dots of the diagram.*



18.6.1 Thermal Deformation of the Three-state Tricritical Potts Model

The tricritical version of the three-state Potts model can be identified with a subset of the fields of the minimal conformal model $\mathcal{M}_{6,7}$. As in the ordinary Potts model, its tricritical version is invariant under the permutation group S_3. Its thermal deformation is implemented by adding to the conformal action the energy operator $\Phi_{1,2}$ with conformal weights $(\Delta, \bar{\Delta}) = \left(\frac{1}{7}, \frac{1}{7}\right)$. This is the most relevant field of the Kac table that is invariant under the S_3 symmetry.

The off-critical model is integrable. To compute the S-matrix, let's assume the existence of two douplets $(A_a, A_{\bar{a}})$ and $(A_b, A_{\bar{b}})$ with the bootstrap fusions

$$A_a \times A_a \to A_{\bar{a}} + A_{\bar{b}}, \qquad A_b \times A_b \to A_{\bar{a}} + A_{\bar{b}},$$

and masses m_a, m_b $(m_a < m_b)$. From the analysis of the consistency equations done in the previous chapter, one arrives at the resonance angles

$$\overline{U}^{a}_{a\bar{b}} = \frac{\pi}{12}, \qquad \overline{U}^{\bar{b}}_{a\bar{b}} = \frac{5\pi}{12}, \qquad \overline{U}^{a}_{a\bar{a}} = \frac{\pi}{3}. \tag{18.6.4}$$

Also in this case all reflection amplitudes vanish. The scattering amplitudes relative to the doublet with lower mass $(A_a, A_{\bar{a}})$ are given by

$$\begin{aligned} \mid A_a(\theta_1)A_a(\theta_2)\rangle &= S_{aa}(\theta_{12}) \mid A_a(\theta_2)A_a(\theta_1)\rangle; \\ \mid A_a(\theta_1)A_{\bar{a}}(\theta_2)\rangle &= S^T_{a,\bar{a}}(\theta_{12}) \mid A_a(\theta_2)A_{\bar{a}}(\theta_1)\rangle. \end{aligned}$$

The bootstrap fusion $a \times a \to \overline{a}$ implies

$$\begin{aligned} S^T_{a\bar{a}}(\theta) &= S_{aa}(\theta - i\tfrac{\pi}{3})S_{aa}(\theta + i\tfrac{\pi}{3}) \\ S_{aa}(\theta) &= S^T_{a\bar{a}}(\theta - i\tfrac{\pi}{3})S^T_{a\bar{a}}(\theta + i\tfrac{\pi}{3}). \end{aligned}$$

Equivalently

$$S_{aa}(\theta)S_{aa}\left(\theta - i\frac{2\pi}{3}\right)S_{aa}\left(\theta + i\frac{2\pi}{3}\right) = 1.$$

The minimal solution of these equations, which satisfies the unitarity condition, is

$$S_{aa}(\theta) \;=\; \frac{\sinh(\frac{\theta}{2}+i\frac{\pi}{3})\sinh(\frac{\theta}{2}+i\frac{\pi}{12})\sinh(\frac{\theta}{2}+i\frac{\pi}{4})}{\sinh(\frac{\theta}{2}-i\frac{\pi}{3})\sinh(\frac{\theta}{2}-i\frac{\pi}{12})\sinh(\frac{\theta}{2}-i\frac{\pi}{3})} \equiv s_{\frac{2}{3}}(\theta)s_{\frac{1}{6}}(\theta)s_{\frac{1}{2}}(\theta). \qquad (18.6.5)$$

S_{aa} has two simple poles with positive residue: the first, $\theta = i2\pi/3$, corresponds to the particle $A_{\overline{a}}$ while the other, at $\theta = i\pi/6$, corresponds to the particle $A_{\overline{b}}$. Their mass ratio is

$$m_{\overline{b}} \;=\; m_b \;=\; 2m_a\,\cos\left(\frac{\pi}{12}\right).$$

The additional pole at $\theta = i\pi/2$ has negative residue and corresponds to a bound state in the crossed channel. In fact,

$$S^T_{a\overline{a}}(\theta) \;=\; S_{aa}(i\pi-\theta) = -s_{\frac{1}{3}}(\theta)s_{\frac{1}{2}}(\theta)s_{\frac{5}{6}}(\theta)$$

which has a simple pole with positive residue at $\theta = i\pi/2$. This pole is associated to a new neutral particle A_c, with mass

$$m_c = 2m_a\,\cos\left(\frac{\pi}{4}\right).$$

The scattering amplitute $S_{\overline{a}b}$ is recovered by the equation

$$S_{\overline{a}b}(\theta) \;=\; S_{\overline{a}a}\left(\theta - i\frac{\pi}{12}\right) S_{\overline{a}b}\left(\theta + i\frac{\pi}{12}\right)$$

with the result

$$S_{\overline{a}b}(\theta) \;=\; s_{\frac{3}{4}}(\theta)\, s_{\frac{1}{4}}(\theta)\, s_{\frac{1}{12}}(\theta)\, s_{\frac{5}{12}}(\theta)\, s^2_{\frac{7}{12}}(\theta). \qquad (18.6.6)$$

The pole analysis of $S_{\overline{a}b}$ shows an additional neutral particle A_d that enters the bootstrap fusion

$$A_{\overline{a}} \times A_b \to A_c + A_d$$

with mass

$$m_d \;=\; 4m_a\,\cos\left(\frac{\pi}{12}\right)\cos\left(\frac{\pi}{4}\right).$$

It is easy to show that the set of these six particles $\{A_a, A_{\overline{a}}, A_b, A_{\overline{b}}, A_c, A_d\}$ closes the bootstrap procedure. The masses can be put in correspondence with the dots of the Dynkin diagram of E_6, as shown in Table 18.8. The complete set of the scattering amplitudes is

$$\begin{array}{ll}
S_{aa} = (\frac{1}{6})(\frac{2}{3})(\frac{1}{2}), & S_{\overline{a}\overline{a}} = S_{aa} \\
S^T_{a\overline{a}} = -(\frac{1}{3})(\frac{5}{6})(\frac{1}{2}) & S_{ab} = S_{\overline{a}\overline{b}} = (\frac{1}{4})(\frac{3}{4})(\frac{7}{12})(\frac{11}{12})(\frac{5}{12})^2 \\
S_{a\overline{b}} = S_{\overline{a}b} = (\frac{1}{12})(\frac{1}{4})(\frac{3}{4})(\frac{5}{12})(\frac{7}{12})^2, & S_{ac} = S_{\overline{a}c} = (\frac{1}{4})(\frac{3}{4})(\frac{5}{12})(\frac{7}{12}) \\
S_{ad} = S_{\overline{a}d} = (\frac{1}{6})(\frac{5}{6})(\frac{1}{3})^2(\frac{2}{3})^2(\frac{1}{2})^2, & S_{bb} = (\frac{5}{6})(\frac{1}{6})^2(\frac{1}{3})^2(\frac{1}{6})^2(\frac{2}{3})^3(\frac{1}{2})^3 \\
S^T_{b\overline{b}} = -(\frac{1}{6})(\frac{5}{6})^2(\frac{2}{3})^2(\frac{1}{3})^3(\frac{1}{2})^3, & S_{bc} = S_{\overline{b}c} = (\frac{1}{6})(\frac{5}{6})(\frac{1}{2})^2(\frac{2}{3})^2(\frac{1}{3})^2 \\
S_{b\overline{d}} = S_{\overline{b}d} = (\frac{1}{12})(\frac{11}{12})(\frac{1}{4})^3(\frac{1}{4})^3(\frac{5}{12})^4(\frac{7}{12})^4, & S_{cc} = -(\frac{1}{6})(\frac{5}{6})(\frac{1}{3})(\frac{2}{3})(\frac{1}{2})^2 \\
S_{c\overline{d}} = (\frac{1}{12})(\frac{11}{12})(\frac{1}{4})^2(\frac{3}{4})^2(\frac{5}{12})^3(\frac{7}{12})^3, & S_{dd} = -(\frac{1}{6})^3(\frac{5}{6})^3(\frac{1}{3})^5(\frac{2}{3})^5(\frac{1}{2})^6,
\end{array}$$

where we use the notation

$$(x) \equiv s_x(\theta).$$

E_6 **Toda theory**. The amplitudes above are the minimal S-matrices of the Toda field theory based on the exceptional algebra E_6. This is not surprising, because of the relation between this statistical model and the Toda field theory discussed in Chapter 16. To obtain the exact S-matrix of the Toda field theory with real coupling constant g it is sufficient to multiply the minimal amplitude $S_{aa}(\theta)$ for the Z-factor

$$Z(\theta) \,=\, (-B)\,\left(-\frac{1}{6}+B\right)\left(-\frac{1}{2}-B\right)\left(-\frac{2}{3}+B\right), \tag{18.6.7}$$

where

$$B(g) \,=\, \frac{2}{h}\,\frac{g^2}{8\pi}\,\frac{1}{1+\frac{g^2}{8\pi}} \tag{18.6.8}$$

and $h = 12$ is the Coxeter number of E_6.

18.7 Models with Internal $O(n)$ Invariance

The $O(n)$ statistical models are characterized by an isotropic ferromagnetic interaction among the n components of the spin variables $\vec{S}_i$. For the elastic S-matrix of these theories it is necessary to distinguish three cases: (i) $n > 2$; (ii) $n < 2$; and (iii) $n = 2$. In this section we discuss the first two cases, whereas the discussion of the $n = 2$ case can be found in the next section.

18.7.1 $n > 2$

From the symmetry of the system, let's assume that the spectrum of the theory consists of a multiplet of n particles of equal mass, denoted by the symbols A_i $(i = 1, 2, \ldots, n)$. Enforcing the $O(n)$ invariance of the scattering theory, we can decompose the S-matrix elements as

$$\begin{aligned} A_i(\theta_1)A_j(\theta_2) \,=\, \delta_{ij}\,S_1(\theta)\sum_{k=1}^{n} A_k(\theta_2)A_k(\theta_1) \\ +S_2(\theta)\,A_j(\theta_2)A_i(\theta_1) + S_3(\theta)\,A_i(\theta_2)A_i(\theta_1). \end{aligned} \tag{18.7.1}$$

The functions $S_2(\theta)$ and $S_3(\theta)$ are the transmission and reflection amplitudes respectively, while $S_1(\theta)$ describes the annihilation–creation process $A_i + A_i \to A_j + A_j$, with $i \neq j$. This decomposition is represented in Fig. 18.4. These functions satisfy the unitarity equation

$$\begin{aligned} &S_2(\theta)S_2(-\theta) + S_3(\theta)S_3(-\theta) \,=\, 1 \\ &S_2(\theta)S_3(-\theta) + S_3(\theta)S_2(-\theta) \,=\, 0 \\ &n\,S_1(\theta)S_1(-\theta) + S_1(\theta)S_2(-\theta) + S_1(\theta)S_3(-\theta) \\ &+S_2(\theta)S_1(-\theta) + S_3(\theta)S_1(-\theta) \,=\, 0. \end{aligned} \tag{18.7.2}$$

Moreover, they are related by the crossing symmetry relationships

$$S_2(\theta) \,=\, S_2(i\pi-\theta) \tag{18.7.3}$$

$$S_1(\theta) \,=\, S_3(i\pi-\theta) \tag{18.7.4}$$

as can be seen by looking at the diagrams of Fig. 18.4 from left to right rather than from bottom to top.

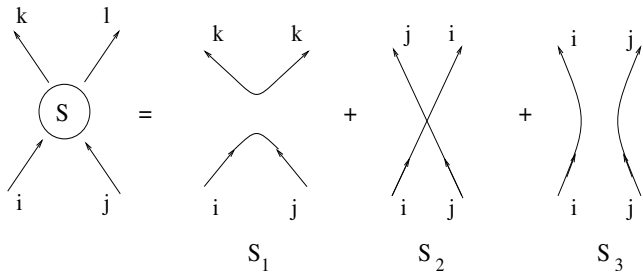


Fig. 18.4 Decomposition of the S-matrix into invariant amplitudes under the $O(n)$ group.

In addition to these basic conditions, the amplitudes satisfy a non-trivial set of Yang–Baxter equations:

$$\begin{aligned} S_2\,S_1\,S_3 + S_2\,S_3\,S_3 + S_3\,S_3\,S_2 &= S_3\,S_2\,S_3 + S_1\,S_2\,S_2 + S_1\,S_1\,S_2 S_3\,S_1\,S_3 + S_3\,S_2\,S_3 \\ &= S_3\,S_3\,S_1 + S_3\,S_3\,S_2 + S_3\,S_3\,S_1 + S_2\,S_2\,S_3 + 2S_1\,S_3\,S_1 \\ &\quad + S_1\,S_3\,S_2 + S_1\,S_3\,S_3 + S_1\,S_2\,S_2 + S_1\,S_1\,S_1 \end{aligned}$$

where the arguments in each generic term $S_a S_b S_c$ of these equations are θ for the first factor S_a, $\theta + \theta'$ for the second factor S_b, and θ' for the third one S_c. The general solution of these equations has the functional form

$$\begin{aligned} S_3(\theta) &= -\frac{i\lambda}{\theta}\, S_2(\theta), \\ S_1(\theta) &= -\frac{i\lambda}{i[(n-2)/2]\lambda - \theta}\, S_2(\theta). \end{aligned} \tag{18.7.5}$$

Substituting these expressions into the crossing equations, one can determine the parameter λ

$$\lambda = \frac{2\pi}{n-2}. \tag{18.7.6}$$

Substituting in the unitarity equations, one arrives at the condition

$$S_2(\theta)\, S_2(-\theta) = \frac{\theta^2}{\theta^2 + \lambda^2}. \tag{18.7.7}$$

In order to solve this equation, together with (18.7.3) coming from the crossing symmetry, one can follow an iterative strategy. Notice that a solution of (18.7.7) is given by

$$Q(\theta) = \frac{\theta}{\theta + i\lambda}.$$

However this spoils the crossing symmetry equation (18.7.3), which can however be re-established by writing

$$Q(\theta) = \frac{\theta}{\theta + i\lambda}\,\frac{i\pi - \theta}{i\pi - \theta + i\lambda}.$$

In turn, this new expression spoils the unitarity condition (18.7.7), which can be saved by rewriting $Q(\theta)$ as

$$Q(\theta) \;=\; \frac{\theta}{\theta+i\lambda}\,\frac{i\pi-\theta}{i\pi-\theta+i\lambda}\,\frac{i\pi+\theta+i\lambda}{i\pi+\theta}.$$

Iterating these two operations to satisfy simultaneously the unitarity and crossing symmetry equations, one ends up in an infinite product. Using the identity

$$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\gamma)\Gamma(\beta-\gamma)} \;=\; \prod_{k=0}^{\infty}\left[\left(1+\frac{\gamma}{\alpha+k}\right)\left(1-\frac{\gamma}{\beta+k}\right)\right]$$

the final result can be concisely expressed in terms of the Γ functions as

$$\begin{aligned} S_2(\theta) \;&=\; \mathcal{U}^{(+)}(\theta)\,\mathcal{U}^{(+)}(i\pi-\theta); \\ \mathcal{U}^{(+)}(\theta) \;&=\; \frac{\Gamma\left(\frac{\lambda}{2\pi}-i\frac{\theta}{2\pi}\right)\Gamma\left(\frac{1}{2}-i\frac{\theta}{2\pi}\right)}{\Gamma\left(\frac{1}{2}+\frac{\lambda}{2\pi}-i\frac{\theta}{2\pi}\right)\Gamma\left(-i\frac{\theta}{2\pi}\right)}. \end{aligned} \tag{18.7.8}$$

With the determination of this amplitude, one can use eqn (18.7.5) to determine the remaining two amplitudes. The S-matrix so obtained does not have a pole in the physical sheet. Hence the theory does not present additional bound states and the only excitations are given by the original particle A_i. It is possible to show that this scattering theory is in agreement with the perturbative computations done using the bosonic lagrangian

$$\mathcal{L} \;=\; \frac{1}{2}\partial_\mu\vec{S}\cdot\partial^\mu\vec{S}, \qquad |\vec{S}| \;=\; 1 \tag{18.7.9}$$

This is a nonlinear σ-model: although the lagrangian looks like that of a free massless theory, the constraint on the components of the field induces, ipso facto, a mass term and a series of interactions. The nonlinear σ-model is renormalizable, asymptotically free and explicitly $O(n)$ symmetric. The simplest way of showing the mass generation in this theory is to study the large n limit: introducing a coupling constant g and enforcing the constraint using the Fourier representation of the δ function, we can write the lagrangian of the model as

$$\mathcal{L} \;=\; \frac{n}{2g}\left[(\partial_\mu\vec{S})^2+i\lambda(x)\,(\vec{S}^2-1)\right].$$

In this expression $\lambda(x)$ is the lagrangian multiplier field associated to the constraint and furthermore we have parameterized the coupling constant in such a way as to have a factor n in front of the lagrangian. In the path integral of this theory we can now integrate out $\vec{S}$, which is no longer constrained, obtaining an effective action for the field $\lambda(x)$:

$$\mathcal{S}_{eff}(\lambda) \;=\; \frac{n}{2}\left[-\int d^2x\,\left(i\frac{\lambda(x)}{g}+\mathrm{tr}\,\log(-\partial^2+i\lambda)\right)\right]. \tag{18.7.10}$$

Because of the presence of a factor n in front of $\mathcal{S}_{eff}$, in the large n-limit we can ignore the fluctuations of $\lambda(x)$ and evaluate it at the action saddle point. This can be

done by deforming the functional integration contour of λ into the complex plane: a saddle point is found at a constant, imaginary value of $\lambda = i\lambda_0$. Imposing $\lambda_0 = m^2$, the saddle point equation is expressed by

$$\frac{1}{g} = \int \frac{d^2k}{(2\pi)^2} \frac{1}{k^2 + m^2} = \frac{1}{2\pi} \log \frac{\Lambda}{m}$$

where Λ is an ultraviolet cut-off. This equation determines the mass parameter m of the theory, in terms of the cut-off and the bare coupling g:

$$m = \Lambda\, e^{-2\pi/g}.$$

At lowest order in $1/n$, the theory consists of just n free boson particles of mass m. This results is consistent with the S-matrix formulation given above.

18.7.2 $n < 2$

As seen in Section 14.6, for $-2 < n < 2$ the $O(n)$ model has a critical point. The conformal theory has central charge

$$c = 1 - \frac{6}{p(p+1)},$$

where p is a function of the index n

$$n = 2 \cos\left(\frac{\pi}{p}\right). \tag{18.7.11}$$

In the continuum limit, the energy operator corresponds to the primary field $\Phi_{1,3}$. Correspondingly, the off-critical theory

$$\mathcal{S} = \mathcal{S}_{CFT} + \tau \int \Phi_{1,3}(x)\, d^2x.$$

defines an integrable theory. Let's compute the S-matrix based on the following assumptions. As in the previous case, the particles A_i are associated to a vector representation of $O(n)$, even though n can now take continuous values. Their S-matrix can still be decomposed into the invariant amplitudes, as done in eqn (18.7.1). However, in this range of values of n, there is also $n = 0$ which corresponds to self-avoiding random walks. Associating the world-lines of the particles A_i to the lines that enter the high-temperature expansion of the lattice model, it seems natural to conjecture that for all n in the interval $-2 < n < 2$ we have the condition

$$S_2(\theta) = 0. \tag{18.7.12}$$

Notice that this equation is consistent with the crossing symmetry of the problem, while for the other amplitudes we have

$$S_1(\theta) = S_3(i\pi - \theta). \tag{18.7.13}$$

The general solution of the Yang–Baxter equation is given in this case by

$$\begin{aligned} S_1(\theta) &= i \sinh\left(\tfrac{\theta}{p}\right) R(\theta); \\ S_3(\theta) &= i \sinh\left(\tfrac{i\pi-\theta}{p}\right) R(\theta), \end{aligned} \tag{18.7.14}$$

where $R(\theta)$ is a function that satisfies

$$R(\theta) = R(i\pi - \theta). \tag{18.7.15}$$

It can be determined by imposing the unitarity equations

$$\begin{aligned} &S_3(\theta)\, S_3(-\theta) = 1 \\ &S_3(\theta)\, S_1(-\theta) + S_1(\theta)\, S_3(-\theta) + n\, S_1(\theta)\, S_1(-\theta) = 0. \end{aligned} \tag{18.7.16}$$

Notice that the second equation of this system is authomatically satisfied by virtue of (18.7.11). The first equation implies instead

$$R(\theta)\, R(-\theta) = -\left[\sinh\left(\frac{i\pi-\theta}{p}\right) \sinh\left(\frac{i\pi+\theta}{p}\right)\right]^{-1}. \tag{18.7.17}$$

The minimal solution of both eqns (18.7.15) and (18.7.17) is given by

$$\begin{aligned} R(\theta) = {} & \frac{1}{\sinh\left(\frac{i\pi-\theta}{p}\right)} \frac{\Gamma\left(1-\frac{\theta}{i\pi p}\right)}{\Gamma\left(1+\frac{\theta}{i\pi p}\right)} \\ & \times \prod_{k=1}^{\infty} \frac{\Gamma\left(\frac{2k}{p}-\frac{\theta}{i\pi p}\right)\Gamma\left(1+\frac{2k}{p}-\frac{\theta}{i\pi p}\right)\Gamma\left(\frac{2k-1}{p}+\frac{\theta}{i\pi p}\right)\Gamma\left(1+\frac{2k-1}{p}+\frac{\theta}{i\pi p}\right)}{\Gamma\left(\frac{2k}{p}+\frac{\theta}{i\pi p}\right)\Gamma\left(1+\frac{2k}{p}+\frac{\theta}{i\pi p}\right)\Gamma\left(\frac{2k-1}{p}-\frac{\theta}{i\pi p}\right)\Gamma\left(1+\frac{2k-1}{p}-\frac{\theta}{i\pi p}\right)}. \end{aligned} \tag{18.7.18}$$

As shown in Problem 6, the infinite product admits an integral representation, so that $R(\theta)$ can also be expressed as

$$R(\theta) = \frac{1}{\sinh\left(\frac{i\pi-\theta}{p}\right)} \exp\left[i \int_0^{\infty} \frac{dt}{t} \frac{\sinh\frac{\pi(p-1)t}{2}}{\sinh\frac{\pi p t}{2}\cosh\frac{\pi t}{2}} \sin\theta t\right]. \tag{18.7.19}$$

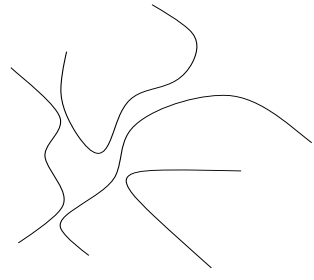


Fig. 18.5 *Configurations of self-avoiding polymers associated to the world-lines of the particles of the $O(n)$ scattering theory for $-2 < n < 2$.*

Notice that for $n = 1$

$$S_1(\theta) + S_3(\theta) \;=\; -1, \tag{18.7.20}$$

which coincides with the S-matrix of the thermal Ising model, as it should do. In the limit $n \to 0$, the two amplitudes can be interpreted as the two possible interactions among the chains of the polymers, as shown in Fig. 18.5.

18.8 S-Matrix of the Sine–Gordon Model

For $n = 2$ the scattering theory of the $O(n)$ model coincides with that associated to the Sine–Gordon model, whose lagrangian is

$$\mathcal{L}_{SG} \;=\; \frac{1}{2}(\partial_\mu \phi)^2 + \frac{m^2}{\beta^2}(\cos\beta\phi - 1). \tag{18.8.1}$$

It is useful to define

$$\xi \;\equiv\; \frac{\beta^2}{8}\,\frac{1}{1-\frac{\beta^2}{8\pi}}, \tag{18.8.2}$$

a quantity that plays the role of the renormalized coupling constant. It is worth mentioning that Sidney Coleman discovered that the quantum Sine–Gordon model is equivalent to the massive Thirring model for a Dirac field

$$\mathcal{L}_{MTM} \;=\; i\bar\psi\gamma^\mu\partial_\mu\psi - m_0\bar\psi\psi - \frac{g}{2}(\bar\psi\gamma^\mu\psi)^2, \tag{18.8.3}$$

with the position

$$\frac{g}{\pi} \;=\; \frac{4\pi}{\beta^2} - 1\,, \qquad \text{for} \qquad 0 \le \frac{\beta^2}{8\pi} < 1. \tag{18.8.4}$$

In this mapping, the Sine–Gordon soliton is identified with the Thirring fermion. Note that the Thirring interaction becomes

$$\begin{array}{llllll} \text{attractive} & (g>0) & \text{for} & \beta^2 < 4\pi & \text{i.e.} & \xi < \pi \\ \text{repulsive} & (g<0) & \text{for} & \beta^2 > 4\pi & \text{i.e.} & \xi > \pi. \end{array} \tag{18.8.5}$$

It is important to keep this in mind for understanding the structure of the bound states of the Sine–Gordon model that will be discussed below.

The best way to reveal the structure of the scattering theory of this model is to define its basic excitations by the complex linear combinations

$$A(\theta) \;=\; A_1(\theta) + iA_2(\theta), \qquad \bar A(\theta) \;=\; A_1(\theta) - iA_2(\theta)$$

where A_1 and A_2 are the degenerate particles of the original $O(2)$ model. In terms of these new excitations, the scattering amplitudes can be written as

$$\begin{aligned} A(\theta_1)\bar A(\theta_2) &= S_T(\theta)\bar A(\theta_2)A(\theta_1) + S_R(\theta)A(\theta_2)\bar A(\theta_1) \\ A(\theta_1)A(\theta_2) &= S(\theta)A(\theta_2)A(\theta_1) \\ \bar A(\theta_1)\bar A(\theta_2) &= S(\theta)\bar A(\theta_2)\bar A(\theta_1). \end{aligned} \tag{18.8.6}$$

They can be collected into a 4×4 matrix with non-zero entries given by

$$S^{SG}(\theta) = \begin{pmatrix} S & & & \\ & S_T & S_R & \\ & S_R & S_T & \\ & & & S \end{pmatrix}. \tag{18.8.7}$$

The quantities above can be interpreted as the scattering amplitudes of the soliton A and the antisoliton $\bar{A}$. S_T and S_R are the transmission and reflection amplitudes, respectively, in the soliton–antisoliton scattering process, while S, by charge conjugation symmetry, is the common transmission amplitude in the soliton–soliton and antisoliton–antisoliton scatterings. Notice the close structure between the S-matrix (18.8.7) and the R-matrix of the six-vertex model given in eqn (6.4.3).

The amplitudes satisfy the crossing symmetry equations

$$S(\theta) = S_T(i\pi - \theta), \qquad S_R(\theta) = S_R(i\pi - \theta) \tag{18.8.8}$$

and those coming from the unitarity condition

$$\begin{aligned} &S(\theta)S(-\theta) = 1, \\ &S_T(\theta)S_T(-\theta) + S_R(\theta)S_R(-\theta) = 1, \\ &S_T(\theta)S_R(-\theta) + S_R(\theta)S_T(-\theta) = 0. \end{aligned} \tag{18.8.9}$$

Using the Yang–Baxter equations satisfied by the amplitudes, they can be expressed as

$$\begin{aligned} S_T(\theta) &= \frac{\sinh\frac{\pi\theta}{\xi}}{\sinh\frac{\pi(i\pi-\theta)}{\xi}}\, S(\theta), \\ S_R(\theta) &= i\frac{\sin\frac{\pi^2}{\xi}}{\sinh\frac{\pi(i\pi-\theta)}{\xi}}\, S(\theta). \end{aligned} \tag{18.8.10}$$

Substituting them into the unitarity and crossing symmetry equations, we get the equations satisfied by $S(\theta)$:

$$\begin{aligned} &S(\theta)S(-\theta) = 1, \\ &S(i\pi - \theta) = \frac{\sinh\frac{\pi\theta}{\xi}}{\sinh\frac{\pi(i\pi-\theta)}{\xi}}\, S(\theta). \end{aligned}$$

Its solution can be written in terms of an infinite product

$$\begin{aligned} S(\theta) = \prod_{k=0}^{\infty} &\frac{\Gamma\left(1+(2k+1)\frac{\pi}{\xi} - i\frac{\theta}{\xi}\right)\Gamma\left(1+2k\frac{\pi}{\xi} + i\frac{\theta}{\xi}\right)}{\Gamma\left(1+(2k+1)\frac{\pi}{\xi} + i\frac{\theta}{\xi}\right)\Gamma\left(1+2k\frac{\pi}{\xi} - i\frac{\theta}{\xi}\right)} \\ &\times\frac{\Gamma\left((2k+1)\frac{\pi}{\xi} - i\frac{\theta}{\xi}\right)\Gamma\left((2k+2)\frac{\pi}{\xi} + i\frac{\theta}{\xi}\right)}{\Gamma\left((2k+1)\frac{\pi}{\xi} + i\frac{\theta}{\xi}\right)\Gamma\left((2k+2)\frac{\pi}{\xi} - i\frac{\theta}{\xi}\right)}. \end{aligned} \tag{18.8.11}$$

This expression admits the integral representation

$$S(\theta) \;=\; -\exp\left[-i\int_0^\infty \frac{dt}{t}\,\frac{\sinh\frac{t(\pi-\xi)}{2}}{\sinh\frac{\xi t}{2}\cosh\frac{\pi t}{2}}\,\sin\theta t\right]. \tag{18.8.12}$$

Another useful representation of this amplitude is the mixed one

$$S(\theta) = -(-1)^n \prod_{k=1}^{n}\left[\frac{\theta+ik\xi}{\theta-ik\xi}\right] \tag{18.8.13}$$
$$\times\exp\left\{-i\int_0^\infty \frac{dt}{t}\,\frac{\left[2\sinh\frac{t(\pi-\xi)}{2}\,e^{-n\xi t}+\left(e^{-n\xi t}-1\right)\left(e^{(\xi-\pi)t/2}+e^{-(\pi+\xi)t/2}\right)\right]}{2\sinh\frac{\xi t}{2}\cosh\frac{\pi t}{2}}\sin\theta t\right\}.$$

The mixed representation is particularly helpful for determining the numerical values of $S(\theta)$: notice that the convergence of the integral increases by increasing the integer n, with the only price to pay of having more power factors in the first product term. Note that the integer n can be varied arbitrarily without changing the value of $S(\theta)$ and, in particular, for $n = 0$ one recovers the previous expression (18.8.12). The proof of the integral and the mixed representation of $S(\theta)$ is suggested in Problem 6.

The pole structure of the S-matrix is determined by the various terms that enter its expression. It is important to focus attention on the poles that may belong to the physical sheet $0 < \theta < i\pi$: using the results of Appendix A in Chapter 2 for the Γ functions or simply looking at the mixed representation (18.8.13), it is easy to see that $S(\theta)$ has a set of poles at

$$\theta = i\,n\,\xi, \quad n = 0, 1, \ldots \tag{18.8.14}$$

Other poles of the S-matrix come from the factor $\sinh(\pi(i\pi-\theta)/\xi)$ in the denominator of the right-hand side of eqn (18.8.10), placed at

$$\theta \;=\; i(\pi - n\,\xi), \quad n = 0, 1, \ldots \tag{18.8.15}$$

Both sets of poles fall in the physical sheet if

$$\xi \;<\; \pi. \tag{18.8.16}$$

As we show below, if the condition (18.8.16) is satisfied, the poles (18.8.15) lead to the bound states in the s-channel of the SG model whereas the poles (18.8.14) lead to the bound states of the crossed t-channel. The number of these bound states is

$$\bar{N} \;=\; \left[\frac{\pi}{\xi}\right], \tag{18.8.17}$$

where $[x]$ is the integer part of the number x. To support the interpretation of the poles given above, it is convenient to define the amplitudes

$$S_\pm(\theta) \;=\; (S_R \pm S_T)(\theta). \tag{18.8.18}$$

These quantities correspond to scattering processes with a well-defined quantum number under the charge conjugation operator: S_- has charge conjugation $C = -1$ while S_+ has $C = +1$, as can be seen by writing the scattering processes as

$$\begin{aligned}\left[A(\theta_1)\bar{A}(\theta_2) + \bar{A}(\theta_1)A(\theta_2)\right] &= S_+(\theta)\,\left[A(\theta_2)\bar{A}(\theta_1) + \bar{A}(\theta_2)A(\theta_1)\right],\\ \left[A(\theta_1)\bar{A}(\theta_2) - \bar{A}(\theta_1)A(\theta_2)\right] &= S_-(\theta)\,\left[A(\theta_2)\bar{A}(\theta_1) - \bar{A}(\theta_2)A(\theta_1)\right].\end{aligned}$$

The explicit expressions of these amplitudes are

$$S_\pm(\theta) \;=\; -\frac{1}{\sinh\frac{\pi(\theta - i\pi)}{\xi}}\,\left[i\sin\frac{\pi^2}{\xi} \pm \sinh\frac{\pi\theta}{\xi}\right]\,S(\theta), \tag{18.8.19}$$

and their residue at the poles (18.8.15) is

$$S_\pm(\theta) \;\simeq\; -\frac{i}{\theta - i\pi + in\xi}\,(-1)^n\,\xi\,\sin\frac{\pi^2}{\xi}\,\left[1 \pm (-1)^n\right]\,S(i\pi - in\xi) \tag{18.8.20}$$

Hence, $S_+(\theta)$ has poles only when n is an even number, with $S_-(\theta)$ only when n is an odd number. Both sets of poles have a positive residue[3] and therefore, as anticipated, they correspond to the poles of the s-channel associated to the bound states B_n. These are ordinary scalar particles, called *breathers*, with eigenvalues $C = (-1)^n$ under charge conjugation. If M is the mass of the solitons, the mass spectrum of the bound states is given by

$$m_n \;=\; 2M\,\sin\frac{k\xi}{2}, \quad k = 1, 2, \ldots, \bar{N}. \tag{18.8.21}$$

The S-matrix elements that involve the bound states can be computed by the bootstrap equations. For their scattering processes with the solitons

$$\begin{aligned}A(\theta_1)B_n(\theta_2) &= S^{(n)}(\theta)\,B_n(\theta_2)\,A(\theta_1),\\ \bar{A}(\theta_1)B_n(\theta_2) &= S^{(n)}(\theta)\,B_n(\theta_2)\,\bar{A}(\theta_1),\end{aligned} \tag{18.8.22}$$

we have

$$S^{(n)}(\theta) \;=\; \frac{\sinh\theta + i\cos\frac{n\xi}{2}}{\sinh\theta - i\cos\frac{n\xi}{2}}\,\prod_{k=1}^{n-1}\frac{\sin^2\left(\frac{(n-2k)\xi}{4} - \frac{\pi}{4} + i\frac{\theta}{2}\right)}{\sin^2\left(\frac{(n-2k)\xi}{4} - \frac{\pi}{4} - i\frac{\theta}{2}\right)}. \tag{18.8.23}$$

For the scattering processes that involve only the particles B_n

$$B_n(\theta_1)B_m(\theta_2) \;=\; S^{(n,m)}(\theta)\,B_m(\theta_2)B_n(\theta_1), \tag{18.8.24}$$

[3]The mixed representation (18.8.13) is particularly useful in the evaluation of these residues. Attention has to be paid when $\xi = \pi/m$, with m an integer, for the simultaneous presence of a pole in $S(i\pi - in\xi)$ and a zero in $\sin\pi^2/\xi$.

we have the amplitudes

$$S^{(n,m)}(\theta) = \frac{\sinh\theta + i\sin\left(\frac{(n+m)\xi}{2}\right)}{\sinh\theta - i\sin\left(\frac{(n+m)\xi}{2}\right)} \frac{\sinh\theta + i\sin\left(\frac{(n-m)\xi}{2}\right)}{\sinh\theta - i\sin\left(\frac{(n-m)}{\xi}2\right)} \times \prod_{k=1}^{n-1} \frac{\sin^2\left(\frac{(m-n-2k)\xi}{4} + i\frac{\theta}{2}\right)}{\sin^2\left(\frac{(m-n-2k)\xi}{4} - i\frac{\theta}{2}\right)} \frac{\cos^2\left(\frac{(m+n-2k)\xi}{4} + i\frac{\theta}{2}\right)}{\cos^2\left(\frac{(m+n-2k)\xi}{4} - i\frac{\theta}{2}\right)} \tag{18.8.25}$$

with $n \geq m$. From the analysis of the poles of these expressions it is easy to see that the particle B_n can be regarded as a bound state of $B_k + B_l$, with $k+l = n$. Consequently, iterating this relation, the particle B_n can be seen as the bound state of n elementary particles B_1. The lowest particle B_1 can be associated to the excitation created by the field ϕ that enters the lagrangian (18.8.1). When $n = m = 1$, the amplitude of the fundamental particle is given by

$$S^{(1,1)}(\theta) = \frac{\sinh\theta + i\sin\xi}{\sinh\theta - i\sin\xi}. \tag{18.8.26}$$

With the expression for ξ given in (18.8.2), this amplitude can be expanded in power series of β^2 and successfully compared with the perturbative series coming from the Lagrangian (18.8.1). Notice that making the analytic continuation $\beta \to ig$ the Sine–Gordon model becomes the Sinh–Gordon model: the formula (18.8.2) given for ξ becomes the expression (18.4.5) previously obtained for the function $B(g)$ of the latter model, while the amplitude (18.8.26) reduces to the S-matrix (18.4.4) of the Sinh–Gordon model.

Let's close this section with a comment that the reader should reflect upon. Notice that when $\xi > \pi$, the pole in the soliton–antisoliton amplitude falls outside the physical sheet. Correspondingly, there is no longer a bound state B_1 associated to the field ϕ, despite the fact that the lagrangian is expressed in terms of this field! This observation shows that the spectrum of a quantum field theory is a question of dynamical nature and less intuitive than it would appear.

18.9 S-Matrices for $\Phi_{1,3}$, $\Phi_{1,2}$, $\Phi_{2,1}$ Deformation of Minimal Models

As seen previously, the $\Phi_{1,3}$, $\Phi_{1,2}$, and $\Phi_{2,1}$ deformations of the minimal models $\mathcal{M}_m$ of CFT generally lead to integrable massive field theories with kink behavior. This means that such deformations give rise to an effective potential of the theory with a finite number of degenerate vacua. The basic massive excitations are the kinks that interpolate between different vacua. There may also be kink bound states.

There is a general approach for dealing with such massive theories, deeply related to the Sine–Gordon model (for the $\Phi_{1,3}$ deformation) and to the Bullogh–Dodd model with model imaginary coupling (for the $\Phi_{1,2}$ and $\Phi_{2,1}$ deformations) (see Chapter 16). The main idea is based on the well-known relation between the S-matrices and the R-matrices entering the transfer matrix of lattice integrable models (see Sections 6.4 and 17.2). An important feature of both quantities is their invariance under the quantum group $SL_q(2)$. The q-parameter is a function of the coupling constant and when

q is a root of unity, it is possible to restrict the Hilbert space of the original models, preserving both the integrability and the locality of an invariant set of operators. Let's see how this procedure is implemented for the Sine–Gordon model and for the Bullogh–Dodd model.

18.9.1 Quantum Group Symmetry of the Sine–Gordon

The quantum group $SL_q(2)$ is the deformation of the algebra of functions over $SL(2)$. It is defined by the universal enveloping algebra $\mathcal{U}_q[sl(2)]$ with the commutation relations

$$[\, J_+, J_- \,] \;=\; \frac{q^H - q^{-H}}{q - q^{-1}} \;, \qquad [\, H, J_\pm \,] \;=\; \pm 2\, J_\pm . \tag{18.9.1}$$

If the deformation parameter q goes to 1, eqn (18.9.1) reduces to the ordinary $SL(2)$ commutation relations and the quantum group $SL_q(2)$ to the ordinary $SL(2)$ group. $\mathcal{U}_q[sl(2)]$ forms a Hopf algebra with the *comultiplication* Δ_q defined by

$$\begin{aligned} \Delta_q(H) \;&=\; 1 \otimes H + H \otimes 1 \\ \Delta_q(J_\pm) \;&=\; q^{H/2} \otimes J_\pm + J_\pm \otimes q^{-H/2}. \end{aligned} \tag{18.9.2}$$

The comultiplication Δ_q is an algebra homomorphism and is the analog of addition of angular momentum in $SU(2)$, to which it reduces when $q \to 1$. The irreducible representations of $SL_q(2)$ are generated by the comultiplication Δ_q which defines tensor product representations.

Because of the resemblance between the algebraic structure of $SL(2)$ and $\mathcal{U}_q[sl(2)]$, the representation theory of the quantum group is quite similar to the classical theory. The irreducible representations of $\mathcal{U}_q[su(2)]$ are labelled by $j = 0, \frac{1}{2}, 1, \ldots$ acting on a Hilbert space V_j with basis vectors $\mid j, m\rangle$ $(-j \le m \le j)$ as follows:

$$J_3 \mid j, m\rangle \;=\; m \mid j, m\rangle, \qquad J_\pm \mid j, m\rangle \;=\; \sqrt{[j \mp m]_q \, [j \pm m + 1]_q} \mid j, m\rangle$$

where all the usual numbers have turned into q-numbers, so defined:

$$[n]_q \;\equiv\; \frac{q^n - q^{-n}}{q - q^{-1}} \qquad \text{and} \qquad [n]_q \to n \quad \text{as} \quad q \to 1. \tag{18.9.3}$$

All these representations can be obtained by starting with the fundamental representation $j = \frac{1}{2}$ and using eqn (18.9.2), with the relation

$$\mid J, M; j_1, j_2\rangle = \sum_{m_1, m_2} \begin{bmatrix} j_1 & j_2 & J \\ m_1 & m_2 & j \end{bmatrix}_q \mid j_1, m_1\rangle \otimes \mid j_2, m_2\rangle \tag{18.9.4}$$

where the quantum analogue of the Clebsh–Gordan (CG) coefficients appears.[4] To cluster together three representations there are two possibilities: one is related to the configuration $(V_{j_1} \otimes V_{j_2}) \otimes V_{j_3}$, the other one to $V_{j_1} \otimes (V_{j_2} \otimes V_{j_3})$. Both are physically

[4] For the classical values of the Clesh–Gordan and 6-j coefficients, see L.D. Landau and E.M. Lifshitz, *Quantum Mechanics. Non-relativistic Theory*, Pergamon, Oxford, 1991.

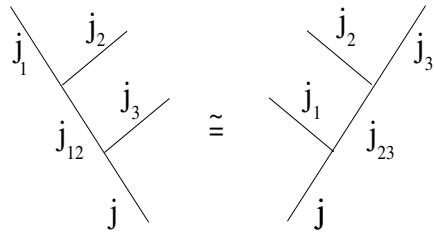


Fig. 18.6 *Equivalence between the two Hilbert spaces* $(V_{j_1} \otimes V_{j_2}) \otimes V_{j_3}$ *(left-hand side) and* $V_{j_1} \otimes (V_{j_2} \otimes V_{j_3})$ *(right-hand side) rules by the* $6-j$ *symbols.*

equivalent and related each to the other by the quantum analogue of the 6-j symbols: denoting by $e_m^{j_{12},j}(j_1, j_2 \mid j_3)$ an orthonormal basis in $V_{j_1} \otimes V_{j_2} \otimes V_{j_3}$ associated to the left-hand side of Fig. 18.6 and by $e_m^{j_{23},j}(j_1 \mid j_2, j_3)$ an orthonormal basis associated to the right-hand side of the same figure, we have

$$e_m^{j_{12},j}(j_1, j_2 \mid j_3) \;=\; \sum_{j_{23}} \left\{ \begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & j & j_{23} \end{matrix} \right\}_q \; e_m^{j_{23},j}(j_1 \mid j_2, j_3).$$

As long as q is not a root of unity, the irreducible representations have dimension $(2j+1)$. However, when q is a root of unity, one can see from eqn (18.9.3) that some of the q-CG coefficients (and the q-6j symbols) become singular. For this case, a sensible representation theory of $SL_q(2)$ is obtained by restricting the allowed spins to $\{0, 1/2, \ldots, j_{max}\}$, where j_{max} is determined by the condition

$$[2j_{max}+1]_q \;=\; 0 \qquad \rightarrow \qquad j_{max} \;=\; \frac{N}{2} - 1 \qquad \text{for} \qquad q^N = \pm 1. \tag{18.9.5}$$

This restriction on the allowed representations of $SU_q(2)$ with q a root of unity leads to the truncation of the Hilbert space. It is this property, in particular, that is responsible for the fusion rules of the minimal models of conformal field theory, discussed in Section 11.4 of Chapter 11.

From the Sinh–Godon model, the restriction on the spins may lead to the truncation of the multikink Hilbert space. Let's see in more detail how this happens. Notice that the quantum group $SL_q(2)$ can be realized by a constant R-matrix defined by

$$R_{12}(q)\,(g \otimes 1)\,(1 \otimes g) \;=\; (1 \otimes g)\,(g \otimes 1)\,R_{12}(q) \qquad \text{with} \qquad g \;\in\; SL_q(2). \tag{18.9.6}$$

Using eqn (18.9.2), this implies $[R(q), \Delta_q(g)] = 0$ for any $g \in SL_q(2)$. In the fundamental representation, g is a 2×2 matrix with q-determinant equal to 1 ($g_{11}g_{22} - q\,g_{12}g_{21} = 1$) and the non-zero entries of the R_{12}-matrix are

$$R_{12}(q) \;=\; \begin{pmatrix} q & & & \\ & 1 & q-q^{-1} & \\ & 0 & 1 & \\ & & & q \end{pmatrix}. \tag{18.9.7}$$

The spectral parameter λ can be introduced in the R-matrix as follows

$$\hat{R}_{12}(\lambda, q) \,=\, \lambda \hat{R}_{12}(q) - \lambda^{-1} \hat{R}_{12}^{-1}(q) \qquad \text{with} \qquad \hat{R}_{12} = \mathcal{P}\, R_{12} \tag{18.9.8}$$

where the permutation operator $\mathcal{P}$ is defined as $\mathcal{P}(V_1 \otimes V_2) = V_2 \otimes V_1$. The matrix $\hat{R}_{12}(p, q)$ acts on the vector space $\mathbf{C}^2 \otimes \mathbf{C}^2$ and is a solution of the Yang–Baxter equation.

The $SL_q(2)$ quantum group symmetry of the Sine–Gordon equation is obtained by noticing that the soliton S-matrix of this model, given in eqn (18.8.7), can be expressed in terms of $\hat{R}_{12}(p, q)$ as

$$S^{SG}(\theta) \,=\, \frac{S(\theta)}{2 \sinh \frac{\pi(i\pi - \theta)}{\xi}}\, \hat{R}_{12}(\lambda = e^{\theta}, q), \qquad q \,=\, -e^{-i\pi^2/\xi} \tag{18.9.9}$$

where $S(\theta)$ is given in eqn (18.8.12). Notice that the deformation parameter q is related to the coupling constant. From $\left[S^{SG}, \Delta_q\right] = 0$, the soliton and antisoliton pair forms the fundamental spin-1/2 representation, whereas the multisoliton states may be regarded as the irreducible representations with higher spins which are created by tensor products like in eqn (18.9.4).

18.9.2 Restricted Sine–Gordon model

For arbitrary values of the coupling constant of the Sine–Gordon model, starting from the spin-1/2 representation of the soliton–antisoliton, we obtain multisoliton states with $j = 1, 3/2, \ldots$ Ordinarily there is no limit to the number of solitons. But, if ξ assume rational values, q given in eqn (18.9.9) becomes a root of unit and j is bounded by j_{max}. This means that the Sine–Gordon model at special rational values of the coupling constant cannot sustain solitons exceeding a certain number. Let's consider the various cases of interest.

Φ_{13} deformation of minimal unitary models $\mathcal{M}_m$. The scattering theory of these models is obtained when

$$\frac{\beta^2}{8\pi} \,=\, \frac{m}{m+1} \qquad \longrightarrow \qquad \xi \,=\, m\pi. \tag{18.9.10}$$

In this case $j_{max} = \frac{m}{2} - 1$ and there are at most $(m-2)$ solitons (or antisolitons). This peculiar aspect of the model can be understood as follows. In the original Hilbert space of the Sinh–Godon theory, there are many sectors, each containing a certain number of solitons. Then there is the sector containing up to $(m-2)$ solitons. That sector decouples from the rest of the Hilbert space if eqn (18.9.10) holds and it can be isolated out. Since solitons connect neighboring vacua, for a system having only up to a certain number of solitons its effective potential is going to have a cut: while there is an infinite degeneracy of vacua in the original Sinh–Godon theory, at the special values (18.9.10) there is a truncation. With $(m-2)$ solitons, the truncated potential has only $(m-1)$ vacua, as shown in Fig. 18.7, and one can imagine that the effective potential turns over at the edges. This agrees with the description of the massive $\Phi_{1,3}$ perturbed unitary minimal models $\mathcal{M}_m$: after perturbation, the original multiple vacua split into $(m-1)$ degenerate ones, all having the same energy.

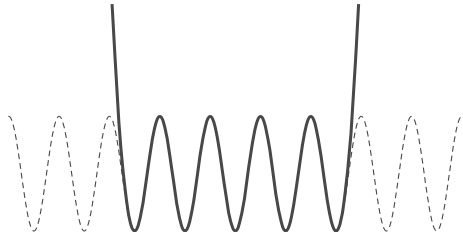


Fig. 18.7 *Effective potential of the Sinh–Godon theory at $\xi = \pi/m$ (for $m = 6$). The dashed line is the original untruncated potential.*

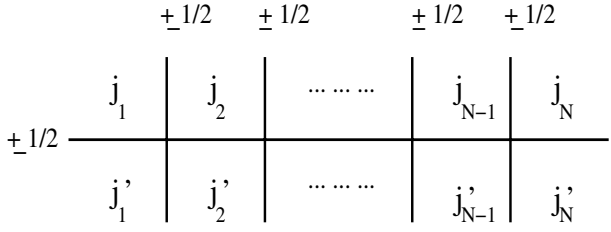


Fig. 18.8 *Change of basis from the vertex to IRF form.*

To get the S-matrix of the truncated theory it is necessary to change the basis in the Hilbert space: since a soliton–antisoliton pair forms a spin-1/2 representation $|\ \frac{1}{2}, \pm\frac{1}{2}\rangle$, we can decompose the multisoliton Hilbert space $\mathcal{H}$ into the irreducible spaces characterized by the higher spin, as shown in Fig. 18.8. In lattice models this is equivalent to changing the Boltzmann weights from the fluctuating variables expressed in terms of the vertices to the so-called RSOS (Restricted Solid On Solid) variables

$$\mathcal{H} = \sum_{m_i = \pm 1/2} |\ \frac{1}{2}, m_1\rangle \otimes |\ \frac{1}{2}, m_2\rangle \otimes \cdots |\ \frac{1}{2}, m_N\rangle = \sum_{\substack{0 \le j_1 < \infty \\ |j_i - j_{i+1}| = 1/2}} |\ j_1, \cdots, j_N\rangle$$

with appropriate q-CG coefficients. In this new basis, the multisoliton Hilbert space is spanned by the kink $K_{ab}(\theta)$, where a, b act as the RSOS vacua satisfying the condition $|a - b| = 1/2$. The kink is therefore a domain wall between two vacua, and a multikink state $|\ K_{ab}(\theta_1)K_{bc}(\theta_2)K_{cd}(\theta_3)\ldots\rangle$ should also have the next neighboring indices equal in order to have no jumps in the field configuration, as shown in Fig. 18.9. The S-matrix of the two-kink scattering can be derived from a restriction of the original S-matrix. For the scattering process

$$|\ K_{da}(\theta_1)K_{ab}(\theta_2)\rangle \to |\ K_{dc}(\theta_2)K_{cb}(\theta_1)\rangle$$

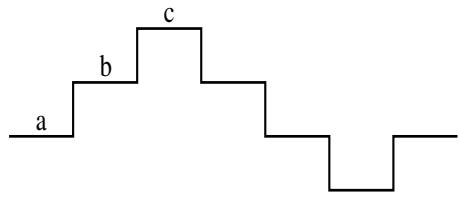
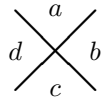


Fig. 18.9 *Multikink configuration.*

whose graphical representation is

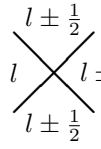


$$= S^{ab}_{dc}(\theta)$$

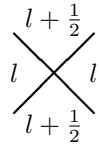
the explicit form of the RSOS S-matrix is

$$S^{ab}_{dc}(\theta) = \frac{S(\theta)}{2\sinh\frac{\pi(i\pi-\theta)}{\xi}} \left(\frac{[2a+1]]_q\,[2c+1]_q}{[2d+1]_q\,[2b+1]_q}\right)^{-\theta/2\pi i} \times \left[\delta_{db}\,\sinh\frac{\theta}{\xi}\left(\frac{[2a+1]]_q\,[2c+1]_q}{[2d+1]_q\,[2b+1]_q}\right)^{1/2} + \delta_{ac}\,\sinh\left(\frac{i\pi-\theta}{\xi}\right)\right]. \tag{18.9.11}$$

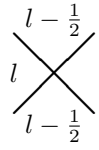
Note that for the values of ξ given in eqn (18.9.10) the original Sinh–Godon S-matrix and also the S-matrix S^{ab}_{dc} of the kinks do not have bound states. It satisfies the crossing condition $S^{ab}_{dc}(\theta) = S^{bc}_{ad}(i\pi-\theta)$. Its non-vanishing basic entries are (up to a common prefactor)



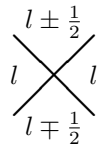
$$= \sinh\left(\frac{\pi(i\pi-\theta)}{\xi}\right)$$



$$= \frac{\sin\pi^2/\xi}{\sin[(2l+1)\pi^2/\xi]}\,\sinh\left(\frac{\pi[i\pi(2l+1)+\theta]}{\xi}\right) \tag{18.9.12}$$



$$= \frac{\sin\pi^2/\xi}{\sin[(2l+1)\pi^2/\xi]}\,\sinh\left(\frac{\pi[i\pi(2l+1)-\theta]}{\xi}\right)$$



$$= \frac{\sqrt{\sin[2l\pi^2/\xi]\,\sin[\pi^2(2l+2)/\xi]}}{\sin[(2l+1)\pi^2/\xi]}\,\sinh\left(\frac{\pi\theta}{\xi}\right).$$

In the formulas above, l labels the different vacuum states and takes the value

$$l = 0, \frac{1}{2}, \ldots, \frac{(m-2)}{2}. \tag{18.9.13}$$

Φ_{13} deformation of minimal non-unitary models $\mathcal{M}_{p_1p_2}$. The scattering theory of these models is obtained for these rational values of the coupling constant:

$$\frac{\beta^2}{8\pi} = \frac{p_1}{p_2} \quad \longrightarrow \quad \xi = \pi\,\frac{p_1}{p_2 - p_1} \tag{18.9.14}$$

(with $p_2 > p_1$). Notice that for these values of the coupling constant, the S-matrix of the Sinh–Godon model has poles corresponding to the bound states. Since the breathers are singlets of $SU_q(2)$, the restriction does not change the breather sector. The S-matrices of the breathers of the restricted Sinh–Godon theory are given exactly by eqns (18.8.25). Furthemore, notice that for these values $q^{p_1} = -1$, and the maximum value $j_{max} = (p_1 - 2)/2$ is determined only by p_1. The labels of the vacuum states are

$$l = 0, \frac{1}{2}, \ldots, \frac{(p_1-2)}{2}. \tag{18.9.15}$$

However, in this case there is an additional constraint coming from the unitarity condition of the RSOS S-matrix: the RSOS S-matrix (18.9.11) satisfies the equation

$$\sum_k S^{ab}_{dk}(\theta)\, S^{kb}_{dc}(-\theta) = \delta_{ac} \tag{18.9.16}$$

and, as long as the condition $S^\dagger(\theta) = S(-\theta)$ is satisfied, the scattering theory is unitary. The problem is with the last term in eqn (18.9.12) that contains square roots. The reality of this term selects the series of values

$$\xi = \pi\,\frac{r}{rk+1}, \qquad \begin{matrix} r = 2, 3, \ldots \\ k = 0, 1, \ldots \end{matrix} \tag{18.9.17}$$

and

$$\xi = \pi\,\frac{3}{3k+2}, \qquad k = 0, 1, \ldots \tag{18.9.18}$$

These are the only values of ξ for which the RSOS S-matrix description of the perturbed conformal models $\mathcal{M}_{p_1p_2}$ admit a self-consistent physical interpretation. For other rational values of ξ one can still use the RSOS S-matrix as monodromy algebra of the asymptotic particles but should be ready to sacrifice some of the usual properties of the S-matrix. Notice that in the series $\xi = \frac{2\pi}{2n+1}$ the solitons disapper completely

and only breathers remain in the spectrum. Their S-matrix was determined in Section 18.2. For the series $\xi = \frac{3\pi}{3n+1}$, there are instead solitons but they behave as ordinary particles because there are only two vacua.[5]

18.9.3 Quantum Group Symmetry of the Bullogh–Dodd Model

As discussed in Section 16.4, the $\Phi_{1,2}$ and $\Phi_{2,1}$ integrable deformations of the minimal models of conformal field theories

$$\mathcal{S}_{\pm}^{(12)} = \mathcal{S}_{CFT} \pm g \int \Phi_{12}(x)\, d^2x \tag{18.9.19}$$

$$\mathcal{S}_{\pm}^{(21)} = \mathcal{S}_{CFT} \pm g \int \Phi_{21}(x)\, d^2x \tag{18.9.20}$$

can be associated to the Bullogh–Dodd model with a charge at infinity and an imaginary coupling. Its lagrangian can be formally written as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2} e^{i\beta\phi} + e^{-i\beta\phi/2}. \tag{18.9.21}$$

An important difference with respect to the Sinh–Godon model (regarded as the Sinh–Gordon model with imaginary coupling) is that the lagrangian (18.9.21) is not a hermitian operator, so the definition itself of a lagrangian as (18.9.21) seems to be problematic. The solution of this problem and the resulting S-matrices for the massive $\Phi_{1,2}$ and $\Phi_{2,1}$ integrable deformations is one of the most beautiful results of the quantum group approach. It is due to F. Smirnov, who has shown that in this case only the restricted RSOS theories have a physical meaning. Here we review the main steps of this analysis, considering first the Φ_{12} deformation.

$\Phi_{1,2}$ deformation. Since the Bullogh–Dodd model is related to the (non-simply laced) Lie algebra $A_2^{(2)}$, the first step is to consider the R-matrix of this algebra. Similarly to the case of the Sinh–Godon model, it contains a spectral parameter λ but is constructed using the spin-1 representation of $SL_q(2)$. Its expression is

$$\hat{R}_{12}(\lambda, q) = (\lambda^{-1} - 1) q^{3/2} R_{12}(q) + (1 - \lambda) q^{-3/2} R_{21}^{-1}(q) + q^{-5/2}(q^2 - 1)(q^3 + 1)\mathcal{P},$$

where $\mathcal{P}$ is the permutation operator. $\hat{R}_{12}(\lambda, q)$ is an operator acting on the vector space $\mathbf{C}^3 \otimes \mathbf{C}^3$. The matrix $R_{12}(q)$ is the constant solution of the Yang–Baxter equation for the spin-1 representation of the quantum group $SL_q(2)$, given by

$$R_{12}(q) = \exp\left(\frac{H \otimes H}{4}\right) \left[I + (q^2 - 1)\, E \otimes F + (q-1)^2 (q+1)\, E^2 \otimes F^2 \right]$$

where

$$H = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \qquad E = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & q^{-1/2} \\ 0 & 0 & 0 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & q^{1/2} & 0 \end{pmatrix}.$$

[5] When there are only two vacua, the kink degrees of freedom are frozen because in the scattering $| K_{da}(\theta_1) K_{ab}(\theta_2) \rangle \to | K_{dc}(\theta_2) K_{cb}(\theta_1) \rangle$ the intermediate indices are forced to be equal, $b = c$.

The second intermediate step is to identify the hypothetical S-matrix of the three-component kink of the Bullogh–Dodd with imaginary coupling. To interpret the matrix $R_{12}(\lambda, q)$ as an S-matrix, one needs to relate the spectral parameter λ to the rapidity variable θ and q to the coupling constant β of the model. With the identification

$$q = e^{i\frac{16\pi^2}{\beta^2}}, \quad \lambda = e^{\frac{2\pi}{\xi}\theta}, \quad \xi = \frac{2}{3}\left(\frac{\pi\beta^2}{16\pi - \beta^2}\right) \tag{18.9.22}$$

the hypothetical S-matrix of the three-component kink is

$$\hat{S}_{12}(\theta) = S_0(\theta)\,\hat{R}_{12}\left(e^{\frac{2\pi}{\xi}\theta}, e^{i\frac{16\pi}{\beta^2}}\right). \tag{18.9.23}$$

The prefactor $S_0(\theta)$ ensures the validity of the "unitarity" equation $\hat{S}_{12}(\theta)\hat{S}_{21}(-\theta)$ and reads

$$\begin{aligned} S_0(\theta) = &\left[\sinh\frac{\pi}{\xi}(\theta - i\pi)\,\sinh\frac{\pi}{\xi}\left(\theta - \frac{2\pi i}{3}\right)\right]^{-1} \\ &\times \exp\left[-2i\int_0^\infty \frac{dt}{t}\,\frac{\sinh\left(\frac{\pi t}{3}\right)\cosh\left[\left(\frac{\pi}{6} - \frac{\xi}{2}\right)t\right]}{\cosh\left(\frac{\pi t}{2}\right)\sinh\left(\frac{\xi t}{2}\right)}\,\sin(\theta t)\right]. \end{aligned} \tag{18.9.24}$$

This prefactor satisfies the crossing relation $S_0(\theta) = S_0(i\pi - \theta)$. It can also be expressed in terms of an infinite product of Γ-functions, using the same procedure as the Sinh–Godon model. For generic value of ξ, the simple poles that lie on the physical sheet $0 \le \theta \le i\pi$ are at the crossing symmetric places

$$\theta = \begin{cases} i\pi - i\xi m, & i\xi m, & m > 0 \\ \frac{2\pi i}{3} - i\xi m, & \frac{i\pi}{3} + i\xi m, & m \le 0, \end{cases} \tag{18.9.25}$$

In both sets, the first poles are the singularities of the s-channel whereas the second poles are those of the crossing t-channel.

For the first set in (18.9.25), the R-matrix degenerates into a one-dimensional projector and the corresponding poles correspond then to the breather bound states. Using the bootstrap equation, the S-matrix of the fundamental breather (corresponding to the pole at $\theta = i\pi - i\xi$) is given by

$$S_{b_1 b_1}(\theta) = f_{\frac{2}{3}}(\theta)\, f_{\frac{\xi}{\pi}}(\theta)\, f_{\frac{\xi}{\pi} - \frac{1}{3}}(\theta), \tag{18.9.26}$$

where the functions $f_x(\theta)$ are those defined in (17.4.8).

For the second set of poles (18.9.25) the R-matrix degenerates instead into a three-dimensional projector and these poles are associated to the higher kinks. But, from a physical point of view, $\hat{S}_{12}(\theta)$ has some drawbacks that prevent it being interpreted it as the correct scattering amplitude of the perturbed conformal field theories. For instance, when q is a root of unity, the $\hat{R}_{12}$-matrix does not satisfy the relation $\hat{R}^*_{12}(\lambda) = \hat{R}_{21}(\lambda^{-1})$, which is crucial to correctly implement the unitarity condition of the scattering amplitudes. Therefore, as it is, the S-matrix (18.9.23) cannot be

assumed as the scattering amplitude of the Bullogh–Dodd model with imaginary coupling. It is only its RSOS restriction that has a physical interpretation and this happens when $q^r = 1$. The RSOS kink states that enter the reduced model

$$|\,\{\theta_1, j_1, a_1\}, \{\theta_2, j_2, a_2\}, \ldots \{\theta_n, j_n, a_n\}\rangle$$

are characterized by their rapidities θ_i, their $SL_q(2)$ spin, and by a string of numbers a_i (that identify the vacua) constrained by

$$a_i \leq \frac{1}{2}(r-2), \qquad |\, a_k - 1\,| \leq\, a_{k+1}\, \leq \min(a_k + 1, r - 3 - a_k). \tag{18.9.27}$$

These constraints formally correspond to the decomposition of tensor products of irreducible representations of

$SL_q(2)$ for $q^r = 1$:

$$V_{j_1} \otimes V_{j_2} \;=\; \sum_{j=|j_1-j_2|}^{\min(j_1+j_2,\, r-j_1-j_2-2)} V_j, \qquad j_1, j_2 \leq \frac{r-2}{2}.$$

The S-matrix of the RSOS kinks is

$$S^{a_k a'_k}_{a_{k-1} a_{k+1}}(\theta) = \frac{i}{4} S_0(\theta) \left[\left\{ \begin{matrix} 1 & a_{k-1} & a_k \\ 1 & a_{k+1} & a'_k \end{matrix} \right\}_q \left(e^{-\frac{2\pi}{\xi}\theta} - 1 \right) q^{(c_{a_{k+1}} + c_{a_{k-1}} - c_{a_k} - c_{a'_k} + 3)/2} \right. \tag{18.9.28}$$

$$\left. - \left(e^{\frac{2\pi}{\xi}\theta} - 1 \right) q^{-(c_{a_{k+1}} + c_{a_{k-1}} - c_{a_k} - c_{a'_k} + 3)/2} + q^{-5/2}(q^3+1)(q^2-1)\, \delta_{a_k a'_k} \right].$$

Here, $c_a \equiv a(a+1)$ and $\{\ldots\}_q$ are the quantum 6-j symbols. As for the RSOS restriction of the Sinh–Godon model, the above S-matrix is unitary if and only if the 6-j symbols are real. This happens for the following cases:

(i)

$$\frac{\beta^2}{8\pi} \;=\; \frac{m}{m+1}, \tag{18.9.29}$$

which corresponds to the $\Phi_{1,2}$ deformation of the minimal unitary models $\mathcal{M}_m$.

(ii)

$$\frac{\beta^2}{8\pi} \;=\; \frac{2}{2n+1}, \qquad \frac{\beta^2}{8\pi} \;=\; \frac{3\pi}{3n \pm 1}, \tag{18.9.30}$$

related to the Φ_{12} deformation of the minimal models $\mathcal{M}_{2,2n+1}$ and $\mathcal{M}_{3,3n\pm1}$. For these values of $\beta^2/8\pi$ the maximal allowed spins are 0 and $\frac{1}{2}$. Hence the kinks disappear and the spectrum is only given by the breathers obtained by closing the bootstrap with the initial S-matrix (18.9.26).

(iii)

$$\frac{\beta^2}{8\pi} \;=\; \frac{4\pi}{4n \pm 1}, \tag{18.9.31}$$

which correspond to the $\Phi_{1,2}$ deformation of the minimal models $\mathcal{M}_{4,4n\pm1}$. For this series the maximal allowed spin is equal to 1 and, according to the RSOS restriction, the kinks behave as a scalar particle.

For other rational values of $\beta^2/8\pi = r/s$ one can still use the RSOS S-matrix as monodromy algebra of the asymptotic particles but should be ready to sacrifice some of the usual properties of the S-matrix. Properly interpreted, they can be assumed as the S-matrix of the Φ_{12} perturbation of the minimal models $\mathcal{M}_{r,s}$.

Let's discuss in more detail the vacuum structure for the values (18.9.29), corresponding to the Φ_{12} deformation of the unitary minimal models. Since the R-matrix is based on a spin-1 representation, there are two closed subspaces V_m^+ and V_m^- containing, respectively, half-integer or integer spins out of the set $a = 0, 1/2, 1, \ldots, m/2 - 1$. Each of these subspaces is associated to the set of vacua and the RSOS reduction gives rise to two quantum field theories. In the Landau–Ginzburg picture, the field $\Phi_{1,2}$ is associated to the composite operator $:\varphi^{m-2}:$. Hence, when m is odd, the field Φ_{12} is odd under the Z_2 spin symmetry and therefore, changing the sign of g in (18.9.19) leads to the same theory. On the contrary, when m is even the two theories $\mathcal{S}_\pm^{(12)}$ are expected to be different.[6] It becomes natural to identify V^+ with the vacuum states of the theory $\mathcal{S}_+^{(12)}$ and V^- with those of $\mathcal{S}_-^{(12)}$. So we have the following situations

- m **odd**. There are $(m-1)/2$ degenerate vacua in both theories $\mathcal{S}_\pm^{(12)}$ labeled as

$$\begin{aligned} a &= \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-2}{2}, \quad g > 0 \\ a &= 0, 1, \ldots, \frac{m-3}{2}, \quad g < 0 \end{aligned}$$

- m **even.** The number of vacua of $\mathcal{S}_+^{(12)}$ is equal to $(m-2)/2$, while the number of vacua of $\mathcal{S}_-^{(12)}$ is equal to $m/2$. Their label is

$$\begin{aligned} \mathbf{a} &= \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-3}{2}, \quad g > 0 \\ \mathbf{a} &= 0, 1, \ldots, \frac{m-2}{2}, \quad g < 0. \end{aligned}$$

Finally, some checks of the above formalism. For $m = 3$, there are only breathers and the S-matrix of the first breather, eqn (18.9.26), correctly coincides with the amplitude (18.4.15) of the Ising model perturbed by a magnetic field. The bootstrap closes with eight breathers. For $m = 4$, the two theories $\mathcal{S}_\pm^{(12)}$ are related by duality: for $g < 0$ there are only two vacua and the kinks behaves as a particle. The RSOS S-matrix of the lowest kink and of the lowest breather correctly coincides with the amplitudes (18.5.5) of the tricritical Ising model away from its critical temperature and the bootstrap closes with seven particles.

Φ_{21} **deformation**. The discussion of the RSOS S-matrix of the Φ_{21} deformation is similar to the one above, the only difference being in the definition of the q-parameter and the spectral parameter λ. In this case, the corresponding values are

$$q = e^{i\frac{\pi^4\beta^2}{4}}, \qquad \lambda = e^{\frac{2\pi}{\tilde{\xi}}\theta}, \qquad \tilde{\xi} = \frac{8}{3}\left(\frac{\pi^2}{\beta^2 - 4\pi}\right) \tag{18.9.32}$$

[6] It can be proved, however, that the two theories are related by duality.

and in all previous formulas ξ has to be changed to $\tilde{\xi}$. In this case we shall also ensure that $\beta^2/4\pi > 1$, in such a way that the field Φ_{21}, with conformal weight $\Delta_{21} = \frac{1}{2} + \frac{6\pi}{\beta^2}$, is a relevant operator. In the Landau–Ginzburg picture, the field Φ_{21} is associated to the composite operator $:\varphi^{m-1}:$. Therefore, under the Z_2 spin symmetry, Φ_{21} is odd if m is even and even if m is odd. In this case the structure of the vacua of the theories $\mathcal{S}_{\pm}^{(21)}$ is as follows:

- when m is even, for both $\mathcal{S}_{\pm}^{(21)}$ the number of the vacua is $m/2$ and

$$\begin{aligned} a &= \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-1}{2}, \quad g > 0 \\ a &= 0, 1, \ldots, \frac{m-2}{2}, \quad g < 0; \end{aligned}$$

- when m is odd, the number of the vacua is $(m-1)/2$ for $g > 0$, and $(m+1)/2$ for $g < 0$, with

$$\begin{aligned} a &= \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-2}{2}, \quad g > 0 \\ a &= 0, 1, \ldots, \frac{m-1}{2}, \quad g < 0. \end{aligned}$$

A significant example. Sub-leading magnetization of the tricritical Ising model. An interesting example of the formalism above is provided by the Tricritical Ising model, i.e. the unitary minimal model $\mathcal{M}_4$, perturbed by the sub-leading magnetization operator Φ_{21}. This is a field odd under the Z_2 spin-reversal transformation: since this deformation explicitly breaks the Z_2 symmetry of the tricritical point, the corresponding massive theory can exhibit the "Φ^3-property". The counting argument supports this picture, giving for the spin of the conserved currents the values s = (1, 5, 7, 11, 13). The RSOS picture predicts for such a theory two vacuum states (hereafter denoted by $|\, 0\rangle$ and $|\, 1\rangle$), which can be associated to the minima of the asymmetric double-well Landau–Ginzburg potential in Fig. 18.10. The twofold degeneracy of the vacua permits two fundamental kink configurations $|\, K_+\rangle$ and $|\, K_-\rangle$ and, possibly, bound states $|\, B\rangle$ thereof. If the two vacua were related by a symmetry transformation, i.e. if we were in the situation of a Z_2 spontaneously broken symmetry, there would be a double degeneracy of the breather-like bound state $|\, B\rangle$. But the absence of a Z_2 symmetry makes it possible that in this case only one of the two asymptotic states $|\, K_+K_-\rangle$ or $|\, K_-K_+\rangle$ couples to the bound state $|\, B\rangle$. This is confirmed by the explicit solution of the model, given by the RSOS S-matrix. In this case, the only possible values of a_i which label the vacuum states in the RSOS S-matrix are 0 and 1. The one-particle states are thus the vectors $|\, K_{01}\rangle$, $|\, K_{10}\rangle$, and $|\, K_{11}\rangle$. They correspond to the states that we previously denoted as $|\, K_+\rangle$, $|\, K_-\rangle$, and $|\, B\rangle$, respectively. All of

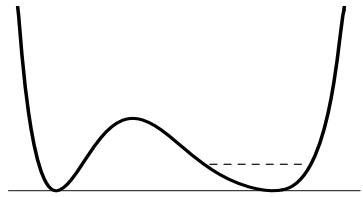


Fig. 18.10 *Landau–Ginzburg potential for the sub-leading magnetic deformation of the tri-critical Ising model.*

them have the same mass m. Notice that the state $\mid K_{00}\rangle$ is projected out because of the reduction. The scattering processes are given by

$$
\begin{aligned}
&\mid K_{01}(\theta_1)K_{10}(\theta_2)\rangle = S_{00}^{11}(\theta_1-\theta_2) \mid K_{01}(\theta_2)K_{10}(\theta_1)\rangle \\
&\mid K_{01}(\theta_1)K_{11}(\theta_2)\rangle = S_{01}^{11}(\theta_1-\theta_2) \mid K_{01}(\theta_2)K_{11}(\theta_1)\rangle \\
&\mid K_{11}(\theta_1)K_{10}(\theta_2)\rangle = S_{10}^{11}(\theta_1-\theta_2) \mid K_{11}(\theta_2)K_{10}(\theta_1)\rangle \\
&\mid K_{11}(\theta_1)K_{11}(\theta_2)\rangle = S_{11}^{11}(\theta_1-\theta_2) \mid K_{11}(\theta_2)K_{11}(\theta_1)\rangle + S_{11}^{10}(\theta_1-\theta_2) \mid K_{10}(\theta_2)K_{01}(\theta_1)\rangle \\
&\mid K_{10}(\theta_1)K_{01}(\theta_2)\rangle = S_{11}^{00}(\theta_1-\theta_2) \mid K_{10}(\theta_2)K_{01}(\theta_1)\rangle + S_{11}^{10}(\theta_1-\theta_2) \mid K_{11}(\theta_2)K_{11}(\theta_1)\rangle .
\end{aligned}
\tag{18.9.33}
$$

Explicitly, the above amplitudes are given by

$$
\begin{array}{l}
\begin{array}{ccc} & 1 & \\ 0 & \times & 0 \\ & 1 & \end{array} \;=\; S_{00}^{11}(\theta) \;=\; \dfrac{i}{2}\, S_0(\theta)\, \sinh\left(\dfrac{9}{5}\theta - i\dfrac{\pi}{5}\right) \\[2ex]
\begin{array}{ccc} & 1 & \\ 0 & \times & 1 \\ & 1 & \end{array} \;=\; S_{01}^{11}(\theta) \;=\; -\dfrac{i}{2}\, S_0(\theta)\, \sinh\left(\dfrac{9}{5}\theta + i\dfrac{\pi}{5}\right) \\[2ex]
\begin{array}{ccc} & 1 & \\ 1 & \times & 1 \\ & 1 & \end{array} \;=\; S_{11}^{11}(\theta) \;=\; \dfrac{i}{2}\, S_0(\theta)\, \dfrac{\sin\left(\frac{\pi}{5}\right)}{\sin\left(\frac{2\pi}{5}\right)}\, \sinh\left(\dfrac{9}{5}\theta - i\dfrac{2\pi}{5}\right) \\[2ex]
\begin{array}{ccc} & 1 & \\ 1 & \times & 1 \\ & 0 & \end{array} \;=\; S_{11}^{01}(\theta) \;=\; -\dfrac{i}{2}\, S_0(\theta)\, \left(\dfrac{\sin\left(\frac{\pi}{5}\right)}{\sin\left(\frac{2\pi}{5}\right)}\right)^{\frac{1}{2}}\, \sinh\left(\dfrac{9}{5}\theta\right) \\[2ex]
\begin{array}{ccc} & 0 & \\ 1 & \times & 1 \\ & 0 & \end{array} \;=\; S_{11}^{00}(\theta) \;=\; -\dfrac{i}{2}\, S_0(\theta)\, \dfrac{\sin\left(\frac{\pi}{5}\right)}{\sin\left(\frac{2\pi}{5}\right)}\, \sinh\left(\dfrac{9}{5}\theta + i\dfrac{2\pi}{5}\right).
\end{array}
$$

The function $S_0(\theta)$ which implements the unitarity condition reads

$$S_0(\theta) = -\left(\sinh\frac{9}{10}(\theta - i\pi)\sinh\frac{9}{10}\left(\theta - \frac{2\pi i}{3}\right)\right)^{-1} \times w_{-\frac{1}{5}}(\theta))\, w_{\frac{1}{10}}(\theta)\, w_{\frac{3}{10}}(\theta)\, s_{\frac{2}{9}}(\theta)\, s_{-\frac{8}{9}}(\theta)\, s_{\frac{7}{9}}(\theta)\, s_{-\frac{1}{9}}(\theta)\,,$$

where

$$w_x(\theta) = \frac{\sinh\left(\frac{9}{10}\theta + i\pi x\right)}{\sinh\left(\frac{9}{10}\theta - i\pi x\right)}, \qquad s_x(\theta) = \frac{\sinh\ \frac{1}{2}(\theta + i\pi x)}{\sinh\ \frac{1}{2}(\theta - i\pi x)}.$$

The amplitudes are periodic along the imaginary axis of θ with period $10\,\pi i$. The whole structure of poles and zeros is quite rich. On the physical sheet, $0 \le \mathrm{Im}\,\theta \le i\pi$, the poles of the S-matrix are located at $\theta = 2\pi i/3$ and $\theta = i\pi/3$. The first pole corresponds to a bound state in the direct channel whereas the second one is the singularity due to the particle exchanged in the crossed process. The residues at $\theta = 2\pi i/3$ are given by

$$\begin{aligned}
r_1 &= \mathrm{Res}_{\theta=\frac{2\pi i}{3}}\ \ S_{00}^{11}(\theta) = 0;\\
r_2 &= \mathrm{Res}_{\theta=\frac{2\pi i}{3}}\ \ S_{01}^{11}(\theta) = i\,\left(\frac{s\left(\frac{2}{5}\right)}{s\left(\frac{1}{5}\right)}\right)^2\,\omega;\\
r_3 &= \mathrm{Res}_{\theta=\frac{2\pi i}{3}}\ \ S_{11}^{11}(\theta) = i\,\omega;\\
r_4 &= \mathrm{Res}_{\theta=\frac{2\pi i}{3}}\ \ S_{11}^{01}(\theta) = i\,\left(\frac{s\left(\frac{2}{5}\right)}{s\left(\frac{1}{5}\right)}\right)^{\frac{1}{2}}\,\omega;\\
r_5 &= \mathrm{Res}_{\theta=\frac{2\pi i}{3}}\ \ S_{11}^{00}(\theta) = i\,\frac{s\left(\frac{2}{5}\right)}{s\left(\frac{1}{5}\right)}\,\omega;
\end{aligned} \tag{18.9.34}$$

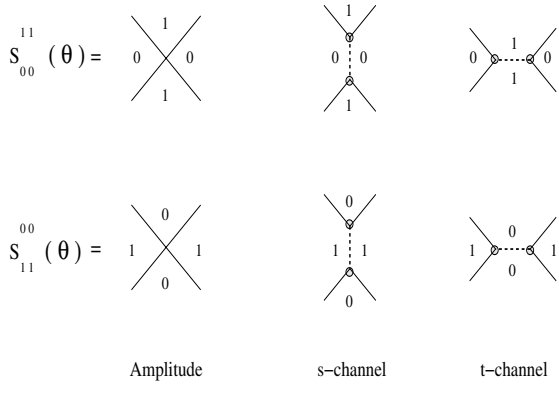


Fig. 18.11 *Elastic scattering amplitudes of the kinks in an asymmetric well potential and their intermediate states in the s-channel and in the t-channel.*

where $s(x) \equiv \sin(\pi x)$ and

$$\omega = \frac{5}{9} \, \frac{s\left(\frac{1}{5}\right) s\left(\frac{1}{10}\right) s\left(\frac{4}{9}\right) s\left(\frac{1}{9}\right) s^2\left(\frac{5}{18}\right)}{s\left(\frac{3}{10}\right) s\left(\frac{1}{18}\right) s\left(\frac{7}{18}\right) s^2\left(\frac{2}{9}\right)}.$$

Hence, in the s-channel of the amplitude S_{00}^{11}, there is no bound state related to $\mid K_{00} \rangle$ (a state that does not exist): its only singularity comes from the bound state $\mid K_{11} \rangle$, exchanged, however, in the t-channel. In the amplitude S_{11}^{00} the situation is reversed (the two amplitudes are related by crossing): there is the s-channel singularity due to the bound state $\mid K_{11} \rangle$ while that of the t-channel is absent. This is easily seen from Fig. 18.11, where the original amplitudes are streched along the vertical direction (s-channel) and along the horizontal one (t-channel).

References and Further Reading

The determination of scattering theory of the Ising model in a magnetic field is due to A. B. Zamolodchikov:

A. B. Zamolodchikov, *Integrable field theory from conformal field theory*, Adv. Stud. Pure Math., 19 (1989), 641.

The S-matrix of the Yang–Lee model has been computed in:

J.L. Cardy, G. Mussardo, *S-matrix of the Yang–Lee edge singularity in two dimensions* Phys. Lett. B 225 (1989), 275.

The S-matrices of the Toda field theories and their relation with the minimal S-matrices of statistical models have been studied by several authors. We suggest consulting the papers:

A. Arishtein, V. Fateev, A.B. Zamolodchikov, *Quantum S-matrix of the* $(1 + 1)$ *dimensional Toda chain*, Phys. Lett. B 87 (1979), 389.

P. Christe, G. Mussardo, *Integrable systems away from criticality: The Toda field theory and S-matrix of the tricritical Ising model*, Nucl. Phys. B 330 (1990), 465.

P. Christe, G. Mussardo, *Elastic S-matrices in* $(1 + 1)$ *dimensions and Toda field theories*, Int. J. Mod. Phys. A 5 (1990), 4581.

H.W. Braden, E. Corrigan, P.E. Dorey, R. Sasaki, *Affine Toda field theory and exact S-matrices*, Nucl. Phys. B 338 (1990), 689.

H.W. Braden, E. Corrigan, P.E. Dorey, R. Sasaki, *Multiple poles and other features of affine Toda field theories*, Nucl. Phys. B 356 (1991), 469.

V.A. Fateev, A.B. Zamolodchikov, *Conformal field theory and purely elastic S-matrices*, Int. J. Mod. Phys. A 5 (1990), 1025.

G. Sotkov, C.J. Zhu, *Bootstrap fusions and tricritical Potts model away from criticality*, Phys. Lett. B 229 (1989), 391.

The scattering theory of geometrical statistical models, as self-avoiding walks or percolation, has been studied in the papers:

A.B. Zamolodchikov, *Exact S-matrix associated with self-avoiding polymer problem in two dimensions*, Mod. Phys. Lett. A 6 (1991), 1807.

L. Chim, A.B. Zamolodchikov, *Integrable field theory of the q-state Potts model with* $0 < q < 4$, Int. J. Mod. Phys. A 7 (1992), 5317.

The S-matrix of the Sinh–Godon model has a long history that can be traced back by reading the article in which the exact S-matrix of the model was determined:

A.B. Zamolodchikov, Al.B. Zamolodchikov, *Factorized S-matrices in two dimensions as the exact solutions of certain relativistic quantum field theories*, Ann. Phys. 120 (1979), 253.

An interesting article, from a simple minded point of view, on the properties of the Sinh–Godon S-matrix is:

C.J. Goebel, *On the Sinh–Godon S-matrix*, Prog. Theor. Phys. Suppl. 86 (1986), 261.

Scattering theory for the integrable $\Phi_{1,3}$, $\Phi_{1,2}$ and $\Phi_{2,1}$ deformations of the minimal models admits a general formulation. It requires, though, the sophisticated tools of quantum groups. The backbone of our understanding of this approach comes mainly from F.A. Smirnov and A. LeClair, and their writings are a fine place to learn about the subject:

F.A. Smirnov, *Reductions of the Sinh–Godon model as perturbation of minimal models of conformal field theory*, Nucl. Phys. B 337 (1990), 156.

F.A. Smirnov, *Exact S-matrices for* $\Phi_{1,2}$ *perturbed minimal models of conformal field theory*, Int. J. Mod. Phys. A 6 (1991), 1407.

A. LeClair, *Restricted Sinh–Godon theory and the minimal conformal series*, Phys. Lett. B 230 (1989), 103.

For an introduction to quantum groups see:

S. Majid, *Foundations of Quantum Group Theory*, Cambridge University Press, Cambridge, 1995.

Problems

1. Bootstrap equations

Prove that the most general solution of the bootstrap equation relative to a particle bound state of itself

$$S_{AA}(\theta) \,=\, S_{AA}\left(\theta - \frac{i\pi}{3}\right) S_{AA}\left(\theta + \frac{i\pi}{3}\right)$$

is given by

$$S_{AA}(\theta) \,=\, f_{\frac{2}{3}}(\theta) \prod_i f_{-x_i}(\theta)\, f_{\frac{2}{3}-x_i}(\theta).$$

Study the motion of the poles of the function $S_{AA}(\theta)$ under the shifts induced by the bootstrap equation.

2. Analytic structure of the S-matrix of the Bullogh–Dodd model

a Study the structure of the poles and zeros of the S-matrix of the Bullogh–Dodd model

$$S(\theta) \,=\, f_{\frac{2}{3}}(\theta)\, f_{-\frac{B}{3}}(\theta)\, f_{\frac{B-2}{3}}(\theta)$$

with

$$B(\lambda) \,=\, \frac{\lambda^2}{2\pi}\frac{1}{1+\frac{\lambda^2}{4\pi}}$$

by varying the coupling constant λ.

b Make the analytic continuation

$$B \to 1 + \frac{3}{i\pi}\beta_0$$

with β_0 real and show that in the limit $\beta_0 \to \infty$ the S-matrix of the Bullogh–Dodd model reduces to the S-matrix of the Yang–Lee model.

3. Multiple poles

Prove that the amplitude S_{11} of the fundamental particle cannot have higher order poles by noticing that the resonance angle of two heavier masses is larger than $2\pi/3$. This makes it impossible to draw a diagram such as the one in Fig. 18.2.

4. Double poles

Use the values of the resonance angles of the S-matrix of the thermal tricritical Ising model to explain the double poles that enter the amplitude $S_{1,6}$ in terms of multi-scattering processes.

5. S-matrix of the Gross–Neveu model

The Gross–Neveu model is a model of the n-component neutral Fermi field $\psi_k(x)$ $k = 1, 2, \ldots, n$ $(n \geq 3)$, with four-fermion interaction

$$\mathcal{L} = \frac{i}{2} \sum_{k=1}^{n} \bar{\psi}_k \gamma^\mu \partial_\mu \psi_k + \frac{g}{8} \left[\sum_{k=1}^{n} \bar{\psi}_k \psi_k \right]^2$$

where $\bar{\psi}_k = \psi_k \gamma^0$ and the 2×2 γ^μ matrices satisfy the anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. Like the bosonic $O(n)$ σ model, the Gross–Neveu model is massive, renormalizable, asymptotically free, and explicitly $O(n)$ symmetric. It is also integrable.

With the notation of Section 18.7, the exact S-matrix of the Gross–Neveu model can be obtained by solving the unitarity and crossing equations for $S_2(\theta)$

$$S_2(\theta) S_2(-\theta) = \frac{\theta^2}{\theta^2 + \lambda^2}, \qquad S_2(\theta) = S_2(i\pi - \theta)$$

with the initial seed $Q(\theta) = \frac{\theta}{\theta - i\lambda}$ where $\lambda = 2\pi/(n-2)$.

a With the notation of eqn (18.7.8), show that in this case one ends up with

$$\mathcal{U}^{(-)}(\theta) = \frac{\Gamma\left(-\frac{\lambda}{2\pi} - i\frac{\theta}{2\pi}\right) \Gamma\left(\frac{1}{2} - i\frac{\theta}{2\pi}\right)}{\Gamma\left(\frac{1}{2} - \frac{\lambda}{2\pi} - i\frac{\theta}{2\pi}\right) \Gamma\left(-i\frac{\theta}{2\pi}\right)}.$$

b Prove that the amplitudes $\mathcal{U}^{(\pm)}$ are related as

$$\mathcal{U}^{(-)}(\theta) = \frac{\sinh\theta + i \sin\lambda}{\sinh\theta - i \sin\lambda} \mathcal{U}^{(+)}(\theta).$$

c Consider the amplitudes with definite isospin channel

$$\begin{aligned} S_{\text{isoscalar}} &= N S_1 + S_2 + S_3 \\ S_{\text{antisym}} &= S_2 - S_3 \\ S_{\text{sym}} &= S_2 + S_3. \end{aligned}$$

Bound states exist only in isoscalar and antisymmetric isospin channels. Denote these new particles by B and B_{ij} and show that their masses are

$$m_B = m_{B_{ij}} = m \frac{\sin\left(\frac{2\pi}{n-2}\right)}{\sin\left(\frac{\pi}{n-2}\right)}.$$

where m is the mass of the elementary fermion.

6. Integral representation

Use the expansions

$$\frac{1}{\cosh x} \;=\; 2\sum_{k=0}^{\infty}(-1)^k\, e^{-(2k+1)x}\,, \qquad \frac{1}{\sinh x} \;=\; 2\sum_{k=0}^{\infty} e^{-(2k+1)x},$$

the infinite-product

$$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\gamma)\Gamma(\beta-\gamma)} \;=\; \prod_{k=0}^{\infty}\left[\left(1+\frac{\gamma}{\alpha+k}\right)\left(1-\frac{\gamma}{\beta+k}\right)\right],$$

and the integral

$$\int_0^{\infty}\frac{dt}{t}\,e^{-\beta t}\,\sin(\alpha t) \;=\; \frac{1}{2i}\log\left[\frac{1+i\alpha/\beta}{1-i\alpha/\beta}\right]$$

to prove the integral representation of eqns (18.7.19) and (18.8.12). Isolate a finite number of poles and also recover the mixed representation given in eqn (18.8.13).

7. Sine–Gordon

a Study the analytic structure of the S-matrix of the solitons of the Sine–Gordon model, identifying all the sequences of the poles in the amplitudes.

b Using the following definition of the breathers B_n

$$B_n\left(\frac{\theta_1+\theta_2}{2}\right) \;=\; \lim_{\theta_1-\theta_2\to in\xi}\left[A(\theta_2)\bar{A}(\theta_1)+(-1)^n\bar{A}(\theta_2)A(\theta_1)\right]$$

compute the S-matrix of these particles by means of the residue of the S-matrix of the solitons.

8. Reflectionless points

At $\xi=\pi/n$ $(n=1,2,\ldots)$, the amplitude S_R of the Sine–Gordon model vanishes and soliton–antisoliton scattering reduces to a pure transmission. Use the properties of the Γ function to prove that for these values of the coupling constant the transmission amplitude becomes

$$S_T(\theta) \;=\; e^{in\pi}\prod_{k=1}^{n-1}\frac{e^{\theta-i(\pi k/n)}+1}{e^{\theta}+e^{-i(\pi k/n)}}.$$

9. Bound states and semiclassical limit

It can be proved that the renormalized coupling constant

$$\xi \;=\; \frac{\beta^2}{8}\,\frac{1}{1-\frac{\beta^2}{8\pi}}$$

of the Sine-Gordon model comes from the quantum correction of the classical action. By the same token, it is possible to prove that the exact mass of the soliton–

antisoliton is

$$M = \frac{m}{\xi},$$

where m is the parameter in the lagrangian. Keeping m fixed, the semiclassical limit $\beta^2 \to 0$ of the Sine–Gordon model gives rise to a non-trivial theory.

a Use the expression above of the mass of the soliton to express differently the mass of the breathers B_n of the Sine–Gordon model, given in eqn (18.8.21).

b Expand m_n in powers of β^2 and show, that to lowest order, $m_n \simeq n m_1$. So, all these states can be considered as loosely bound states of n "elementary" bosons B_1.

c Compute to order β^4 the binding energy $\Delta E_n \equiv n m_1 - m_n$ of these states.

10. Sine–Gordon and non-unitary models

a Find the value of ξ for which the S-matrix element $S^{(1,1)}(\theta)$ of the Sine–Gordon model coincides with the S-matrix of the Yang–Lee model. Explain why the restriction of the Sine–Gordon model produces a negative residue at the pole $\theta = 2\pi i/3$.

b Generalize the result above, finding the value of ξ that leads to the equality of the S-matrices $S^{(n,m)}(\theta)$ of the Sine–Gordon model with those of the integrable deformation of the minimal non-unitary models $\mathcal{M}_{2,2n+1}$.

19
Thermodynamical Bethe Ansatz

In quantum mechanics there are principles that are certain and these are much more important for the world and for us than the uncertainty principle.

Hans A. Bethe

19.1 Introduction

The thermodynamics of a quantum field theory in an infinite volume can be determined by its S-matrix. This idea, originally proposed by R. Dashen, S.K. Ma, and H.J. Berstein, has been widely used to study the thermal properties of the integrable field theories in $(1+1)$ dimensions. The reason consists of the particularly simple properties of these scattering matrices, as discussed in the previous two chapters, and the possibility to generalize to the relativistic case the thermodynamics Bethe ansatz (TBA) techniques successfully applied to non-relativistic problems by C.N. Yang and C.P. Yang. In the TBA approach, the derivation of the thermodynamics of a purely elastic scattering theory reduces to finding the solution of a set of nonlinear integral equations that rule the energies of the particle excitations and their statistical distribution.

The TBA equations for the relativistic models with a diagonal S-matrix have been derived by A.B. Zamolodchikov. Several applications have been made by Zamolodchikov himself and many other authors. In addition to the generalization of the TBA to the non-diagonal S-matrices, further advances have been accomplished in the computation of the energies of the excited states, in the analysis of systems with generic integrable boundary conditions, and also in the discovery of interesting relations with the Schröedinger equation in quantum mechanics. In this chapter we discuss the main ideas of this approach; for all the advanced topics of the subject we refer the reader to the articles listed at the end of the chapter.

19.2 Casimir Energy

Consider a $(1+1)$-dimensional euclidean quantum field theory defined on a cylinder, with periodic boundary conditions in both the R and L directions. There are two equivalent ways to quantize the theory on such a geometry: from the symmetry of the two directions, one can equivalently choose as the time direction one of the two axes and consider the other as the space direction[1] Hence, the partition function can be

[1] In the context of conformal field theory this is the basis of modular invariance, see Section 11.7.

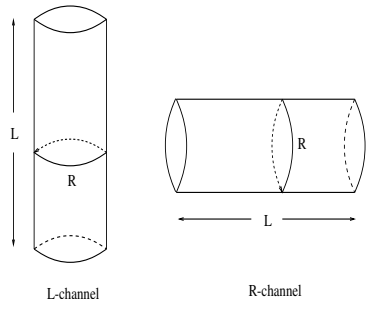


Fig. 19.1 *Cylinder geometry with periodic boundary conditions on both directions, with the two different channels of quantization.*

written either as

$$Z(R,L) \;=\; \mathrm{Tr}\; e^{-L\,\mathcal{H}_R}, \tag{19.2.1}$$

or as

$$Z(R,L) \;=\; \mathrm{Tr}\; e^{-R\,\mathcal{H}_L}, \tag{19.2.2}$$

where $\mathcal{H}_R$ and $\mathcal{H}_L$ are the hamiltonians of the system quantized along the R and L axes, and the trace is a sum done over their eigenstates (See Fig. 19.1). The two hamiltonians can be expressed in terms of the stress–energy tensor $T_{\mu\nu}$, where x and y denote the coordinates along the R and L axes, respectively. In fact we have

$$\mathcal{H}_R \;=\; \frac{1}{2\pi}\int T_{yy}\,dx,$$

while

$$\mathcal{H}_L \;=\; \frac{1}{2\pi}\int T_{xx}\,dy.$$

The quantization scheme in which the role of the time direction is played by the L axis will be denoted as the *L-channel*, while the other one is the *R-channel*.

When $L \to \infty$, the expression in (19.2.1) clearly reduces only to the lowest term, given by the ground state energy $E_0(R)$ of $\mathcal{H}_R$:

$$Z(R,L) \;\simeq\; e^{-L\,E_0(R)}. \tag{19.2.3}$$

But, taking the limit $L \to \infty$ in the second expression (19.2.2) is equivalent to the thermodynamic limit of a one-dimensional quantum system defined along the L axis at temperature $T \equiv 1/R$. In this case, the limiting form of the partition function can be written as

$$Z(R,L) \;\simeq\; e^{-LRf(R)}, \tag{19.2.4}$$

where $f(R)$ is the free energy per unit length of the system at temperature $1/R$. Comparing the two limiting expressions (19.2.3) and (19.2.4) of the partition function, we find

$$E_0(R) \;=\; R\,f(R). \tag{19.2.5}$$

This equation states the important relation between the Casimir energy $E_0(R)$ of the ground state on a finite volume and the free energy $f(R)$ of the one-dimensional quantum system at infinite volume but at temperature $T = 1/R$.

From the translation invariance of the cylinder geometry along the two axes, the one-point correlation functions are independent of the coordinates. In particular, we have

$$\langle T_{yy} \rangle \;=\; 2\pi \frac{E_0(R)}{R}, \quad \langle T_{xx} \rangle \;=\; 2\pi \frac{dE_0(R)}{dR},$$

and, for the one-point correlation function of the trace of the stress–energy tensor $\Theta = (T_{xx} + T_{yy})$, we have

$$\langle \Theta \rangle \;=\; \frac{2\pi}{R} \frac{d}{dR}[RE_0(R)]. \tag{19.2.6}$$

Furthermore, for theories that are invariant under parity and with a unique ground state, we have

$$\langle T_{xy} \rangle \;=\; \langle T_{yx} \rangle \;=\; 0.$$

It is convenient to parameterize the ground state energy as

$$E_0(R) \;=\; -\frac{\pi \tilde{c}(r)}{6R}, \tag{19.2.7}$$

where $r = m_1 R$ is a purely dimensionless variable, with m_1 the lowest mass gap of the theory. As we will show later, the scaling function $\tilde{c}(r)$ can be determined for any value of r by the TBA equations based on the scattering data. There is, however, a simple limit of this expression: in the ultraviolet limit $r \to 0$, the behavior of the ground state energy is controlled by the underlying conformal field theory

$$E_0(R) \;=\; \frac{2\pi}{R} \left(\Delta_{min} + \overline{\Delta}_{min} - \frac{c}{12} \right), \tag{19.2.8}$$

and, for a theory in which $\Delta_{min} = \overline{\Delta}_{min}$, the function $\tilde{c}(r)$ goes to the effective central charge

$$\lim_{r \to 0} \tilde{c}(r) \;=\; c - 24\Delta_{min}. \tag{19.2.9}$$

Notice that this limit establishes an important relation between the scattering theory of a massive quantum field theory and the conformal theory that rules its short-distance behavior. The confirmation and the validity of many scattering theories proposed to describe the deformations of conformal field theories can be accomplished thanks to the relation above.

In the sections to come, we will derive the TBA equations following the original proposals by Yang and Yang and by Zamolodchikov, discussing all their important consequences. In the last two sections, using the simple example of a free massive theory, we will show how to derive the ground state energy $E_0(R)$ at a finite volume by directly quantizing the theory in its L-channel.

19.3 Bethe Relativistic Wave Function

Consider a $(1+1)$-dimensional integrable theory defined on a circumference of length L. Let's assume that the spectrum consists of a set of particles A_a $(a = 1, 2, \ldots, n)$ with masses m_a, and that their scattering amplitudes are purely diagonal and characterized by their phase shifts $\delta_{ab}(\theta)$, where $S_{ab}(\theta) = e^{i\delta_{ab}(\theta)}$. The lowest mass determines the correlation length of the system through $\xi = 1/m_1$. The particles can be either bosons or fermions.

The Hilbert space of such a theory is rather simple. In fact, given any N-particle state, the integrability of the theory ensures that the time evolution of this state preserves both the identity of the particles and their momenta. In this case, it makes sense to associate to any state of such a relativistic system a wavefunction $\Psi(x_1, \ldots, x_N)$. In the configurational space of the N-particle state, we can select $N!$ regions where the particles are well separated from each other, i.e. $\mid x_i - x_{i+1} \mid \gg \xi$, so that we can neglect all relativistic effects induced by virtual processes. Each of these domains is identified by the ordering $x_{i_1} \ll x_{i_2} \ll x_{i_3} \ll \cdots \ll x_{i_N}$ of the coordinates of the particles. In this region, the expression for the wavefunction is particularly simple, since it is given by plane waves

$$\Psi(x_{i_1}, x_{i_2}, \cdots x_{i_N}) \;=\; \prod_{k=1}^{N} e^{i\, p_{i_k}\, x_{i_k}}. \tag{19.3.1}$$

Notice that the exchange of two particles maps one domain into another, and each of these transitions is equivalent to multiplying the wavefunction by the corresponding scattering amplitude. Imposing the periodic (antiperiodic) boundary condition for the wavefunction of the bosonic (fermionic) particles, we have the quantization condition[2] for the momenta p_i:

$$e^{i\, p_i L} \prod_{j \neq i}^{N} S(\theta_i - \theta_j) \;=\; \pm 1, \qquad i = 1, 2, \ldots, N. \tag{19.3.2}$$

Using the rapidity variable to express the momenta and considering the terms in the exponents of this equation, we can write it as

$$m_i\, L\, \sinh\theta_i + \sum_{j \neq i}^{N} \delta_{ij}(\theta_i - \theta_j) \;=\; 2\pi n_i, \tag{19.3.3}$$

where

$$\delta_{ij}(\theta) \;=\; -i \ln\, S_{ij}(\theta).$$

[2]Notice that, in the absence of interactions, this leads to the usual quantization condition of the momenta in a finite volume, $p_i \;=\; 2\pi n_i / L$.

The numbers $\{n_i\}$ assume integer values for the bosons and half-integers for the fermions. Together with the rapidity values that solve eqn (19.3.3), they identify the states of the Bethe ansatz

$$| \, n_1, \theta_1; \, n_2, \theta_2; \ldots; n_N, \theta_N \rangle.$$

The energy and the momentum of these states are

$$E = \sum_{i=1}^{N} m_i \, \cosh \theta_i, \qquad p = \sum_{i=1}^{N} m_i \, \sinh \theta_i. \tag{19.3.4}$$

Both these quantities have the *same* expression as in the case of N free particles. The difference, though, is that in the free case, the rapidities of the particles can take arbitrary values, whereas in the interacting case, their values are determined by the quantization relation (19.3.3), in which the phase shifts $\delta_{ab}(\theta)$ of the scattering processes enter.

19.3.1 Selection Rules

The Bethe wavefunction must be symmetric (antisymmetric) under the exchange of two identical bosons (fermions) with the same value of their rapidities. It is then necessary to consider the selection rules coming from the identity of the particles. Since for the diagonal S-matrices the unitarity condition implies $S_{aa}^2(0) = 1$, there could be two different cases:

1. In the first case,
$$S_{aa}(0) = -1,$$
and this leads to a wave-function that is *antisymmetric* under the exchange of two particles with the same rapidity. If the two particles are bosons, this is clearly in conflict with their Bose statistics. This implies that two bosons A_a cannot have the same value of the rapidity, namely each value of θ can be assigned at most to one particle only. Hence all integers $n_i^{(a)}$ of the species a in eqn (19.3.3) must be different. Vice versa, if the identical particles are fermions, the antisymmetry of the wavefunction perfectly matches their Fermi–Dirac statistics and there is no restriction on the integers $n_i^{(a)}$. In the context of the Bethe ansatz, the condition $S = -1$ is called *fermionic type*, independently of the bosonic or the fermionic nature of the particles A_a.
2. In the second case,
$$S_{aa}(0) = 1, \tag{19.3.5}$$
the situation is opposite to the previous one: this condition gives rise to a *symmetric* wavefunction under the exchange of two particles of the same species with the same rapidity. Hence, if the two particles are bosons, this is compatible with their Bose statistics and there is no restriction on the integers $n_i^{(a)}$. Vice versa, if the two particles are fermions, each value of the rapidity can be taken only by one particle, i.e. all integers $n_i^{(a)}$ of the species a must necessarily be different. In the context of the Bethe ansatz, the condition $S = +1$ is called *bosonic type*, independently of the bosonic or the fermionic nature of the particles A_a.

19.4 Derivation of Thermodynamics

The quantization conditions (19.3.3) for the rapidities of the particles form a complicated set of transcendental equations. They simplify in the thermodynamic limit, on which both $L \to \infty$ and the total number of particles $N_a \to \infty$ but keeping their ratio fixed. In such a limit, the spectrum of the rapidities, the solutions of eqn (19.3.3), becomes dense and the distance between two adjacent levels is of order $(\theta_i - \theta_{i+1}) \sim 1/mL$. It is convenient to introduce the continuous densities $\rho_a^{(r)}(\theta)$ relative to the distribution of the rapidities of the particles, defined as the number of particles A_a with rapidity between θ and $\theta + \Delta\theta$ divided by $L\Delta\theta$. In terms of these densities, the energy per unit length of the system can be written as

$$E[\rho^{(r)}] \;=\; \sum_{a=1}^{n} \int_{-\infty}^{+\infty} m_a \,\cosh\theta\, \rho_a^{(r)}(\theta)\, d\theta. \tag{19.4.1}$$

For $m_1 L = L/\xi \gg 1$, the quantization equation (19.3.3) becomes

$$\frac{m_a}{2\pi}\, \sinh\theta_i^{(a)} + \sum_{b=1}^{n} (\delta_{ab} * \rho_b^{(r)})(\theta) \;=\; \frac{n_i^{(a)}}{L}, \tag{19.4.2}$$

where $*$ denotes the convolution of the functions

$$(f * g)(\theta) \;=\; \int_{-\infty}^{+\infty} \frac{d\theta'}{2\pi} f(\theta - \theta')\, g(\theta').$$

Each time that $n_i^{(a)}$ is a set of admissible quantum numbers, the corresponding solution $\theta_i^{(a)}$ of (19.4.2) is said to be a *root* of the species a and the density of these solutions around the value θ is denoted by the function $\rho_a^{(r)}(\theta)$ introduced above. However, these equations admit solutions in $\theta_i^{(a)}$ also for integer values of $\tilde{n}_i^{(a)}$ that are necessarily in relation to the occupied states. Such solutions, associated to the integers $\tilde{n}_i^{(a)}$ that do not correspond to the admissible quantum numbers, are called *holes* of the species a and their density around the value θ is denoted by $\rho_a^{(h)}(\theta)$. The possibility of having these two types of solution is due, in definitive, to two circumstances. The first is that

$$J_a(\theta) \;=\; \frac{m_a}{2\pi}\, \sinh\theta_i^{(a)} + \sum_{b=1}^{n} (\delta_{ab} * \rho_b^{(r)})(\theta) \tag{19.4.3}$$

are monotonically increasing functions (see Fig. 19.2), as we will show later. The second is the absence of certain integers in the sequence of the quantum numbers $n_i^{(a)}$ of the physical states. This derives from the previous discussion on the selection rules: for instance, in the case of bosonic particles but with $S(0) = -1$, choosing an ordering for the variables $\theta_i^{(a)}$, the integers $n_i^{(a)}$ of the physical states must necessarily be a strictly increasing sequence and some integers may be missed in this sequence.

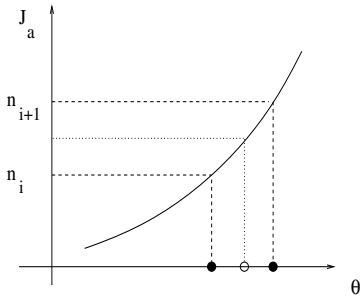


Fig. 19.2 *Plot of the function $J_a(\theta)$ and graphical solution of eqn (19.4.2) for three different integers, where two of them, n_i and n_{i+1}, are admissible quantum numbers. The roots are given by • while the holes are ∘.*

Therefore, in the thermodynamic limit there are densities of both roots and holes. The total density ρ_a of the occupied and empty levels of the particle A_a is equal to the derivative of the functions $J_a(\theta)$

$$\rho_a(\theta) \;=\; \rho_a^{(r)} + \rho_a^{(h)} \;=\; \frac{d}{d\theta} J_a(\theta) \;=\; \frac{1}{2\pi}\, m_a \,\cosh\theta + \sum_{b=1}^{n} (\varphi_{ab} * \rho_a^{(r)})(\theta), \qquad (19.4.4)$$

where

$$\varphi_{ab}(\theta) \;=\; \frac{d}{d\theta}\delta_{ab}(\theta). \qquad (19.4.5)$$

Properties of the functions $\varphi_{ab}(\theta)$. The functions $\varphi_{ab}(\theta)$ satisfy

$$\varphi_{ab}(-\theta) \;=\; \varphi_{ab}(\theta),$$

as can be seen using the unitarity of the amplitudes S_{ab}. For an S-matrix

$$S_{ab}(\theta) \;=\; \prod_{\alpha \in \mathcal{A}_{ab}} s_\alpha(\theta)$$

expressed in terms of the functions $s_\alpha(\theta)$

$$s_\alpha(\theta) = \frac{\sinh\theta + i\sin\alpha\pi}{\sinh\theta - i\sin\alpha\pi}$$

we have

$$\varphi_{ab}(\theta) \;=\; \sum_{\alpha \in \mathcal{A}_{ab}} \varphi_\alpha(\theta), \qquad (19.4.6)$$

where we have defined

$$\varphi_\alpha(\theta) \;=\; -i\,\frac{d}{d\theta}\log s_\alpha(\theta) \;=\; -\,\frac{\sin\alpha\pi}{\cosh\theta-\cos\alpha\pi}. \tag{19.4.7}$$

It is easy to see that $\varphi_{ab}(\theta)$ are periodic functions with period $2\pi i$. For $\theta \neq 0$, they can be written as

$$\varphi_{ab}(\theta) = -\sum_{s=1}^{\infty} \varphi_{ab}^{(k)} e^{-k|\theta|}, \tag{19.4.8}$$
$$\varphi_{ab}^{(k)} = 2 \sum_{\alpha\in\mathcal{A}_{ab}} \sin(k\pi\alpha).$$

Notice that, inserting this expansion into the logarithm derivative of the bootstrap equation (17.4.12), we have

$$\varphi_{il}^{(k)} \;=\; \varphi_{ij}^{(k)}\, e^{-ik\overline{u}_{jl}^{k}} + \varphi_{ik}^{(k)}\, e^{ik\overline{u}_{lk}^{j}}. \tag{19.4.9}$$

Comparing now with the consistency equations of the conserved charges, eqn (17.5.3), one sees that the linearly independent columns and rows of the matrix $\varphi^{(k)} = (\varphi_{ab}^{(k)})$ are solutions (although sometimes trivial) of these equations. The connection between $\varphi_{ab}^{(k)}$ and the eigenvalues $\chi_s^{(a)}$ of the conserved charges is established by

$$\varphi_{ab}^{s} \;=\; \varphi_{11}^{(s)}\,\chi_s^{(a)}\,\chi_s^{(b)}. \tag{19.4.10}$$

The index 1 of this formula refers to the particle of the theory with the lowest mass, with the normalization of the conserved charges set by $\chi_s^{(1)} = 1$. Note that for $s = 1$, eqn (19.4.10) reduces to ($\hat{m}_a = m_a/m_1$)

$$\varphi_{ab}^{1} \;=\; \varphi_{11}^{(1)}\,\hat{m}_a\,\hat{m}_b. \tag{19.4.11}$$

In the thermodynamic limit, there is a large number $N_a \sim L\,\rho_a(\theta)\,\Delta\theta$ of levels in each interval $\Delta\theta$ of the rapidities and there are about $n_a \sim L\,\rho_a^{(r)}\Delta\theta$ particles distributed among them. Since these densities are not strongly influenced by the local redistributions of the particles, the number of different ways of distributing the particles among these levels is given by

$$\Omega_a \;=\; \frac{[L\,\rho_a(\theta)\,\Delta\theta]!}{[L\rho_a^{(r)}(\theta)\Delta\theta]!\,[L\rho_a^{(h)}(\theta)\Delta\theta]!},$$

in the fermionic case and by

$$\Omega_a \;=\; \frac{[L(\rho_a(\theta)+\rho_a^{(r)}(\theta)-1)\,\Delta\theta]!}{[L\rho_a^{(r)}(\theta)\Delta\theta]!\,[L(\rho_a(\theta)-1)\Delta\theta]!},$$

in the bosonic case. Correspondingly, the entropy per unit length $\mathcal{S} = \ln(\prod_a \Omega_a)$ is expressed by

$$\mathcal{S}_{fermi}[\rho, \rho^{(r)}] \;=\; \sum_{a=1}^{n} \int_{-\infty}^{+\infty} d\theta [\rho_a \,\ln \rho_a - \rho_a^{(r)} \,\ln \rho_a^{(r)} - (\rho_a - \rho_a^{(r)}) \,\ln(\rho_a - \rho_a^{(r)})],$$

$$\mathcal{S}_{bose}[\rho, \rho^{(r)}] \;=\; \sum_{a=1}^{n} \int_{-\infty}^{+\infty} d\theta [(\rho_a + \rho_a^{(r)}) \,\ln(\rho_a + \rho_a^{(r)}) - \,\rho_a \ln \rho_a - \rho_a^{(r)} \,\ln(\rho_a^{(r)})].$$

In terms of the densities ρ_a and $\rho_a^{(r)}$, the free energy per unit length is given by the functional

$$f[\rho, \rho^{(r)}] \;=\; E[\rho^{(r)}] - T\,\mathcal{S}[\rho, \rho^{(r)}]. \tag{19.4.12}$$

To derive the thermodynamics of the system at its thermal equilibrium with temperature $T = 1/R$, it is necessary to minimize the free energy with respect the two densitites ρ_a and $\rho_a^{(r)}$, subjected to the constraint (19.4.4). This minimization problem can be solved by using a Lagrange multiplier and can be elegantly expressed by introducing the *pseudo-energies* $\epsilon_a(\theta)$, defined in the two cases by the formulas

$$\frac{\rho_a^{(r)}}{\rho_a} \;=\; \frac{e^{-\epsilon_a}}{1 + e^{-\epsilon_a}}, \quad e^{-\epsilon_a} \;=\; \frac{\rho_a^{(r)}}{\rho_a - \rho_a^{(r)}} \qquad \text{fermionic case} \tag{19.4.13}$$

$$\frac{\rho_a^{(r)}}{\rho_a} \;=\; \frac{e^{-\epsilon_a}}{1 - e^{-\epsilon_a}}, \quad e^{-\epsilon_a} \;=\; \frac{\rho_a^{(r)}}{\rho_a + \rho_a^{(r)}} \qquad \text{bosonic case.} \tag{19.4.14}$$

Using these quantities, the extremum condition reduces to the integral equation

$$m_a\, R \cosh\theta \;=\; \epsilon_a(\theta) \pm \sum_{b=1}^{n} \int \varphi_{ab}(\theta - \theta') \,\log(1 \pm e^{-\epsilon_b(\theta')}) \,\frac{d\theta'}{2\pi}, \tag{19.4.15}$$

where the upper sign refers to the fermionic case and the lower one to the bosonic case. The free energy at equilibrium is then given by

$$f(R) \;=\; \mp \frac{1}{R} \sum_{a=1}^{n} \int_{-\infty}^{+\infty} m_a \,\cosh\theta \,\log\left(1 \pm e^{-\epsilon_a(\theta)}\right) \frac{d\theta}{2\pi}, \tag{19.4.16}$$

where $\epsilon_a(\theta)$ is solution of the integral equation (19.4.15). Therefore the partition function is expressed by

$$Z(L, R) \;=\; \exp\left[\pm L \sum_{a=1}^{n} \int_{-\infty}^{+\infty} m_a \cosh\theta \,\log(1 \pm e^{-\epsilon_a(\theta)}) \,\frac{d\theta}{2\pi}\right]. \tag{19.4.17}$$

We have thus achieved the complete determination of the thermodynamics of the integrable models with diagonal S-matrix. In the following sections we will analyze the behavior of the free energy in different regimes of r and we will study some significant examples.

It is useful to accompany the derivation of thermodynamics given above with a series of comments. The first comment concerns the conceptual difference that exists between the energy levels of free theories and interacting integrable theories. For free theories the levels are simply determined by the quantization of the states of one particle and they can be either empty or occupied by one or more of the N particles of the system, while in integrable models the levels are instead determined in a self-consistent way with the statistical distribution of the particles themselves. This is the reason behind the nonlinear integral equations (19.4.15) for the pseudo-energies ϵ_a. Notice that these quantities determine the distribution of the particles but, in turn, they are also determined by them, as can be seen from their definition, eqns (19.4.13) and (19.4.14).

The second comment concerns some mathematical properties of the pseudo-energies. It is interesting to note that, even though $\epsilon_a(\theta)$ satisfy a nonlinear integral equation, their derivatives $\partial_R \epsilon_a$ and $\partial_\theta \epsilon_a$ are instead solutions of linear integral equations. Differentiating eqn (19.4.15) with respect to R, we have in fact

$$\partial_R \epsilon_a \;=\; m_a \,\cosh\theta + \sum_{b=1}^{n} \int_{-\infty}^{+\infty} \varphi_{ab}(\theta-\theta')\, \frac{e^{-\epsilon_b(\theta')}}{1 \pm e^{-\epsilon_b(\theta')}}\, \partial_R \epsilon_b(\theta')\, \frac{d\theta'}{2\pi}, \tag{19.4.18}$$

and, analogously

$$\partial_\theta \epsilon_a \;=\; m_a\, R\, \sinh\theta + \sum_{b=1}^{n} \int_{-\infty}^{+\infty} \varphi_{ab}(\theta-\theta')\, \frac{e^{-\epsilon_b(\theta')}}{1 \pm e^{-\epsilon_b(\theta')}}\, \partial_\theta \epsilon_b(\theta')\, \frac{d\theta'}{2\pi}. \tag{19.4.19}$$

These equations can be written in compact form by defining the integral operators $\hat{K}_a^{\pm}$, whose kernel is

$$\hat{K}_a^{\pm}(\theta,\theta') \;=\; \sum_{b=1}^{n} \varphi_{ab}(\theta-\theta')\; \frac{e^{-\epsilon_b(\theta')}}{1 \pm e^{-\epsilon_b(\theta')}}. \tag{19.4.20}$$

Hence

$$\begin{array}{l} (1 - \hat{K}_a^{\pm})\, \partial_R \epsilon_a \;=\; e_a, \\ (1 - \hat{K}_a)^{\pm})\, \frac{1}{R} \partial_R \epsilon_a \;=\; k_a, \end{array} \tag{19.4.21}$$

where

$$(\hat{K}\, \partial_R \epsilon) = \int \hat{K}(\theta,\theta')\, \partial_R \epsilon(\theta')\, \frac{d\theta'}{2\pi}$$

(analogously for $(\hat{K}\partial_\theta \epsilon)$), with the notation $e_a = m_a \cosh\theta$, $k_a = m_a \sinh\theta$. Equations (19.4.21) are linear integral equations for the quantities $\partial_R \epsilon_a$ and $\partial_\theta \epsilon_a$ since the functions $\epsilon_b(\theta)$ entering the definition of the kernel $\hat{K}_a^{\pm}$ are regarded as assigned functions, known once the original integral equation (19.4.15) is solved. It is possible to invert eqn (19.4.21) by introducing the resolvents $\hat{L}_a^{\pm}$ that satisfy

$$(1 - \hat{K}_a^{\pm})(1 - \hat{L}_a^{\pm}) \;=\; 1.$$

In this way

$$\partial_R \epsilon_a \,=\, (1+\hat{L}_a^{\pm})\, e_a, \quad \frac{1}{R}\partial_\theta \epsilon_a \,=\, (1+\hat{L}_a^{\pm})\, k_a. \tag{19.4.22}$$

Since $1+\hat{L}_a \,=\, \sum_{n=0}^{\infty} \hat{K}^n$, $\partial_R \epsilon_a$ are then expressed by the Fredholm series

$$\partial_R \epsilon_a \,=\, e_a + \sum_{b=1}^{n} \varphi_{ab} * \left(\frac{e^{-\epsilon_b}}{1+e^{-\epsilon_b}}\, e_a \right) + \cdots,$$

with an analogous result for $\frac{1}{R}\partial_\theta \epsilon_a$.

The third comment refers to the nature of the system of TBA equations for the S-matrices of fermionic and bosonic type. Till now we have presented on the same footing the two cases but there is strong reason to believe that the only consistent interacting theories are those of fermionic type, with $S_{aa}(0) = -1$. In other words, the only diagonal bosonic type S-matrix that gives rise to a consistent set of TBA equations is given by the free theory, for which we have identically $S = 1$. From the mathematical point of view, the problem with the bosonic type TBA equations comes from the term $\log(1 - e^{-\epsilon_a})$, present in the integral of eqn (19.4.15) that determined the pseudo-energies. If it happens that, varying r, one of the ϵ_a becomes negative in an interval of θ, the TBA equations give rise to complex solutions that do not have a natural physical interpretation.

In light of these remarks, hereafter we focus our attention on TBA systems of fermionic type. In this case, it is easy to prove the statement previously made on the monotonic nature of the functions $J_a(\theta)$. In the TBA of fermionic type, eqn (19.4.15) satisfied by $\epsilon_a(\theta)$ implies that these are real functions of θ for any value of r. This also implies the positivity of the densities $\rho_a^{(r)}$ and $\rho_a(\theta)$. But these functions are just the derivatives of the functions $J_a(\theta)$, so that $\frac{d}{d\theta} J_a(\theta) > 0$ and $J_a(\theta)$ are strictly increasing functions.

19.5 The Meaning of the Pseudo-energy

The pseudo-energies $\epsilon_a(\theta)$ admit an interesting physical interpretation. It is necessary to initially note that the final expression of the partition function given by the TBA, eqn (19.4.17), is formally identical to the partition function of a gas of *free* quasi-particles,[3] the only difference being that the energy of each of these particles is given by $\epsilon_a(\theta)/R$ rather than $m_a \cosh\theta$. This observation suggests that, in integrable theories, the only effect of the temperature consists of modifying the excitation energies of the particles, which are now measured with respect to the thermal ground state of the system. To better clarify this observation and to understand the nature of the pseudo-energies, it is convenient to analyze the simplest case of a system with only one species of particles. In the following we are going to show that there exists a one-to-one correspondence between the partition function as obtained by the TBA

$$Z(L,R) = \exp\left[\pm mL \int \frac{d\theta}{2\pi} m_a \cosh\theta \, \log(1 \pm e^{-\epsilon(\theta)}) \right]. \tag{19.5.1}$$

[3]The dispersion relations of the free quasi-particle differ from those of the usual particle and they are given in eqn (19.5.7) below.

and the partition function as given by the usual sum over the states

$$Z(L,R) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi} \langle \theta_n \cdots \theta_1 | \theta_1 \cdots \theta_n \rangle \prod_{i=1}^{n} e^{-\epsilon(\theta_i)}, \tag{19.5.2}$$

where the scalar product is computed using the standard rules relative to the free fermionic or bosonic cases (depending on the type of TBA equations), while the energy of the particles is given by the pseudo-energy $\epsilon(\theta)/R$. Let's study the fermionic case, leaving the derivation of the bosonic case as an exercise to the reader. Let's start by defining

$$F(R) = \int \frac{d\theta}{2\pi} \cosh\theta \log(1 + e^{-\epsilon(\theta)}). \tag{19.5.3}$$

This functions can be expanded in series as

$$F(R) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{n} I_n(R), \tag{19.5.4}$$

where

$$I_n(R) \equiv \int \frac{d\theta}{2\pi} \cosh\theta \, e^{-n\epsilon(\theta)}.$$

The TBA partition function has a power series expansion in powers of (mL):

$$Z(L,R) = 1 + (mL)F(R) + \frac{(mL)^2}{2!}(F(R))^2 + \cdots \frac{(mL)^n}{n!}(F(R))^n + \cdots \tag{19.5.5}$$

Let's now compute the partition function using its alternative definition, given in eqn (19.5.2). We need to employ a regularization of the square of the δ function (which enters the scalar product $\langle \theta_n \cdots \theta_1 | \theta_1 \cdots \theta_n \rangle$) provided by the free fermionic theory defined on a sufficiently large volume L

$$[\delta(\theta - \theta')]^2 \equiv \frac{mL}{2\pi} \cosh(\theta) \, \delta(\theta - \theta'). \tag{19.5.6}$$

We then have

$$Z(L,R) = 1 + Z_1 + Z_2 + \cdots Z_n + \cdots$$

where the first terms are given by

$$\begin{aligned} Z_1 &= \int \frac{d\theta}{2\pi} \langle \theta | \theta \rangle e^{-\epsilon(\theta)} \; = \; \int \frac{d\theta}{2\pi} d\theta' \delta(\theta - \theta') \langle \theta' | \theta \rangle e^{-\epsilon(\theta)} \\ &= mL \int \frac{d\theta}{2\pi} \cosh\theta \, e^{-\epsilon(\theta)} \; = \; (mL) \, I_1; \end{aligned}$$

$$\begin{aligned} Z_2 &= \frac{1}{2} \int \frac{d\theta_1}{2\pi} \frac{d\theta_2}{2\pi} \langle \theta_2 \theta_1 | \theta_1 \theta_2 \rangle e^{-\epsilon(\theta_1) - \epsilon(\theta_2)} \\ &= \frac{1}{2} \int \frac{d\theta_1}{2\pi} \frac{d\theta_2}{2\pi} \left[(2\pi)^2 (\delta(\theta_1 - \theta_1)\delta(\theta_2 - \theta_2) - \delta(\theta_1 - \theta_2)\delta(\theta_2 - \theta_1)) \right] e^{-\epsilon(\theta_1) - \epsilon(\theta_2)} \\ &= \frac{1}{2}(mL)^2 I_1^2 - \frac{1}{2}(mL) I_2. \end{aligned}$$

Similarly,

$$Z_3 = \frac{(mL)^3}{3!} I_1^4 - \frac{(mL)^2}{2} I_1 I_2 + \frac{(mL)}{3} I_3,$$

$$Z_4 = \frac{(mL)^4}{4!} I_1^4 - (mL)^3 I_1^2 I_2 + \frac{(ML)^2}{2} \left[\frac{2}{3} I_1 I_3 + \left(\frac{I_2}{2} \right)^2 \right] - \frac{(mL)}{4} I_4.$$

It is not difficult to extend the computation to a higher order and show that the series (19.5.2) precisely coincides with that given in eqn (19.5.5), the only difference being the different arrangement of their terms. In fact, summing all those proportional to (mL) present in each Z_n, one recovers the function $F(R)$, while the sum of all the terms proportional to $(mL)^k$ present in each Z_n precisely reproduces the higher powers $(F(R))^k$.

This important result implies that all physical properties of the system depend on the quasi-particle excitations with respect to the ground state of the TBA. These excitations have an *effective energy* $\tilde{e} = \epsilon(\theta)/R$ and an *effective momentum* $\tilde{k}(\theta)$ given by

$$\begin{aligned} \tilde{e}(\theta) &= \epsilon(\theta)/R, \\ \tilde{k}(\theta) &= k(\theta) + 2\pi(\delta * \rho_1)(\theta). \end{aligned} \tag{19.5.7}$$

Hence, in the presence of the temperature, one may regard the rapidity θ as the parameter that expresses the dispersion relations of the quasi-particle excitations. This result, derived in the non-relativistic case by C.N. Yang and C.P. Yang, can be easily generalized to the relativistic case, as shown in the box below.

Dressed energy and momentum Let (n_j, θ_j) and (n'_j, θ'_j) be two Bethe states satisfying eqn (19.3.3), where $n'_j = n_j$ except for $j = \alpha$. Subtracting the two equations, we have

$$mL\,(\sinh\theta'_j - \sinh\theta_j) = \sum_i [\delta(\theta_j - \theta_i) - \delta(\theta'_j - \theta'_i)] \tag{19.5.8}$$

$(j \neq \alpha)$. Since $\theta'_j \approx \theta_j$, we can introduce a function $\chi(\theta)$ and write

$$L\,(\sinh\theta'_j - \sinh\theta_j) \approx \chi(\theta_j)\cosh\theta_j. \tag{19.5.9}$$

In the thermodynamic limit, eqn (19.5.8) can be written as

$$2\pi(1 - \hat{K})\,(\rho\chi) = \delta(\theta - \theta_\alpha) - \delta(\theta - \theta'_\alpha), \tag{19.5.10}$$

where ρ is the density of the levels and $\hat{K}$ is the integral operator defined in eqn (19.4.20), here restricted to the case in which there is only one species of particles. We also have

$$\sigma(\theta-\theta_\alpha)-\sigma(\theta-\theta'_\alpha) \;=\; \int_{\theta_\alpha}^{\theta'_\alpha} d\theta'\,\hat{K}(\theta,\theta')\,(1+e^{\epsilon(\theta')}). \tag{19.5.11}$$

Using the resolvent $\hat{L}$ of $\hat{K}$, we can invert (19.5.10):

$$\rho\,\chi(\theta) \;=\; \int_{\theta_\alpha}^{\theta'_\alpha} \frac{d\theta'}{2\pi}\,\hat{L}(\theta,\theta')\,(1+e^{\epsilon(\theta')}). \tag{19.5.12}$$

Consider now the difference in energy ΔE between the two Bethe states

$$\begin{aligned}\Delta E &= \sum_j m\,[\cosh\theta'_j - \cosh\theta_j] \qquad (19.5.13)\\ &= m\,\cosh\theta'_\alpha - m\,\cosh\theta_\alpha + m\int d\theta\,\sinh\theta\,\frac{\chi(\theta)\rho(\theta)}{1+e^{\epsilon(\theta)}}.\end{aligned}$$

Substituting (19.5.12) and using the property $\hat{L}(\theta,\theta')(1+e^{\epsilon(\theta')}) = (1+e^{\epsilon(\theta)})\hat{L}(\theta',\theta)$, together with (19.4.22), we have

$$\Delta E \;=\; \frac{1}{R}\,(\epsilon(\theta'_\alpha)-\epsilon(\theta_\alpha)). \tag{19.5.14}$$

The dressed momentum is obtained in a similar way

$$\begin{aligned}\Delta P &= \sum_j m\,[\sinh\theta'_j - \sinh\theta_j] \qquad (19.5.15)\\ &= m\,\sinh\theta'_\alpha - m\,\sinh\theta_\alpha + m\int d\theta\,\cosh\theta\,\frac{\chi(\theta)\rho(\theta)}{1+e^{\epsilon(\theta)}}.\end{aligned}$$

Substituting again (19.5.12) and using eqn (19.4.22), one finds

$$\Delta P = \tilde{k}(\theta'_\alpha)-\tilde{k}(\theta_\alpha) \tag{19.5.16}$$

where $\tilde{k}$ is defined in eqn (19.5.7).

The interpretation given above of the pseudo-energy finds interesting application in the computation of the correlation functions of the integrable models at finite temperature.

19.6 Infrared and Ultraviolet Limits

In this section we study in more detail the scaling function $\tilde{c}(r)$ for the TBA systems of fermionic type. Using eqns (19.2.5) and (19.2.7), this function is given by

$$\tilde{c}(r) \;=\; \frac{3}{\pi^2} r \sum_{a=1}^{n} \hat{m}_a \int_{-\infty}^{+\infty} L_a(\theta)\, \cosh\theta \, d\theta, \tag{19.6.1}$$

where we have defined $\hat{m}_a = m_a/m_1$ and

$$L_a(\theta) \;=\; \log\left(1 + e^{-\epsilon_a(\theta)}\right).$$

It is easy to find the behavior of this function for large values of r (the infrared limit): we have in fact

$$\epsilon_a(\theta) \;\simeq\; \hat{m}_a r\, \cosh\theta, \quad L_a(\theta) \;\simeq\; e^{-\hat{m}_a\, r \cosh\theta}, \tag{19.6.2}$$

so that $\tilde{c}(r)$, for $r \to \infty$, behaves as

$$\tilde{c}(r) \;\simeq\; \frac{6}{\pi^2}\, r \sum_{a=1}^{n} \hat{m}_a \int_0^{+\infty} d\theta\, \cosh\theta\, e^{-r\hat{m}_a \cosh\theta} \;=\; \frac{6}{\pi^2}\, r \sum_{a=1}^{n} \hat{m}_a\, K_1(\hat{m}_a r), \tag{19.6.3}$$

where $K_1(z)$ is the modified Bessel function. For $r \to \infty$ the Bessel function decreases exponentially and the behavior of the system is that of a free theory made of n particles of masses m_a (see Section 19.11).

The opposite limit, $r \to 0$, corresponds to the ultraviolet or conformal limit[4] of the massive theory. To compute the value of the scaling function $\tilde{c}(r)$ of this limit, we need to study some properties of the integral equation (19.4.15).

The solutions $\epsilon_a(\theta)$ are even functions of θ. For $r \to 0$, they flatten and become constant in the region $-\ln\frac{2}{r} \ll \theta \ll \ln\frac{2}{r}$, whereas they tend to the free values (19.6.2) outside this interval.[5] The constant values ϵ_a can be found by solving the transcendental equation

$$\epsilon_a \;=\; \sum_{b=1}^{n} N_{ab}\, \ln\left(1 + e^{-\epsilon_b}\right), \tag{19.6.4}$$

where N_{ab} is the positive symmetric matrix given by

$$N_{ab} \;=\; -\int_{-\infty}^{+\infty} \frac{d\theta}{2\pi} \varphi_{ab}(\theta) \;=\; -\frac{1}{2\pi}\left[\delta_{ab}(+\infty) - \delta_{ab}(-\infty)\right]. \tag{19.6.5}$$

For $r \to 0$, the plateau of the curve enlarges and the situation appears as in Fig. 19.3. In this limit, the curve rapidly decreasing outside the plateau assumes a universal shape. This can be determined noting that, for large values of θ, the right-hand side of eqn (19.4.15) can be written as

$$\hat{m}_a\, r\, \cosh\theta \;\sim\; \frac{1}{2}\hat{m}_a\, r\, e^{\theta} \;=\; \hat{m}_a\, e^{\left(\theta - \ln\frac{2}{r}\right)},$$

[4]Since r is a scaling variable, given by $r = m_1 R = R/\xi$, the limit $r \to 0$ can be equivalently regarded as the limit in which the correlation length ξ diverges, $\xi \to \infty$.

[5]This is actually the situation for the TBA equations relative to the minimal S-matrices, while for the S-matrices associated to lagrangian models, as for instance the S-matrix of the Sinh–Gordon or Toda models, the functions $\epsilon_a(\theta)$ diverge as $\log(m_a r)$ in the limit $r \to 0$.

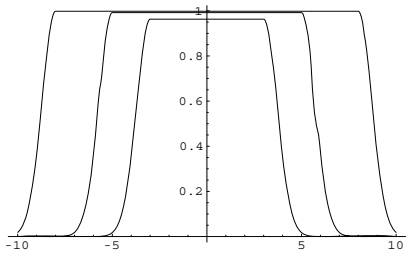


Fig. 19.3 *Behavior of the function $L_a(\theta)$ when $r \to 0$.*

and the dependence on r of the functions $\epsilon_a(\theta)$ reduces then to a simple shift[6]

$$\theta \;\to\; \theta - \frac{2}{r}.$$

Hence, for $r \to 0$, their behavior at the edges of the interval is universal and dictated by the equation

$$\hat{m}_a\, e^{\theta} \;=\; \tilde{\epsilon}_a(\theta) + \sum_{b=1}^{n} (\varphi_{ab} * \tilde{L}_b)(\theta), \tag{19.6.6}$$

where the functions $\tilde{\epsilon}_a(\theta)$ assume the constant values ϵ_a for $\theta \ll \frac{2}{r}$ and increase exponentially to infinity when $\theta \to \infty$. The corresponding functions $\tilde{L}_a(\theta)$ interpolate bewteen 0 and their limiting value given in eqn (19.6.4). For this reason, the universal functions $\tilde{\epsilon}_a$ are called the *kink solutions* of the TBA equations. Expressed in terms of these functions, the value of the scaling function $\tilde{c}(r)$ at $r = 0$ assumes the form

$$\tilde{c}(0) \;=\; \frac{6}{\pi^2} \sum_{a=1}^{n} \int_0^{\infty} d\theta\, \tilde{L}_a(\theta)\, \hat{m}_a\, e^{\theta}. \tag{19.6.7}$$

Substituting for $\hat{m}_a e^{\theta}$ the derivative of the left-hand side of eqn (19.6.6)

$$\hat{m}_a\, e^{\theta} \;=\; \frac{d\tilde{\epsilon}_a(\theta)}{d\theta} - \sum_{b=1}^{n} \left(\varphi_{ab} * \frac{e^{-\tilde{\epsilon}_b}}{1 + e^{-\tilde{\epsilon}_b}} \frac{d\tilde{\epsilon}_b}{d\theta} \right)(\theta), \tag{19.6.8}$$

we have

$$\tilde{c}(0) \;=\; \frac{6}{\pi^2} \sum_{a=1}^{n} \int_0^{\infty} d\theta\, \tilde{L}_a(\theta) \left[\frac{d\tilde{\epsilon}_a(\theta)}{d\theta} - \sum_{b=1}^{n} \left(\varphi_{ab} * \frac{e^{-\tilde{\epsilon}_b}}{1 + e^{\tilde{\epsilon}_b}} \frac{d\tilde{\epsilon}_b}{d\theta} \right)(\theta) \right]. \tag{19.6.9}$$

Since $\tilde{\epsilon}_a$ are monotonically increasing functions, the first term on the right-hand side simply becomes

$$\int_0^{\infty} d\theta\, \tilde{L}(\theta)\, \frac{d\tilde{\epsilon}_a(\theta)}{d\theta} \;=\; \int_{\epsilon_a}^{\infty} d\epsilon\, \tilde{L}(\epsilon). \tag{19.6.10}$$

[6]We discuss the behavior of $\epsilon_a(\theta)$ for positive values of θ, since the behavior for negative values can be recovered by parity.

The convolution term in (19.6.9) can be analogously substituted using the same equation (19.6.6). After an integration by parts, the final result is given by

$$\tilde{c}(0) \;=\; \sum_{a=1}^{n} \tilde{c}_a(\epsilon_a), \tag{19.6.11}$$

where

$$\tilde{c}_a(\epsilon_a) \;=\; \frac{6}{\pi^2}\left[\int_{\epsilon_a}^{\infty} dx\, \ln(1+e^{-x}) + \frac{1}{2}\epsilon_a\, \ln(1+e^{-\epsilon_a})\right] = \frac{6}{\pi^2} L\left(\frac{1}{1+e^{\epsilon_a}}\right),$$

and $L(x)$ is the dilogarithm function

$$L(x) \;=\; -\frac{1}{2}\int_0^x dt\, \left[\frac{\ln t}{1-t} + \frac{\ln(1-t)}{t}\right].$$

In conclusion, the effective central charge of the conformal limit of the massive theory with purely elastic S-matrix is obtained through the following steps:

1. solve the transcendental equation (19.6.4);
2. substitute their constant solutions ϵ_a in eqn (19.6.11).

In the next sections we will see some significant examples of this result.

19.7 The Coefficient of the Bulk Energy

In a theory with a mass scale, the additivity of the energy requires a linear growth of the energy of the ground state with the dimension R of the system

$$E_0 \;\sim\; \mathcal{E}_0\, R.$$

$\mathcal{E}_0$, the bulk term, can be interpreted as the singular part of the infinite volume energy due to the fluctuations present in the system. Usually this is not a universal quantity, since it depends on the regularization scheme adopted. However, in a perturbed conformal field theory, the regularization scheme is fixed by the requirement that the off-critical quantities adiabatically go to their conformal values. Hence, in this case, it is possible to extract a universal term $\mathcal{E}_0$ that only depends on the scattering data. Since E_0 is directly related to the scaling function $\tilde{c}(r)$, the bulk term $\mathcal{E}_0$ is given by

$$\mathcal{E}_0 \;=\; -\;\frac{\pi}{12}\, m_1^2\, \frac{1}{r}\frac{d\tilde{c}}{dr}\bigg|_{r=0}. \tag{19.7.1}$$

For the computation of this limit, let's introduce the functions

$$\psi_a(\theta) \;=\; \left(\partial_r + \frac{1}{r}\partial_\theta\right)\epsilon_a(\theta).$$

As discussed at the end of Section 19.4, the functions ψ_a satisfy the linear integral equations

$$\psi_a(\theta) \,=\, \hat{m}_a e^{\theta} + \sum_{b=1}^{n} \left(\varphi_{ab} * \frac{\psi_b}{e^{\epsilon_b}+1} \right)(\theta).$$

Using eqn (19.6.1), one has

$$\frac{1}{r}\frac{d\tilde{c}(r)}{dr} \,=\, -\frac{3}{\pi^2} \sum_{a=1}^{n} \int_{-\infty}^{+\infty} d\theta\, \hat{m}_a\, e^{-\theta} \frac{\psi_a(\theta)}{e^{\epsilon(\theta)}+1}.$$

When $r \to 0$, the integrand is localized near the edge of the flat region and its behavior is determined by the kink solutions $\tilde{L}_a(\theta)$. Hence, we get

$$\frac{1}{r}\frac{d\tilde{c}(r)}{dr}\bigg|_{r=0} \,=\, \frac{3}{\pi^2} \sum_{a=1}^{n} \hat{m}_a \int_{-\infty}^{\infty} d\theta\, e^{-\theta}\, \partial_\theta \tilde{L}_a(\theta) \equiv -\frac{3}{\pi^2} \sum_{a=1}^{n} \hat{m}_a T_a. \tag{19.7.2}$$

To compute the right-hand side of this equation, let's proceed as follows. Considering initially the asymptotic expansion for $\theta \to -\infty$ of the convolution term, we have

$$\sum_{b=1}^{n} (\varphi_{ab} * \tilde{L}_b)(\theta) \,=\, -\epsilon_a + \frac{e^{\theta}}{2\pi} \sum_{b=1}^{n} \varphi_{ab}^{(1)} T_b + \cdots$$

where $\varphi_{ab}^{(1)}$ is the first term of the expansion of these functions, as given in eqn (19.4.8). Comparing with the exponential term in eqn (19.6.8), we arrive at

$$\sum_{b=1}^{n} \varphi_{ab}^{(1)} T_b \,=\, 2\pi.$$

Using now eqn (19.4.11), we obtain

$$\sum_{b=1}^{n} \varphi_{ab}^{(1)} T_b \,=\, \varphi_{11}^{(1)} \hat{m}_a \sum_{b=1}^{n} \hat{m}_b T_b,$$

where $\varphi_{11}^{(1)}$ is the corresponding quantity relative to the lightest particle. Hence the bulk energy term is determined by the S-matrix of the lightest particle as

$$\mathcal{E}_0 \,=\, \frac{m_1^2}{2\varphi_{11}^{(1)}}. \tag{19.7.3}$$

A direct measurement of this quantity can be achieved by a numerical diagonalization of the transfer matrix of the theory.

19.8 The General Form of the TBA Equations

From the diagonal scattering theories related to the simply laced Lie algebras, there is an extremely elegant formulation of the TBA equations. Besides the great level

of generality, this formulation also has the advantage of highlighting the common structure of all these theories. As shown in Chapter 14, the number of particles of these theories is equal to the rank r of the algebra $\mathcal{A}$. Another important quantity to keep in mind is the Coxeter number h of these algebras.

In order to give such a formulation, let's introduce the notation

$$\nu_a(\theta) \;=\; m_a R \cosh\theta, \tag{19.8.1}$$

and note that the original TBA equations of these theories, expressed by

$$-\nu_a + \epsilon_a + \sum_{b=1}^{r} \varphi_{ab} \star \log[1+\exp(-\epsilon_b)] \;=\; 0, \tag{19.8.2}$$

can be rewritten in the universal form

$$-\nu_a + \epsilon_a + 2\sum_{b=1}^{r} I_{ab}\varphi_h \star \{\nu_b - \log[1+\exp(\epsilon_b)]\}, \tag{19.8.3}$$

where I_{ab} is the incidence matrix of the Dynkin diagram of the corresponding algebra $\mathcal{A}$ and $\varphi_h(\theta)$ is the universal kernel

$$\varphi_h(\theta) \;=\; \frac{h}{2\cosh\frac{h\theta}{2}}, \tag{19.8.4}$$

with h the Coxeter number of the relative algebra. The equivalence between the sets of equations (19.8.2) and (19.8.3) is based on the important identity

$$\left(\delta_{ab} - \frac{1}{2\pi}\varphi_{ab}(k)\right)^{-1} \;=\; \delta_{ab} - \frac{1}{2\cosh(k/h)}\, I_{ab} \tag{19.8.5}$$

which holds for the Fourier transforms

$$\varphi_{ab}(k) \;=\; \int_{-\infty}^{+\infty} \varphi_{ab}(\theta)\, \exp(ik\theta)\, d\theta$$

of the original kernels. Another important relation in the derivation of eqn (19.8.3) is provided by

$$\sum_{b=1}^{r} I_{ab}\, m_b \;=\; 2m_a\, \cos\frac{\pi}{h}. \tag{19.8.6}$$

As a by-product of eqn (19.8.5), one of its remarkable consequences is the universal expression of the matrix N_{ab} for the scattering theories associated to the Lie algebras

$$N \;=\; I\,(2-I)^{-1}. \tag{19.8.7}$$

To derive this expression it is sufficient to substitute $k = 0$ in (19.8.5), taking into account that $N_{ab} = -\frac{1}{2}\varphi_{ab}(0)$.

Equation (19.8.3) can now be analytically continued to the values $\theta \pm i\pi/h$. Using the relation

$$\nu_a\left(\theta + i\frac{\pi}{h}\right) + \nu_a\left(\theta - i\frac{\pi}{h}\right) = \sum_{b=1}^{r} I_{ab}\nu_b(\theta), \tag{19.8.8}$$

they can be written in terms of a set of functional equations for the quantities $Y_a(\theta) = \exp[\epsilon_a(\theta)]$

$$Y_a\left(\theta + i\frac{\pi}{h}\right) Y_a\left(\theta + i\frac{\pi}{h}\right) = \prod_{b=1}^{r}[1 + Y_b(\theta)]^{I_{ab}}. \tag{19.8.9}$$

These equations are completely independent of the energy terms $\nu_a(\theta)$ of the particles and involve only the basic information of the algebras encoded in the incidence matrix I_{ab}. Furthermore, they present the periodic properties

$$\begin{array}{ll} Y_a\left(\theta + i\pi\frac{h+2}{h}\right) = Y_{n-a+1}(\theta), & A_n \text{ series} \\ Y_a\left(\theta + i\pi\frac{h+2}{h}\right) = Y_a(\theta), & D_n \text{ and } E_n \text{ series.} \end{array} \tag{19.8.10}$$

For the series A_n, one should keep in mind that the symmetry of these algebras imposes $Y_a(\theta) = Y_{n-a+1}(\theta)$, so that the periodicity condition is also satisfied for this series.

The periodic properties of $Y_a(\theta)$ have important consequences. First of all, they imply that the solutions of the original equations (19.8.2), with $\nu_a(\theta)$ given in (19.8.1), are entire functions of θ and, consequently, they can be expanded in Laurent series

$$Y_a(\theta) = \sum_{n=-\infty}^{\infty} Y_a^{(n)}\, t^n, \tag{19.8.11}$$

where $t = \exp([(2h/(h+2)]\theta)$. These series are convergent on all the complex plane of θ, except at $t = 0$ and $t = \infty$. In particular, for the solutions of eqn (19.8.2), the symmetry $\theta \to -\theta$ requires that $Y_a^{(n)} = Y_a^{(-n)}$. In the t-plane, the functional equation (19.8.9) becomes

$$Y_a(\Omega\, t)Y_a(\Omega^{-1}\, t) = \prod_{b=1}^{r}[1 + Y_b(t)]^{I_{ab}}. \tag{19.8.12}$$

where $\Omega = \exp[2i\pi/(h+2)]$. These are the most general form of the TBA equations and they may have several classes of solutions. Obviously, among the solutions of eqns (19.8.12), there are also those that are entire functions in t. The kink solution, for instance, corresponding to an energy term given by $\nu_a(\theta) = m_a R\exp(\theta)$ instead of (19.8.1), is an example of this set of solutions. Notice that imposing $t = 0$ in (19.8.12), one obtains the algebraic transcendental equations (19.6.4) for the quantities $z_a = \exp[\epsilon_a(0)]$, that are crucial quantities to obtain the effective central charge.

The second consequence of the periodicity of the functions $Y_a(\theta)$ concerns the behavior of the solutions of eqn (19.8.2) when $R \to 0$. In this limit we saw that the functions $\log[1 + \exp[-\epsilon(\theta)]]$ acquire a plateau of height $\log[1 + 1/z_a]$ in the central interval $-\log(1/m_1R) \ll \theta \ll \log(1/m_1R)$ and rapidly tend to zero outside this interval. For $R \to 0$ this plateau enlarges and this implies that the integral equations (19.8.2) or (19.8.3) are somehow local in the central interval in the rapidity space

while, for large R, the two edges of the plateau influence each other only through the wavelength terms fixed by the periodicity. Hence, the function $f(R) = RE_0(R)/2\pi$ admits a regular series expansion with respect the variable $G^2 = (m_1 R)^{4h/(h+2)}$, except for a bulk term energy proportional to R^2

$$f(R) = -\frac{c_{eff}}{12} - \frac{\mathcal{E}_0}{2\pi} R^2 + \sum_{n=1}^{\infty} f_{2n} G^{2n}. \tag{19.8.13}$$

This analytic structure of $f(R)$ is in full agreement with conformal perturbation theory, set in this case by a perturbing operator of conformal weight $\Delta = 1 - h/(h+2)$ (for the scattering theory based on the Lie algebras, the corresponding perturbative terms are given by even powers of the coupling constants).

The analysis of the TBA done for the Lie algebras can be extended to the most general case. In particular, one can show that, besides the bulk term, the expansion of the free energy dictated by the TBA is a regular function with respect to the variable $G = (m_1 R)^{2-2\Delta}$, where Δ is the conformal weight of the perturbing field

$$f(R) = -\frac{c_{eff}}{12} - \frac{\mathcal{E}_0}{2\pi} R^2 + \sum_{n=1}^{\infty} f_n G^n. \tag{19.8.14}$$

Comparing with the perturbative expression of this quantity one can derive an important relationship between the coupling constant and the lowest mass of the theory.

19.9 The Exact Relation $\lambda(m)$

The TBA permits us to determine the exact relation that holds between the coupling constant of a perturbed conformal field theory and its lowest mass. In this section we present the basic idea that leads to this formula, listing afterward the formulas of the various integrable theories.

Consider the action of a perturbed conformal field theory, with the perturbation given by a relevant field of conformal weight Δ

$$\mathcal{S} = \mathcal{S}_0 + \lambda \int d^2x\, \phi(x). \tag{19.9.1}$$

Let's assume that such a deformation defines a massive integrable field theory, characterized by its S-matrix. The coupling constant λ is a dimensional quantity, expressed in terms of the lowest mass m_1 by the relation

$$\lambda = \mathcal{D}\, m_1^{2-2\Delta}. \tag{19.9.2}$$

Once the normalization of the operator $\phi(x)$ is fixed, the coefficient $\mathcal{D}$ is a pure number that can be extracted by the comparison between the TBA and the perturbative series. For the normalization of the operator we take the conformal one, given by

$$\langle \phi(x_1)\phi(x_2) \rangle \simeq \frac{1}{|x_{12}|^{4\Delta}}, \qquad x_{12} \to 0.$$

As seen in the previous sections, the free energy of an integrable theory can be computed in terms of the Bethe ansatz equations and this leads to the general expression

(19.8.14). On the other hand, this quantity can be computed in conformal perturbation theory, using eqn (19.9.1). The corresponding series is

$$f_{pert}(R) \;=\; -c_{eff} - \frac{R^2}{2\pi}\sum_{n=1}^{\infty}\frac{(-\lambda)^n}{n!}\int\langle\phi(X_1)\,\ldots\,\phi(x_n)\rangle_c\,d^2X_1\ldots d^2X_n, \qquad (19.9.3)$$

where $X_i = (x_i, y_i)$ are the coordinates on the cylinder and the connected correlation functions are those of the conformal theory on the cylinder. Using the mapping $z = \exp(-2\pi i\zeta/R)$ where $\zeta = x + iy$ is the complex coordinate of the cylinder, the perturbative terms can be written as integrals of the connected correlation functions on the euclidean plane

$$\begin{aligned} f_{pert}(R) \;=\; -c_{eff} - \frac{R^2}{2\pi}\sum_{n=1}^{\infty}\frac{(-\lambda)^n}{n!}\left(\frac{2\pi}{R}\right)^{2(\Delta-1)n+2} \qquad (19.9.4)\\ \times\int\langle V(0)\phi(z_1,\bar z_1)\ldots\phi(z_n,\bar z_n)V(\infty)\rangle_c\prod_{i=1}^{n}(z_i\,\bar z_i)^{\Delta-1}\,d^2z_1\ldots d^2z_n. \end{aligned}$$

In this expression, $V(z,\bar z)$ is the operator that creates the lowest energy state on the cylinder, i.e. the field associated to Δ_{min} (for the unitary theories, $V = \mathbf{1}$). If the field ϕ is odd under a Z_2 symmetry of the original conformal model – an hypothesis that we will make in the following for simplicity – we have an even series in λ.

All the integrals of the perturbative series are ultraviolet convergent if $\Delta < 1/2$ and, on the cylinder because of its finite size, they are also convergent in the infrared. Using a dimensional argument, it is not difficult to see that the perturbative series is an expansion in the parameter $g^2 \equiv \lambda^2\,R^{2(2-2\Delta)}$

$$f_{pert}(R) = -c_{eff} + \mathcal{F}_2\,g^2 + \mathcal{F}_4 g^4 + \cdots$$

Let's assume that this series converges in a finite domain around the origin where it defines a function $\mathcal{F}(g)$. On the other hand, for thermodynamics reasons, we know that in the limit $R \to \infty$ we have

$$f_{pert}(R) \sim \frac{\mathcal{E}_0}{2\pi}R^2.$$

This behavior is related to the analytic continuation of $\mathcal{F}(g)$ outside the domain of convergence of the original series. Since in quantum field theory the normalization of the free energy is chosen in such a way as to vanish at infinity, it is necessary to subtract the quantity above from the perturbative series, so that the final expression is

$$f(R) \;=\; -c_{eff} - \frac{\mathcal{E}_0}{2\pi}R^2 + \sum_{n=1}\mathcal{F}_{2n}g^{2n}. \qquad (19.9.5)$$

The relation between the coupling constant and the lowest mass is obtained by comparing this series with the original series of the TBA, eqn (19.8.14). Taking the first

term of both and simplifying the common factor $R^{2(2-2\Delta)}$ we have

$$\mathcal{F}_2 \lambda^2 \;=\; f_2\, m_1^{2(2-2\Delta)}. \tag{19.9.6}$$

The proportional coefficient $\mathcal{D}$ between the coupling constant λ and the mass m_1 is then

$$\mathcal{D}^2 \;=\; f_2/\mathcal{F}_2. \tag{19.9.7}$$

This coefficient has been determined exactly for many integrable models. For all the theories related to the simply laced Toda models this has been achieved by Fateev. Let's present the relevant formula. From Section 16.7 we know that the Toda field theories for particular imaginary values of the coupling constant given by

$$\beta^2 \;=\; \frac{p}{p+1} \qquad p = k+h, k+h+1, \ldots$$

describe the integrable deformation of the field of conformal weight

$$\Delta \;=\; 1 - \frac{h}{k+h+1}$$

of the coset model

$$\frac{G_k \times G_1}{G_{k+1}}.$$

Let's denote generically the perturbed action of these models as

$$\mathcal{S} \;=\; \mathcal{S}_{CFT} + \lambda \int d^2x\, \Phi_\Delta(x).$$

For these theories the relation that links λ to the lowest mass of the system is given by

$$\begin{aligned} \lambda \;=\; & \left[\frac{\pi\, m_1\, k(\mathcal{G})\, \Gamma\left(\frac{k+h+1}{h}\right)}{\Gamma\left(\frac{1}{h}\right)\, \Gamma\left(\frac{k+h}{h}\right)} \right]^{4hu} \\ & \times \frac{(1-hu)^{-2}(1-(h+1)u)^2\, \gamma(qu)\, \gamma\left(\frac{h-2q+2}{2}u\right)}{\pi^2\, \gamma(u)\, \gamma\left(\frac{h+2q-2}{2}u\right)\, \gamma((h-q)u))\, \gamma((h+1)u)} \end{aligned} \tag{19.9.8}$$

where h is the Coxeter number of the algebra $\mathcal{G}$, $\gamma(x)$ is given by $\gamma(x) = \Gamma(x)/\Gamma(1-x)$, and

$$k(\mathcal{G}) \;=\; \left(\prod_{i=1}^{r} q_i^{q_i} \right)^{1/2h}, \qquad u \;=\; \frac{1}{k+h+1}.$$

The quantities q_i are the integer numbers that enter the definition of the maximal root of the simply laced algebra, see eqn (16.6.2), whereas $q = \max q_i$. For the ADE one has $q(A) = 1$, $q(D) = 2$, $q(E_6) = 3$, $q(E_7) = 4$, $q(E_8) = 6$.

19.10 Examples

In this section we use the TBA associated to several integrable models to compute the effective central charge that emerges in the ultraviolet limit. As we will see, this gives strong support to the S-matrix description of the off-critical deformations.

19.10.1 Yang–Lee

Besides the free theories, which will be discussed in Section 19.11, the simplest TBA system is provided by the Yang–Lee S-matrix presented in Chapter 18. In this case the kernel is given by

$$\varphi(\theta) \;=\; -\sqrt{3}\left(\frac{1}{2\cosh\theta+1}+\frac{1}{2\cosh\theta-1}\right). \tag{19.10.1}$$

The bulk energy term is then

$$\mathcal{E}_0 \;=\; \frac{m^2}{2\sin\frac{2\pi}{3}}.$$

To discuss the conformal limit of this scattering theory we need first to find the plateau value of the pseudo-energy, the solution of the transcendental equation

$$\epsilon_0 \;=\; \log\left(1+e^{-\epsilon_0}\right). \tag{19.10.2}$$

Taking the exponential of both terms and imposing $x = e^{\epsilon_0}$, it reduces to the algebraic equation

$$x^2 - x - 1 \;=\; 0 \qquad \epsilon_0 \;=\; \log\left(\frac{\sqrt{5}+1}{2}\right)$$

and, because

$$L\left(\frac{2}{3+\sqrt{5}}\right) \;=\; \frac{\pi^2}{15}$$

for the effective central charge we get the value

$$\tilde{c}(0) \;=\; \frac{2}{5}. \tag{19.10.3}$$

Notice that for the Yang–Lee conformal model, $c = -22/5$ while $\Delta_{min} = -1/5$. The effective central charge is then $c_{eff} = c - 24\Delta_{min} = 2/5$, in agreement with the value above. The exact relation between the coupling constant λ of the field that perturbs the conformal theory is given by $i\lambda = \mathcal{D}\, m^{12/5}$ with

$$\mathcal{D} \;=\; -\frac{25}{12}\left(\frac{\pi}{12}\right)^{1/5}\left(\frac{\Gamma\left(\frac{5}{6}\right)}{\Gamma\left(\frac{1}{3}\right)}\right)^{12/5}\left(\frac{\Gamma\left(\frac{2}{5}\right)\Gamma\left(\frac{6}{5}\right)}{\Gamma\left(-\frac{1}{5}\right)\Gamma\left(\frac{3}{5}\right)}\right)^{1/2} \;=\; 0.09704845\ldots\, i. \tag{19.10.4}$$

Note the explicit presence of the imaginary number i by the non-unitarity of the model.

19.10.2 The Ising Model in a Magnetic Field

The S-matrix proposed for the Ising model in an external magnetic field involves eight particles of different masses, and the S-matrix amplitude of the lowest particle is (with the notation of Chapter 17)

$$S_{11}(\theta) = f_{2/3}(\theta)\, f_{2/5}(\theta)\, f_{1/15}(\theta).$$

This amplitude determines the bulk energy term

$$\mathcal{E}_0 = \frac{m_1^2}{2(\sin\frac{2\pi}{3} + \sin\frac{2\pi}{5} + \sin\frac{\pi}{15})}.$$

From the exact expressions of all other amplitudes, given in Chapter 18, we can determine the N matrix

$$N_{ab} = \begin{pmatrix} 3 & 4 & 6 & 6 & 8 & 8 & 10 & 12 \\ 4 & 7 & 7 & 10 & 12 & 14 & 16 & 20 \\ 6 & 8 & 11 & 12 & 16 & 16 & 20 & 24 \\ 6 & 10 & 12 & 15 & 18 & 20 & 24 & 30 \\ 8 & 12 & 16 & 18 & 23 & 24 & 30 & 36 \\ 8 & 14 & 16 & 20 & 24 & 27 & 32 & 40 \\ 10 & 16 & 20 & 24 & 30 & 32 & 39 & 48 \\ 12 & 20 & 24 & 30 & 36 & 40 & 48 & 59 \end{pmatrix}.$$

This permits us to derive the plateu values of the pseudo-energy solving eqn (19.6.4):

$$\begin{array}{lll} e^{\epsilon_1} = 2 + 2\sqrt{2} & e^{\epsilon_2} = 5 + 4\sqrt{2} & e^{\epsilon_3} = 11 + 8\sqrt{2} \\ e^{\epsilon_4} = 16 + 12\sqrt{2} & e^{\epsilon_5} = 42 + 30\sqrt{2} & e^{\epsilon_6} = 56 + 40\sqrt{2} \\ e^{\epsilon_7} = 152 + 108\sqrt{2} & e^{\epsilon_8} = 543 + 384\sqrt{2}. & \end{array} \tag{19.10.5}$$

Computing the dilogarithmic functions associated to these values, we have

$$\begin{array}{lll} \tilde{c}_1 = 0.2100096.. & \tilde{c}_2 = 0.120269.. & \tilde{c}_3 = 0.068324.. \\ \tilde{c}_4 = 0.0500483.. & \tilde{c}_5 = 0.023056.. & \tilde{c}_6 = 0.018087.. \\ \tilde{c}_7 = 0.0076889.. & \tilde{c}_8 = 0.002515.. & \end{array} \tag{19.10.6}$$

whose sum is $\tilde{c}(0) = \frac{1}{2}$. This is a highly non-trivial check of the validity of the S-matrix proposed for the magnetic deformation of the Ising model. The exact relation that links the lowest mass m_1 of this model to the coupling constant, given in this case by the magnetic field, is $m_1 = \mathcal{C}\, h^{\frac{8}{15}}$ where

$$\mathcal{C} = \frac{4\sin\frac{\pi}{5}\Gamma\left(\frac{1}{5}\right)}{\Gamma\left(\frac{2}{3}\right)\Gamma\left(\frac{8}{15}\right)}\left(\frac{4\pi^2\Gamma\left(\frac{3}{4}\right)\Gamma^2\left(\frac{13}{16}\right)}{\Gamma\left(\frac{1}{4}\right)\Gamma^2\left(\frac{3}{16}\right)}\right)^{\frac{4}{5}} = 4.40490858\ldots \tag{19.10.7}$$

19.10.3 The Tricritical Ising Model

The thermal deformation of the tricritical Ising model is described by an exact S-matrix based on seven particles, where the amplitude of the fundamental particle is

$$S_{11} = -f_{1/9}(\theta)\, f_{4/9}(\theta).$$

Hence, the bulk energy term is

$$\mathcal{E}_0 = \frac{m_1^2}{2(\sin\frac{\pi}{9}\,\sin\frac{4\pi}{9})}.$$

From the other scattering amplitudes we can obtain the N matrix of this model:

$$N_{ab} = \begin{pmatrix} 2 & 2 & 3 & 4 & 4 & 5 & 6 \\ 2 & 3 & 4 & 4 & 6 & 6 & 8 \\ 3 & 4 & 6 & 6 & 8 & 9 & 12 \\ 4 & 4 & 6 & 7 & 8 & 10 & 12 \\ 4 & 6 & 8 & 8 & 11 & 12 & 16 \\ 5 & 6 & 9 & 10 & 12 & 14 & 18 \\ 6 & 8 & 12 & 12 & 16 & 18 & 23 \end{pmatrix}.$$

The solutions of the plateau equations of the pseudo-energies are

$$\begin{array}{lll} e^{\epsilon_1} = 2+\sqrt{5} & e^{\epsilon_2} = (5+3\sqrt{5})/2 & e^{\epsilon_3} = 6+3\sqrt{5} \\ e^{\epsilon_4} = 8+4\sqrt{5} & e^{\epsilon_5} = (33+15\sqrt{15})/2 & e^{\epsilon_6} = 27+12\sqrt{5} \\ e^{\epsilon_7} = 80+36\sqrt{5}. & & \end{array} \tag{19.10.8}$$

Computing the dilogarithmic functions at these values, we get

$$\begin{array}{lll} \tilde{c}_1 = 0.228828.. & \tilde{c}_2 = 0.184429.. & \tilde{c}_3 = 0.1054611.. \\ \tilde{c}_4 = 0.084686.. & \tilde{c}_5 = 0.049684.. & \tilde{c}_6 = 0.0335404.. \\ \tilde{c}_7 = 0.013369.. & & \end{array} \tag{19.10.9}$$

whose sum gives $\tilde{c}(0) = 7/10$, which is the central charge of the tricritical Ising model. This provides explicit confirmation of the validity of the S-matrix for the thermal deformation of this model. The exact relation between the lowest mass m_1 and the coupling constant $\tau = T - T_c$ is $m_1 = \mathcal{C}\,\tau^{\frac{5}{9}}$ where

$$\mathcal{C} = \left(\frac{2\Gamma\left(\frac{2}{9}\right)}{\Gamma\left(\frac{2}{3}\right)\,\Gamma\left(\frac{5}{9}\right)}\right)\left(\frac{4\pi^2\,\Gamma\left(\frac{2}{5}\right)\,\Gamma^3\left(\frac{4}{5}\right)}{\Gamma^3\left(\frac{1}{5}\right)\,\Gamma\left(\frac{3}{5}\right)}\right)^{5/18} = 3.745378362\ldots \tag{19.10.10}$$

19.11 Thermodynamics of the Free Field Theories

A particularly simple case of the TBA equations is associated to the theories where there is only one massive excitation with a constant S-matrix, that is $S = \pm 1$. In these theories it is obviously not necessary to solve the integral equations (19.4.15) to derive the thermodynamics since we have identically

$$\epsilon(\theta) = r\,\cosh\theta, \tag{19.11.1}$$

and, for the central charge,

$$c_{\pm}(r) = \mp\frac{6}{\pi^2}\int_0^\infty d\theta\, r\,\cosh\theta\,\log(1 \mp e^{-r\cosh\theta}). \tag{19.11.2}$$

Apart from the prefactor $-\pi/6R^2$, these expressions are precisely the free energies of a relativistic gas with Bose/Fermi statistics at temperature $T = 1/R$ (see Appendix B of

Chapter 1). Let's discuss their analytic structure. Expanding the logarithm in powers of $\exp[-r\cosh\theta]$ and integrating term by term, we get

$$c_{\pm}(r) \;=\; \frac{6r}{\pi^2}\sum_{n=1}^{\infty}\frac{(\pm 1)^{n-1}}{k}\,K_1(nr), \tag{19.11.3}$$

where $K_1(z)$ is the modified Bessel function. Taking now the limit $r \to 0$, we have

$$c_{\pm}(0) \;=\; \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{(\pm 1)^{n-1}}{n^2} \;=\; \begin{cases} 1 \\ \frac{1}{2}. \end{cases} \tag{19.11.4}$$

Moreover, using

$$\frac{d}{dx}[xK_1(x)] \;=\; -x\,K_0(x),$$

we obtain

$$\frac{1}{r}\frac{d}{dr}c_{\pm}(r) \;=\; -\frac{6}{\pi^2}\sum_{n=1}^{\infty}(\pm 1)^{n-1}\,K_0(nr). \tag{19.11.5}$$

Using the identity

$$\begin{aligned}\sum_{n=1}K_0(nx)\,\cos nxt \;=\; &\frac{1}{2}\left(\gamma_E+\log\frac{x}{4\pi}\right)+\frac{\pi}{2x\sqrt{1+t^2}} \\ &+\frac{\pi}{2}\sum_{l=1}^{\infty}\left\{\frac{1}{\sqrt{x^2+(2l\pi-tx)^2}}-\frac{1}{2l\pi}\right\} \\ &+\frac{\pi}{2}\sum_{l=1}^{\infty}\left\{\frac{1}{\sqrt{x^2+(2l\pi+tx)^2}}-\frac{1}{2l\pi}\right\}\end{aligned}$$

and integrating eqn (19.11.5), we arrive at the expressions:

for the bosonic case $S=1$

$$\begin{aligned}c_+(r) = 1-\frac{3r}{\pi}+\frac{3r^2}{2\pi^2}\left[\log\frac{1}{r}+\frac{1}{2}+\log 4\pi-\gamma_E\right] \\ -\frac{6}{\pi}\sum_{n=1}^{\infty}\left(\sqrt{(2n\pi)^2+r^2}-2n\pi-\frac{r^2}{4n\pi}\right)\end{aligned} \tag{19.11.6}$$

while, for the fermionic case $S=-1$

$$\begin{aligned}c_-(r) = \frac{1}{2}-\frac{3r^2}{2\pi^2}\left[\log\frac{1}{r}+\frac{1}{2}+\log\pi-\gamma_E\right] \\ +\frac{6}{\pi}\sum_{n=1}^{\infty}\left(\sqrt{(2n-1)^2\pi^2+r^2}-(2n-1)\pi-\frac{r^2}{2(2n-1)\pi}\right).\end{aligned} \tag{19.11.7}$$

The plots of these functions are given in Fig. 19.4.

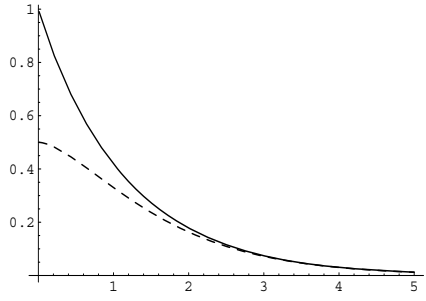


Fig. 19.4 *Plot of the functions $c_+(r)$ (continuous line) and $c_-(r)$ (dashed line).*

19.12 L-channel Quantization

The formulas obtained by the TBA for the finite volume vacuum energy of free theories can be directly derived by quantizing them in the L-channel. In this section we present explicit formulas for the bosonic case, since similar expressions can be easily reproduced for the fermionic case. Let $\phi(x,t) = \phi^\dagger(x,t)$ be the real bosonic field defined in the interval $(-\frac{R}{2}, \frac{R}{2})$, with periodic boundary conditions

$$\phi(x+R,t) \;=\; \phi(x,t) \tag{19.12.1}$$

at any time. The action is

$$\mathcal{A} \;=\; \int dt \int_{-\frac{R}{2}}^{\frac{R}{2}} dx\, \frac{1}{2}\left[(\partial_\mu \phi)^2 - m^2\phi^2\right].$$

Defining the conjugate momentum of the field

$$\Pi(x,t) \;=\; \frac{\partial \phi}{\partial t}(x,t),$$

for the hamiltonian of the system we have

$$H \;=\; \frac{1}{2}\int_{-\frac{R}{2}}^{\frac{R}{2}} dx[\Pi^2 + (\nabla\phi)^2 + m^2\phi^2]. \tag{19.12.2}$$

Let's assume the commutation relations

$$\begin{aligned}[\phi(x,t), \Pi(y,t)] \;&=\; i\,\delta_p(x-y),\\ [\phi(x,t), \phi(y,t)] \;&=\; [\Pi(x,t), \Pi(y,t)] \;=\; 0,\end{aligned} \tag{19.12.3}$$

where $\delta_p(x-y)$ is the periodic version of the usual Dirac delta function: in addition to the usual properties, in this case it also satisfies

$$\delta_p(x+R) \;=\; \delta_p(x).$$

Its explicit representation is given by

$$\delta_p(x) \;=\; \frac{1}{L}\sum_{-\infty}^{\infty}\exp\left[\frac{2\pi i n}{R}x\right]. \tag{19.12.4}$$

It is now necessary to solve the equation of motion of the field $\phi(x,t)$

$$\left[\frac{\partial^2}{\partial t^2}-\nabla^2+m^2\right]\,\phi(x,t) \;=\; 0, \tag{19.12.5}$$

together with the boundary conditions (19.12.1). There is a standard procedure to do so: because the field is periodic along the space direction, it admits a Fourier expansion

$$\phi(x,t) \;=\; \sum_{-\infty}^{\infty} c_n(t)\exp\left[\frac{2ni}{R}x\right]. \tag{19.12.6}$$

It is convenient to introduce the notation

$$p_n \;=\; \frac{2\pi n}{R},\quad \omega_n \;=\; \sqrt{p_n^2+m^2},\quad n=0,\pm1,\pm2\,\ldots$$

Substituting the expansion (19.12.6) in the equation of motion (19.12.5), we obtain

$$\left[\frac{d^2}{dt^2}+\omega_n^2\right]\,c_n(t) \;=\; 0,$$

whose solution is

$$c_n(t) \;=\; a_n\,e^{-i\omega_n t}+a_n^\dagger\,e^{i\omega_n t}.$$

Hence, the field and its conjugate momentum are expressed by

$$\phi(x,t) \;=\; \sum_{-\infty}^{+\infty} N_n\,\left[a_n\,e^{i(p_n x-\omega_n t)}+a_n^\dagger\,e^{-i(p_n x-\omega_n t)}\right], \tag{19.12.7}$$

$$\Pi(x,t) \;=\; \sum_{-\infty}^{+\infty} N_n\,(-i\omega_n)\,\left[a_n\,e^{i(p_n x-\omega_n t)}-a_n^\dagger\,e^{-i(p_n x-\omega_n t)}\right],$$

where N_n is a normalization that can be fixed by imposing the quantization conditions (19.12.3). Choosing

$$N_n \;=\; \frac{1}{\sqrt{2}\omega_n L},$$

eqn (19.12.3) becomes the commutation relation among the a_n modes

$$\begin{aligned}\left[a_n,a_m^\dagger\right] \;&=\; \delta_{n,m},\\ \left[a_n,a_m\right] \;&=\; \left[a_n^\dagger,a_m^\dagger\right] \;=\; 0.\end{aligned} \tag{19.12.8}$$

Substituting the expressions of $\phi(x,t)$ and $\Pi(x,t)$ in the hamiltonian (19.12.2), we have

$$H \,=\, \frac{1}{2}\sum_{-\infty}^{+\infty}\omega_n(a_n a_n^\dagger + a_n^\dagger a_n) \,=\, \frac{1}{2}\sum_{-\infty}^{+\infty}\omega_n\left[a_n^\dagger a_n + \frac{1}{2}\right], \tag{19.12.9}$$

where we used

$$\int_{-\frac{R}{2}}^{\frac{R}{2}} e^{i(p_n-p_m)x}\,dx \,=\, R\,\delta_{n,m}.$$

The ground state energy of the theory is then

$$E_0(R) \,=\, \frac{1}{2}\sum_{-\infty}^{+\infty}\omega_n \,=\, \frac{1}{2}\sum_{-\infty}^{+\infty}\sqrt{\left(\frac{2n\pi}{R}\right)^2 + m^2}. \tag{19.12.10}$$

This expression needs, however, to be regularized by subtracting the term coming from the continuous limit of the infinite volume in order to implement the correct normalization

$$\lim_{R\to\infty}\mathcal{E}_0^{\rm vac}(R) \,=\, 0.$$

Hence, for the finite volume ground state energy we have

$$\mathcal{E}_0^{\rm vac}(R) \,=\, \frac{1}{2}\sum_{n=-\infty}^{\infty}\sqrt{\left(\frac{2\pi n}{R}\right)^2 + m^2} \,-\, \frac{1}{2}\int_{-\infty}^{\infty} dn\,\sqrt{\left(\frac{2\pi n}{R}\right)^2 + m^2}. \tag{19.12.11}$$

Selecting out the zero mode, this expression can be written as

$$\mathcal{E}_0^{\rm vac}(R) \,=\, \frac{m}{2} + \frac{2\pi}{R}\sum_{n=1}^{\infty}\sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} \,-\, \frac{2\pi}{R}\int_0^{\infty} dn\,\sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2}, \tag{19.12.12}$$

where $r \equiv mR$. Since the divergence of the series is due to the large n behavior of the first two terms of the expansion

$$\sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} \,\simeq\, n + \frac{1}{2}\left(\frac{r}{2\pi}\right)^2\frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right),$$

let's start by subtracting and adding these divergent terms

$$\begin{aligned} S(r) \,\equiv\, \sum_{n=1}^{\infty}\sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} &= \sum_{n=1}^{\infty}\left\{\sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} - n - \frac{1}{2}\left(\frac{r}{2\pi}\right)^2\frac{1}{n}\right\} \\ &\quad + \sum_{n=1}^{\infty} n \,+\, \frac{1}{2}\left(\frac{r}{2\pi}\right)^2\sum_{n=1}^{\infty}\frac{1}{n}. \end{aligned} \tag{19.12.13}$$

The first series on the right is now convergent, while the last two terms must be paired with analogous terms coming from the integral, whose divergence has to be treated

in a similar way to the series. Hence, subtracting and adding these divergent terms in the integral

$$I(r) \equiv \int_0^\infty dn \sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} = \\ = \int_0^\infty dn \left\{ \sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} - n \right\} + \int_0^\infty dn\, n, \tag{19.12.14}$$

we can pair the last term of this expression with the one in (19.12.13) and implement the well-known regularization

$$\sum_{n=0}^\infty n - \int_0^\infty n\, dn = \lim_{\alpha\to 0} \left[\sum_{n=0}^\infty n\, e^{-\alpha n} - \int_0^\infty n\, e^{-\alpha n}\, dn \right] = -\frac{1}{12}. \tag{19.12.15}$$

However, the first term in (19.12.14) still contains a logarithmic divergence, as can be seen by explicitly computing the integral by means of a *cut-off* Λ, in the limit $\Lambda \to \infty$

$$\int_0^\Lambda dn \left\{ \sqrt{n^2 + \left(\frac{r}{2\pi}\right)^2} - n \right\} = \frac{1}{2}\left(\frac{r}{2\pi}\right)^2 \ln 2\Lambda + \frac{1}{4}\left(\frac{r}{2\pi}\right)^2 - \frac{1}{2}\left(\frac{r}{2\pi}\right)^2 \ln \frac{r}{2\pi}. \tag{19.12.16}$$

This divergence can be cured by subtracting and adding the term $\frac{1}{2}\left(\frac{r}{2\pi}\right)^2 \ln \Lambda$. Combining this last term with the analogous one coming from the series, we obtain

$$\lim_{\Lambda\to\infty} \left(\sum_{n=1}^\Lambda \frac{1}{n} - \ln \Lambda \right) = \gamma_E,$$

where γ_E is the Euler–Mascheroni constant, whereas the remaining part of (19.12.16), with the subtraction done above, is now finite.

Gathering together all the terms, the finite expression of the finite volume ground state energy is

$$\mathcal{E}_0^{\rm vac}(R) = \frac{1}{R}\left[-\frac{\pi}{6} + \frac{r}{2} + \frac{r^2}{4\pi}\left(\ln \frac{r}{4\pi} + \gamma_E - \frac{1}{2} \right) + \sum_{n=1}^\infty \left(\sqrt{(2\pi n)^2 + r^2} - 2\pi n - \frac{r^2}{4\pi n} \right) \right]. \tag{19.12.17}$$

It is easy to see that eqn (19.12.17) satisfies modular invariance, which imposes its equality with the expression obtained by the TBA

$$\mathcal{E}_0^{\rm vac}(R) = -\frac{\pi c(r)}{6R},$$

where

$$c(r) = -\frac{6r}{\pi^2} \int_0^\infty d\theta\, \cosh\theta\, \ln\left(1 - e^{-r\cosh\theta}\right).$$

Moreover, it is easy to verify that the regularization adopted above ensures perfect agreement between the expressions for the one-point correlation functions $\langle \phi^{2k} \rangle$, done either in the R- or the L-channels.

It is useful to notice that the result (19.12.17) can be obtained in the simplest way by using a prescription that automatically ensures the subtraction of the various divergent terms. This consists of ignoring completely the divergent part of the integral, though keeping its finite part, and regularizing the divergent series according to the formulas

$$\sum_{n=1}^{\infty} n \Bigg|_{\text{reg}} = -\frac{1}{12}, \tag{19.12.18}$$

$$\sum_{n=1}^{\infty} \frac{1}{n} \Bigg|_{\text{reg}} = \gamma_E + \ln \frac{r}{2\pi}. \tag{19.12.19}$$

Equation (19.12.18) is the standard regularization provided by the Riemann zeta function $\zeta(-1)$, where $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$. This regularization corresponds to the normal order of the operators in the infinite volume. However, from the logarithmic divergence, the regularization of the second series is intrinsically ambiguous, and its finite value can be determined according to the earlier discussion.

References and Further Reading

The use of the S-matrix to determine the thermodynamics of a quantum field theory has been proposed in:

R. Dashen, S.K. Ma, H.J. Bernstein, *S-matrix formulation of statistical mechanics*, Phys. Rev. 187 (1969), 345.

The original paper on the thermodynamics Bethe ansatz for the integral models is:

C.N. Yang, C.P. Yang, *Thermodynamics of a one-dimensional system of bosons with repulsive delta-function interaction*, J. Math. Phys. 10 (1969), 1115.

The generalization to the relativistic case has been proposed by Al. Zamolodchikov:

Al.B. Zamolodchikov, *Thermodynamic Bethe ansatz in relativistic models: Scaling 3-state Potts and Lee–Yang models*, Nucl. Phys. B 342 (1990), 695.

On the relation between the TBA and the deformed conformal theories it is useful to read the articles:

T. Klassen, E. Melzer, *Purely elastic scattering theories and their ultraviolet limits*, Nucl. Phys. B 338 (1990), 485.

T. Klassen, E. Melzer, *The thermodynamics of purely elastic scattering theories and conformal perturbation theory*, Nucl. Phys. B 350 (1991), 635.

The relationship between the coupling constant and the lowest mass of the integrable models has been the subject of these papers:

Al.B. Zamolodchikov, *Mass scale in the Sine–Gordon model and its reductions*, Int. J. Mod. Phys. A 10 (1995), 1125.

V. Fateev, *The exact relations between the coupling constants and the masses of particles for the integrable perturbed conformal field theories*, Phys. Lett. B 324 (1994), 45.

The formalism based on the TBA to compute, in addition to the free energy, also the finite temperature correlation functions in integrable models has been proposed in:

A. LeClair, G. Mussardo, *Finite temperature correlation functions in integrable QFT*, Nucl. Phys. B 552 (1999), 624.

The TBA is deeply related to the monodromy matrix of integrable models. This subject has been investigated in a remarkable series of papers:

V. Z. Bazhanov, S. Lukyanov, A.B. Zamolodchikov, *Integrable structure of conformal field theory, quantum KdV theory and thermodynamic Bethe ansatz*, Comm. Math. Phys. 177 (1996), 381.

V. Z. Bazhanov, S. Lukyanov, A.B. Zamolodchikov, *Integrable structure of conformal field theory. 2 Q operator and DDV equation*, Comm. Math. Phys. 190 (1997), 247.

V. Z. Bazhanov, S. Lukyanov, A.B. Zamolodchikov, *Integrable structure of conformal field theory. 3. The Yang–Baxter relation*, Comm. Math. Phys. 200 (1999), 297.

The TBA can also be extended to compute the finite volume energies of the excited states, as shown in the papers:

P. Fendley, *Excited state thermodynamics*, Nucl. Phys. B 374 (1992), 667.

V. Z. Bazhanov, S. Lukyanov, A.B. Zamolodchikov, *Integrable quantum field theories in finite volume: Excited state energies*, Nucl. Phys. B 489 (1997), 487.

P. Dorey and R. Tateo, *Excited states by analytic continuation of TBA equations*, Nucl. Phys. B 482 (1996), 639.

The interesting relationship between the TBA and the spectrum of one-dimensional quantum mechanical systems has been pointed out in the papers:

P. Dorey and R. Tateo, *Anharmonic oscillators, the thermodynamic Bethe ansatz, and nonlinear integral equations*, J. Phys. A 32 (1999), L419.

V. Z. Bazhanov, S. Lukyanov, A.B. Zamolodchikov, *Spectral determinants for the Schrödinger equation and Q operators of conformal field theory*, J. Stat. Phys. 102 (2001), 567.

The TBA equation has also been applied to study the cascade of RG flows $\mathcal{M}_p \to \mathcal{M}_{p-1}$ between the minimal models of CFF. See:

Al. B. Zamolodchikov, *Resonance factorized scattering and roaming trajectories*, J. Phys. A 39 (2006), 12847.

Problems

1. Non-relativistic gas

Consider a one-dimensional gas of N non-relativistic bosons on an interval of length L, with two-body repulsive interaction given by a delta function. The hamiltonian of such a system is

$$H \;=\; -\sum_{i=1}^{N} \frac{\partial^2}{\partial x_i^2} + 2c \sum_{i>j} \delta(x_i - x_j) \qquad c > 0.$$

a Find the phase shift of the two-body scattering process and write the Bethe ansatz equations.

b Analyze the solutions in the thermodynamic limit $N \to \infty$, $L \to \infty$, $N/L = \rho$, ρ finite.

2. Simple TBA system

Consider the TBA equations for a relativistic system made of one massive particle, with kernel

$$\varphi(\theta) \;=\; \frac{1}{2\pi}\delta(\theta).$$

a Solve explicitly the equation for the pseudo-energy $\epsilon(\theta)$ and show that it is given by

$$\epsilon(\theta) \;=\; \log\left[e^{mR\cosh\theta} - 1\right].$$

b Plot the scaling function

$$c(R) \;=\; \frac{6}{\pi^2} mR \int_0^\infty \cosh\theta \, \log(1 + e^{-\epsilon(\theta)})\, d\theta$$

and compute its limit at $R = 0$

3. L-channel for Majorana fermions

Consider the Dirac action of a Majorana massive fermion on a finite volume

$$\mathcal{S} \;=\; \int dt \int_{-\frac{R}{2}}^{\frac{R}{2}} dx\, \bar{\psi}\,(i\,\gamma^\mu\,\partial_\mu - m)\;\psi.$$

Quantize this system in the canonical way and show that the finite volume ground state energy $E_0(R)$ can be written as

$$E_0(R) \;=\; -\frac{\pi\, c_-(r)}{6R},$$

where the scaling function $c_-(r)$ coincides with the expression given by the TBA.

20
Form Factors and Correlation Functions

Elementary, my dear Watson.

Arthur Conan Doyle

One of the fundamental problems of statistical mechanics and its quantum field theory formulation is the characterization of the order parameters and the computation of their correlation functions. Besides the intrinsic interest of this problem, the correlation functions are the key quantities in the determination of the universal ratios of the renormalization group and therefore they can have direct experimental confirmation (see Chapter 8). In general, the computation of correlation functions is a difficult task, usually achieved with partial success through perturbative methods.

As we saw in the previous chapters devoted to conformal field theories, an exact determination of the operator content and the correlation functions of a two-dimensional theory can be obtained only when the model is at its critical point. In this case, in fact, one has a classification of the order parameters in terms of the irreducible representation of the Virasoro algebra and, moreover, one can get an exact expression of the correlators by solving the linear differential equations that they satisfy.

Unfortunately, the simple theoretical scheme of the critical points cannot be generalized once we move away from criticality. In this case, the problem has to be faced with different techniques. As shown in this chapter, significant progress can be made when we deal with integrable theories, characterized by their elastic S-matrix and the spectrum of the asymptotic states. The central quantities are in this case the matrix elements of the various operators on the asymptotic states of the theory, called the *form factors*. The precise definition of these quantities is given below. The general properties related to the unitarity and crossing symmetry lead to a set of functional equations for the form factors that can be explicitly solved in many interesting cases. Once the matrix elements of the operators are known, their correlation functions can be recovered in terms of spectral representation series. It is worth mentioning that these series present remarkable convergence properties.

Hence, the success of the form factor method relies on two points: (a) the possibility of determining exactly the matrix elements of the order parameters on the asymptotic states of the theory, identified by scattering theory; (b) the fast convergence properties of the spectral series. These two steps lead to the determination of the correlation functions away from criticality with a precision that cannot be obtained by other methods.

20.1 General Properties of the Form Factors

An essential quantity for the computation of the matrix elements is the S-matrix of the problem. As shown in the previous chapters, the S-matrix of many two-dimensional systems is particularly simple and can be explicitly found. For an infinite number of conservation laws, the scattering processes of integrable systems are purely elastic and the n-particle S-matrix can be factorized in terms of the $n(n-1)/2$ two-body scattering amplitudes. In the following, for simplicity, we mainly focus our attention on diagonal scattering theories with a non-degenerate spectrum. To characterize the kinematic state of the particles we use the rapidities θ_i that enter the dispersion relations

$$p_i^0 = m_i \cosh\theta_i, \quad p_i^1 = m_i \sinh\theta_i. \tag{20.1.1}$$

The two-body S-matrix amplitudes depend on the difference of the rapidities $\theta_{ij} = \theta_i - \theta_j$ and satisfy the unitary and crossing symmetry equations

$$\begin{aligned} S_{ij}(\theta_{ij}) &= S_{ji}(\theta_{ij}) = S_{ij}^{-1}(-\theta_{ij}), \\ S_{i\bar{j}}(\theta_{ij}) &= S_{ij}(i\pi - \theta_{ij}). \end{aligned} \tag{20.1.2}$$

Possible bound states correspond to simple poles (or higher order odd poles) of these amplitudes, placed at imaginary values of θ_{ij} in the physical strip $0 < \mathrm{Im}\theta < \pi$.

Let's see how the S-matrix allows us to compute the matrix elements of the (semi)-local operators on the asymptotic states. To this end, it is useful to introduce an algebraic formalism.

20.1.1 Faddeev–Zamolodchikov Algebra

A key assumption of the form factor theory is that there exist some operators, both of creation and annihilation type, $V^\dagger_{\alpha_i}(\theta_i)$, $V_{\alpha_i}(\theta_i)$, that implement a generalization of the usual bosonic and fermionic algebraic relations. Let's call them *vertex operators*. Denoting by α_i the quantum number that distinguishes the different types of particles of the theory, these operators satisfy the associative algebra in which enters the S-matrix

$$V_{\alpha_i}(\theta_i)V_{\alpha_j}(\theta_j) = S_{ij}(\theta_{ij})\, V_{\alpha_j}(\theta_j)V_{\alpha_i}(\theta_i) \tag{20.1.3}$$

$$V^\dagger_{\alpha_i}(\theta_i)V^\dagger_{\alpha_j}(\theta_j) = S_{ij}(\theta_{ij})\, V^\dagger_{\alpha_j}(\theta_j)V^\dagger_{\alpha_i}(\theta_i) \tag{20.1.4}$$

$$V_{\alpha_i}(\theta_i)V^\dagger_{\alpha_j}(\theta_j) = S_{ij}(\theta_{ji})\, V^\dagger_{\alpha_j}(\theta_j)V_{\alpha_i}(\theta_i) + 2\pi\delta_{\alpha_i\alpha_j}\delta(\theta_{ij}). \tag{20.1.5}$$

Any commutation of these operators can be interpreted as a scattering process. The Poincaré group, generated by the Lorentz transformations $L(\epsilon)$ and the translations T_y, acts on the operators as

$$U_L V_\alpha(\theta) U_L^{-1} = V_\alpha(\theta + \epsilon) \tag{20.1.6}$$

$$U_{T_y} V_\alpha(\theta) U_{T_y}^{-1} = e^{ip_\mu(\theta)y^\mu} V_\alpha(\theta). \tag{20.1.7}$$

Obviously the explicit form of the creation and annihilation operators depends crucially on the theory in question and their construction is an open problem for most

models. This difficulty does not stop us, however, from deriving the fundamental equations for the matrix elements starting from the algebraic equations given above.

The vertex operators define the space of physical states. The vacuum $|0\rangle$ is the state annihilated by $V_\alpha(\theta)$,

$$V_\alpha(\theta)|0\rangle = 0 = \langle 0|V_\alpha^\dagger(\theta),$$

while the Hilbert space is constructed by applying the various vertex operators $V_\alpha^\dagger(\theta)$ on $|0\rangle$:

$$|V_{\alpha_1}(\theta_1)\dots V_{\alpha_n}(\theta_n)\rangle \equiv V_{\alpha_1}^\dagger(\theta_1)\dots V_{\alpha_n}^\dagger(\theta_n)|0\rangle. \tag{20.1.8}$$

From eqn (20.1.5), the one-particle states have the normalization

$$\langle V_{\alpha_i}(\theta_i)|V_{\alpha_j}(\theta_j)\rangle = 2\pi\,\delta_{\alpha_i\alpha_j}\delta(\theta_{ij}).$$

The algebra of the vertex operators implies that the vectors (20.1.8) are not all linearly independent. To select a basis of linearly independent vectors we need an additional requirement: for the initial states, the rapidites must be ordered in a decreasing way:

$$\theta_1 > \theta_2 > \dots > \theta_n$$

while, for the final states in an increasing way:

$$\theta_1 < \theta_2 < \dots < \theta_n.$$

These orderings select a set of linearly independent vectors that form a basis in the Hilbert space.

20.1.2 Form Factors

In this section we describe the principles of the theory. Unless explicitly stated, in the following we consider the matrix elements between the *in* and *out* states of the particle with the lowest mass of local, scalar, and hermitian operators $\mathcal{O}(x)$

$${}_{\rm out}\langle V(\theta_{m+1})\dots V(\theta_n)|\mathcal{O}(x)|V(\theta_1)\dots V(\theta_m)\rangle_{\rm in}. \tag{20.1.9}$$

We can always place the operator at the origin by using the translation operator, $U_{T_y}\mathcal{O}(x)U_{T_y}^{-1} = \mathcal{O}(x+y)$, and using eqn (20.1.7), the matrix elements above are given by

$$\begin{aligned}\exp\left[i\left(\sum_{i=m+1}^{n} p_\mu(\theta_i) - \sum_{i=1}^{m} p_\mu(\theta_i)\right)x^\mu\right] \\ \times\, {}_{\rm out}\langle V(\theta_{m+1})\dots V(\theta_n)|\mathcal{O}(0)|V(\theta_1)\dots V(\theta_m)\rangle_{\rm in}.\end{aligned} \tag{20.1.10}$$

It is convenient to define the functions

$$F_n^{\mathcal{O}}(\theta_1,\theta_2,\dots,\theta_n) = \langle 0 \mid \mathcal{O}(0) \mid \theta_1,\theta_2,\dots,\theta_n\rangle_{in} \tag{20.1.11}$$

called the *Form Factors* (FF), whose graphical representation is shown in Fig. 20.1: they are the matrix elements of an operator placed at the origin between the n-particle state and the vacuum.[1]

[1] From now on we use the simplified notation $|\dots V(\theta_n)\dots\rangle \equiv |\dots\theta_n\dots\rangle$ to denote the physical states of the particle with the lowest mass.

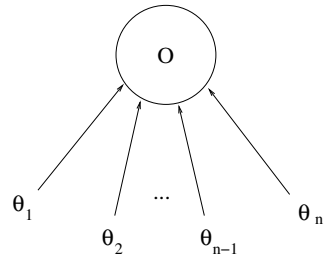


Fig. 20.1 *Form factor of the operator $\mathcal{O}$.*

For local and scalar operators, the relativistic invariance of the theory implies that the FF are functions of the differences of the rapidities θ_{ij}

$$F_n^{\mathcal{O}}(\theta_1, \theta_2, \ldots, \theta_n) = F_n^{\mathcal{O}}(\theta_{12}, \theta_{13}, \ldots, \theta_{ij}, \ldots), i < j. \tag{20.1.12}$$

The invariance under crossing symmetry permits us to recover the most general matrix elements by an analytic continuation of the functions (20.1.11)

$$F_{n+m}^{\mathcal{O}}(\theta_1, \theta_2, \ldots, \theta_m, \theta_{m+1} - i\pi, \ldots, \theta_n - i\pi) = F_{n+m}^{\mathcal{O}}(\theta_{ij}, i\pi - \theta_{sr}, \theta_{kl}) \tag{20.1.13}$$

where $1 \leq i < j \leq m$, $1 \leq r \leq m < s \leq n$, and $m < k < l \leq n$.

Apart from the poles corresponding to the bound states present in all possible channels of this amplitude, the form factors $F_n^{\mathcal{O}}$ are expected to be analytic functions in the strips $0 < \mathrm{Im}\theta_{ij} < 2\pi$.

20.2 Watson's Equations

The FF of a scalar and hermitian operator $\mathcal{O}$ satisfy a set of equations, known as *Watson's equations*, that assume a particularly simple form for the integrable systems

$$F_n^{\mathcal{O}}(\theta_1, \ldots, \theta_i, \theta_{i+1}, \ldots, \theta_n) = F_n^{\mathcal{O}}(\theta_1, \ldots, \theta_{i+1}, \theta_i, \ldots, \theta_n) S(\theta_i - \theta_{i+1}), \tag{20.2.1}$$

$$F_n^{\mathcal{O}}(\theta_1 + 2\pi i, \ldots, \theta_{n-1}, \theta_n) = e^{2\pi i \gamma} F_n^{\mathcal{O}}(\theta_2, \ldots, \theta_n, \theta_1) = \prod_{i=2}^{n} S(\theta_i - \theta_1) F_n^{\mathcal{O}}(\theta_1, \ldots, \theta_n),$$

where γ is the semilocal index of the operator $\mathcal{O}$ with respect to the operator that creates the particles. The first equation is a simple consequence of eqn (20.1.3), because a commutation of two operators is equivalent to a scattering process. Concerning the second equation, it states the nature of the discontinuity of these functions at the cuts $\theta_{1i} = 2\pi i$. The graphical representation of these equations is shown in Fig. 20.2. When $n = 2$, eqns (20.2.1) reduce to

$$\begin{aligned} F_2^{\mathcal{O}}(\theta) &= F_2^{\mathcal{O}}(-\theta)\, S_2(\theta), \\ F_2^{\mathcal{O}}(i\pi - \theta) &= F_2^{\mathcal{O}}(i\pi + \theta). \end{aligned} \tag{20.2.2}$$

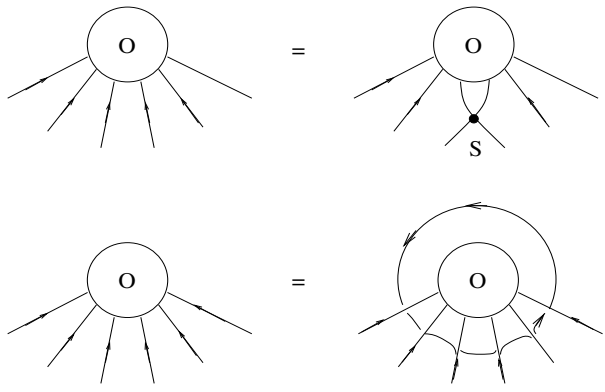


Fig. 20.2 *Graphical form of the Watson equations.*

The most general solution of the Watson equations (20.2.1) is given by

$$F_n^{\mathcal{O}}(\theta_1,\dots,\theta_n) \;=\; K_n^{\mathcal{O}}(\theta_1,\dots,\theta_n)\prod_{i<j} F_{\min}(\theta_{ij}). \tag{20.2.3}$$

Let's discuss the various terms entering this expression.

Minimal two-particle form factor. $F_{\min}(\theta)$ is an analytic function in the region $0 \le \mathrm{Im}\ \theta \le \pi$, the solution of the two equations (20.2.2), with neither zeros nor poles in the strip $0 < \mathrm{Im}\theta < \pi$, and with the mildest behavior at $|\theta| \to \infty$. These requirements determine uniquely this function, up to a normalization factor $\mathcal{N}$. Its explicit expression can be found by writing the S-matrix as

$$S(\theta) \;=\; \exp\left[\int_0^\infty \frac{dt}{t}\, f(t)\, \sinh\frac{t\theta}{i\pi}\right].$$

In fact, it is easy to see that $F_{\min}(\theta)$ is given by

$$F_{\min}(\theta) \;=\; \mathcal{N}\exp\left[\int_0^\infty \frac{dt}{t}\,\frac{f(t)}{\sinh t}\sin^2\left(\frac{t\pi\hat\theta}{2\pi}\right)\right], \qquad \hat\theta = i\pi - \theta. \tag{20.2.4}$$

Note that for interacting theories, $S(0) = -1$, and therefore the first equation in (20.2.2) forces $F_{\min}(\theta)$ to have a zero at the two-particle threshold

$$F(\theta) \;\simeq\; \theta, \qquad \theta \to 0. \tag{20.2.5}$$

$K_n^{\mathcal{O}}$ factors. The remaining factors $K_n^{\mathcal{O}}$ in (20.2.3) satisfy the Watson equation but with $S_2 = 1$: this implies that they are completely symmetric functions in the variables θ_{ij}, periodic with period $2\pi i$. Therefore they can be considered as functions of the variables $\cosh\theta_{ij}$. Let's investigate other properties of the functions $K_n^{\mathcal{O}}$. They must have all physical poles expected for the form factors. We recall that, in general, there

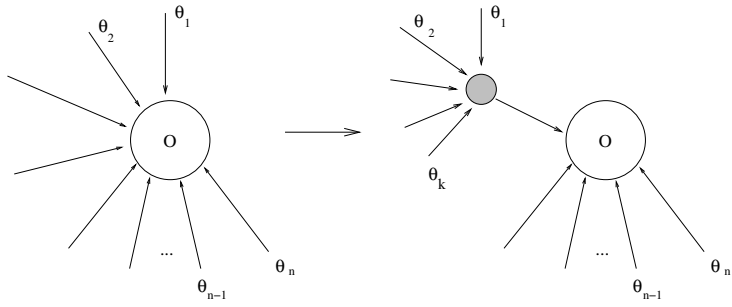


Fig. 20.3 *Kinematic configuration of a k-particle responsable for a pole in the form factors.*

is a simple pole in the form factors when a cluster made of k particles can reach a kinematical configuration that is equivalent to that of a single particle, as shown in Fig. 20.3, with the pole given just by the propagator of the latter particle. If this is the general situation, for the integrable theories there is however an important simplification. In fact, by the factorization property of the S-matrix, it is sufficient to consider only the cases in which the clusters are made of $k = 2$ or $k = 3$: the poles coming from the two-particle clusters are dictated uniquely by the bound states of the S-matrix, while those coming from the three-particle clusters are determined by the crossing processes, although they are also related to the S-matrix (see the discussion in the next section). In conclusion, all the poles of the form factors are determined by the underlying scattering theory and they do not depend on the operator! In the light of this analysis, the functions $K_n^{\mathcal{O}}$ can be parameterized as follows

$$K_n^{\mathcal{O}}(\theta_1, \dots, \theta_n) \;=\; \frac{Q_n^{\mathcal{O}}(\theta_1, \dots, \theta_n)}{D_n(\theta_1, \dots, \theta_n)}, \tag{20.2.6}$$

where the denominator D_n is a polynomial in $\cosh\theta_{ij}$ that is fixed only by the pole structure of the S-matrix while the information on the operator $\mathcal{O}$ is enclosed in the polynomial $Q_n^{\mathcal{O}}$ of the variables $\cosh\theta_{ij}$ placed at the *numerator*. We will come back to this important point in later sections.

Symmetric polynomials. As shown above, the functions $K_n^{\mathcal{O}}$ are symmetric under the permutation of the rapidities of the various particles. In many case it is convenient to change variables, introducing the parameters $x_i \equiv e^{\theta_i}$, so that both numerator and denominator become symmetric polynomials in the x_i variables. A basis in the functional space of the symmetric polynomials in n variables is given by the *elementary symmetric polynomials* $\sigma_k^{(n)}(x_1, \dots, x_n)$, whose generating function is

$$\prod_{i=1}^{n}(x + x_i) \;=\; \sum_{k=0}^{n} x^{n-k}\, \sigma_k^{(n)}(x_1, x_2, \dots, x_n). \tag{20.2.7}$$

Conventionally all $\sigma_k^{(n)}$ with $k > n$ and with $n < 0$ are zero. The explicit expressions for the other cases are

$$\begin{aligned}
&\sigma_0 = 1,\\
&\sigma_1 = x_1 + x_2 + \ldots + x_n,\\
&\sigma_2 = x_1x_2 + x_1x_3 + \ldots x_{n-1}x_n,\\
&\vdots \qquad\qquad \vdots\\
&\sigma_n = x_1x_2\ldots x_n.
\end{aligned} \tag{20.2.8}$$

The $\sigma_k^{(n)}$ are homogeneous polynomials in x_i, of total degree k but linear in each variable.

Total and partial degrees of the polynomials. The polynomials $Q_n^{\mathcal{O}}(x_1,\ldots,x_n)$ in the numerator of the factor $K_n^{\mathcal{O}}$ satisfy additional conditions coming from the asymptotic behavior of the form factors. The first condition simply comes from relativistic invariance: in fact, for a simultaneous translation of all the rapidities, the form factors of a scalar operator[2] satisfy

$$F_n^{\mathcal{O}}(\theta_1+\Lambda,\theta_2+\Lambda,\ldots,\theta_n+\Lambda) \;=\; F_n^{\mathcal{O}}(\theta_1,\theta_2,\ldots,\theta_n). \tag{20.2.9}$$

This implies the equality of the total degrees of the polynomials $Q_n^{\mathcal{O}}(x_1,\ldots,x_n)$ and $D_n(x_1,\ldots,x_n)$. Concerning the partial degree with respect to each variable, it is worth anticipating a result discussed in Section 20.8: in order to have a power law behavior of the two-point correlation function of the operator $\mathcal{O}(x)$, its form factors must behave for $\theta_i \to \infty$ at most as $\exp(k\theta_i)$, where k is a constant (independent of i), related to the conformal weight of the operator $\mathcal{O}$.

20.3 Recursive Equations

The poles in the FF induce a set of recursive equations that are crucial for the explicit determination of these functions. As a function of the difference of the rapidities θ_{ij}, the FF have two kinds of simple pole.[3]

Kinematical poles. The first kind of singularity does not depend on whether the model has bound states. It is in fact associated to the kinematical poles at $\theta_{ij} = i\pi$ that come from the one-particle state realized by the three-particle clusters. In turn, these processes correspond to the crossing channels of the S-matrix, as shown in Fig. 20.4. The residues at these poles give rise to a recursive equation that links the n-particle and the $(n-2)$-particle form factors

$$-i\lim_{\tilde\theta\to\theta}(\tilde\theta-\theta)F_{n+2}^{\mathcal{O}}(\tilde\theta+i\pi,\theta,\theta_1,\theta_2,\ldots,\theta_n) = \left(1-e^{2\pi i\gamma}\prod_{i=1}^{n}S(\theta-\theta_i)\right)F_n^{\mathcal{O}}(\theta_1,\ldots,\theta_n). \tag{20.3.1}$$

[2]For the form factors of an operator $\mathcal{O}(x)$ of spin s, the equation generalizes to $F_n^{\mathcal{O}}(\theta_1+\Lambda,\theta_2+\Lambda,\ldots,\theta_n+\Lambda) = e^{s\Lambda}F_n^{\mathcal{O}}(\theta_1,\theta_2,\ldots,\theta_n)$.

[3]There could also be higher order poles, corresponding to the higher order poles of the S-matrix. Their discussion is however beyond the scope of this book.

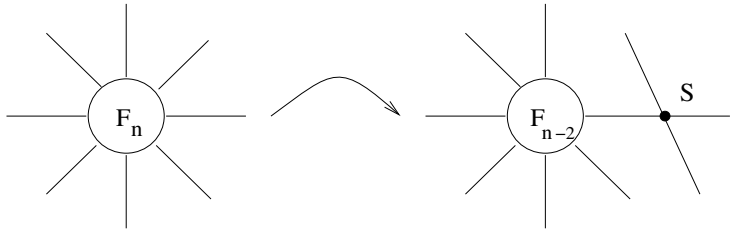


Fig. 20.4 *Recursive equation of the kinematic poles.*

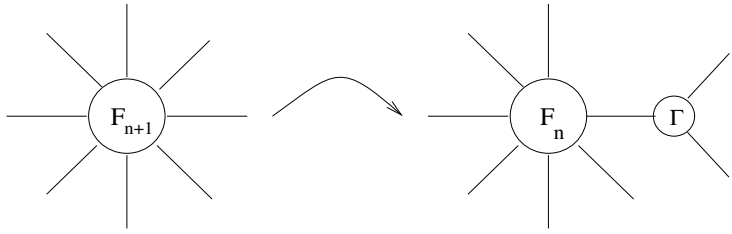


Fig. 20.5 *Recursive equation of the bound state poles.*

Let's denote concisely by $\mathcal{C}$ the map between $F^{\mathcal{O}}_{n+2}$ and $F^{\mathcal{O}}_{n}$ established by the recursive equation

$$F^{\mathcal{O}}_{n+2} \,=\, \mathcal{C}\, F^{\mathcal{O}}_{n}\,. \tag{20.3.2}$$

Bound state poles. There is another family of poles in F_n if the S-matrix has simple poles related to the bound states. These poles are at the values of θ_{ij} corresponding to the resonance angles. Let $\theta_{ij} = iu^k_{ij}$ be one of these poles, associated to the bound state A_k present in the channel $A_i \times A_j$. For the S-matrix we have

$$-i \lim_{\theta \to i u^k_{ij}} (\theta - i u^k_{ij})\, S_{ij}(\theta) \,=\, \left(\Gamma^k_{ij}\right)^2 \tag{20.3.3}$$

where Γ^k_{ij} is the on-shell three-particle vertex and for the residue of the form factor F_{n+1} involving the particles A_i and A_j we have

$$-i \lim_{\epsilon \to 0} \epsilon\, F^{\mathcal{O}}_{n+1}(\theta + i\overline{u}^j_{ik} - \epsilon, \theta - i\overline{u}^i_{jk} + \epsilon, \theta_1, \ldots, \theta_{n-1}) \,=\, \Gamma^k_{ij}\, F^{\mathcal{O}}_{n}(\theta, \theta_1, \ldots, \theta_{n-1}), \tag{20.3.4}$$

where $\overline{u}^c_{ab} \equiv (\pi - u^c_{ab})$. This equation sets up a recursive structure between the $(n+1)$- and the n-particle form factors, as shown in Fig. 20.5. Let's denote by $\mathcal{B}$ the map between F'_{n+1} and $F^{\mathcal{O}}_{n}$ set by this recursive equation

$$F^{\mathcal{O}}_{n+1} \,=\, \mathcal{B}\, F^{\mathcal{O}}_{n}\,. \tag{20.3.5}$$

When the theory presents bound states, it is possible to show that the two kinds of recursive equation are compatible, so that it is possible to reach the $(n+2)$-particle FF by the n-particle FF either using directly the recursive equation shown in Fig. 20.4 or applying the recursive equation of Fig. 20.5 twice. In terms of the mappings $\mathcal{B}$ and $\mathcal{C}$ we have $\mathcal{C} \,=\, \mathcal{B}^2$.

20.4 The Operator Space

At the critical point, one can identify the operator space of a quantum field theory in terms of the irreducible representations of the Virasoro algebra. An extremely interesting point is the characterization of the operator content also away from criticality. As argued below, this can be achieved by means of the form factor theory: although this identification is based on different principles than conformal theories, nevertheless it allows us to shed light on the classification problem of the operators.

Let's start our discussion with some general considerations. In the form factor approach, an operator $\mathcal{O}$ is defined once all its matrix elements $F_n^{\mathcal{O}}$ are known. Notice the particular nature of all the functional equations – the recursive and Watson's equations – satisfied by the form factors: (i) they are all linear; (ii) they do not refer to any particular operator! This implies that, given a fixed number n of asymptotic particles, the solutions of the form factor equations form a linear space. The classification of the operator content is then obtained by putting the vectors of this linear space in correspondence with the operators.

Kernel solutions. Among the functions of these linear spaces, there are those belonging to the kernel of the operators $\mathcal{B}$ and $\mathcal{C}$: these are the functions $\hat{F}_n^{(i)}$ and $\hat{F}_n^{(j)}$ that satisfy

$$\begin{aligned} \mathcal{B}\,\hat{F}_n^{(i)} &= 0 \\ \mathcal{C}\,\hat{F}_n^{(j)} &= 0. \end{aligned} \tag{20.4.1}$$

Their general expression is given in eqn (20.2.3) but, in this case, the function K_n does not contain poles that give rise to the recursive equations. Hence each of the functions $\hat{F}_n^{(i)}$ and $\hat{F}_n^{(j)}$ is simply a symmetric polynomial in the x_i variables. The vector space of the form factors that belong to the kernels can be further specified by assigning the total and partial degrees of these polynomials.

A non-vanishing kernel of the operators $\mathcal{B}$ and $\mathcal{C}$ has the important consequence that at each level n, if $\tilde{F}_n$ is a reference solution of the recursive equation and $\hat{F}_n$ a function of any of the two kernels, the most general form factor can be written as

$$F_n = \tilde{F}_n + \sum_i \alpha_i \hat{F}_n. \tag{20.4.2}$$

Therefore the identification of each operator is obtained by specifying at each level n the constants α_i. If we graphically represent by dots the linearly independent solutions at the level n of the form factor equations, we have the situation of Fig. 20.6. In this graphical representation, each operator is associated to a well-defined path on this lattice, with each step $(n+1) \to n$ (or $(n+2) \to n$) ruled by the operator $\mathcal{B}$ (or $\mathcal{C}$). We will see explicit examples of this operator structure when we discuss the form factors of the Ising and the Sinh–Gordon models.

20.5 Correlation Functions

Once we have determined the form factors of a given operator, its correlation functions can be written in terms of the spectral representation series using the completeness

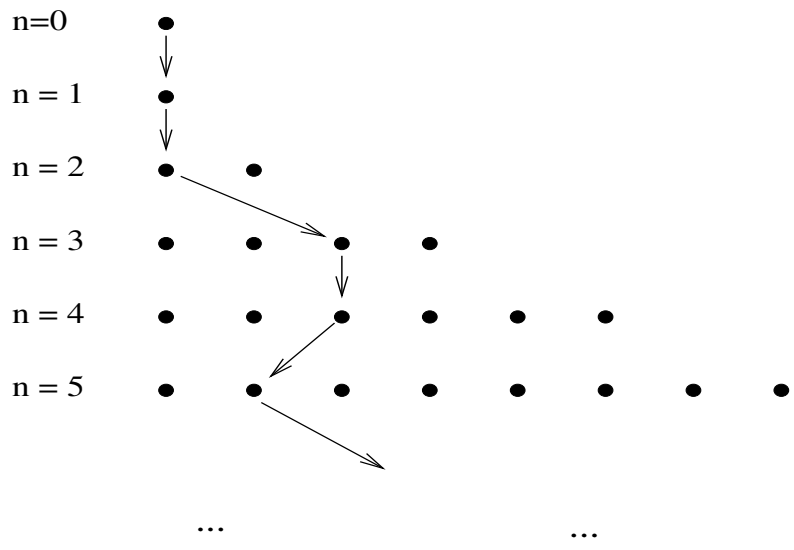


Fig. 20.6 *Vector spaces of the solutions of the form factor equations (the number of dots at each level is purely indicative). An operator is associated to the sequence of its matrix elements* F_n.

relation of the multiparticle states

$$1 = \sum_{n=0}^{\infty} \int \frac{d\theta_1 \dots d\theta_n}{n!(2\pi)^n} |\theta_1, \dots, \theta_n\rangle \langle \theta_1, \dots, \theta_n|. \tag{20.5.1}$$

For instance, for the two-point correlation function of the operator $\mathcal{O}(x)$ in euclidean space, we have

$$\begin{aligned} \langle \mathcal{O}(x)\, \mathcal{O}(0) \rangle &= \sum_{n=0}^{\infty} \int \frac{d\theta_1 \dots d\theta_n}{n!(2\pi)^n} \langle 0|\mathcal{O}(x)|\theta_1, \dots, \theta_n\rangle_{\text{in}} \langle \theta_1, \dots, \theta_n|\mathcal{O}(0)|0\rangle \\ &= \sum_{n=0}^{\infty} \int \frac{d\theta_1 \dots d\theta_n}{n!(2\pi)^n} \mid F_n(\theta_1 \dots \theta_n) \mid^2 \exp\left(-mr \sum_{i=1}^{n} \cosh\theta_i \right) \end{aligned} \tag{20.5.2}$$

where r is the radial distance $r = \sqrt{x_0^2 + x_1^2}$ (Fig. 20.7). Similar expressions, although more complicated, hold for the n-point correlation functions. It is worth making some comments to clarify the nature of these expressions and their advantage.

- The integrals that enter the spectral series are all convergent. This is in sharp contrast with the formalism based on the Feynman diagrams, in which one encounters the divergences of the various perturbative terms. In a nutshell, the deep reason of this difference between the two formalisms can be expressed as follows. The Feynman formalism is based on the quantization of a *free* theory and on the *bare* unphysical parameters of the lagrangian. What the renormalization

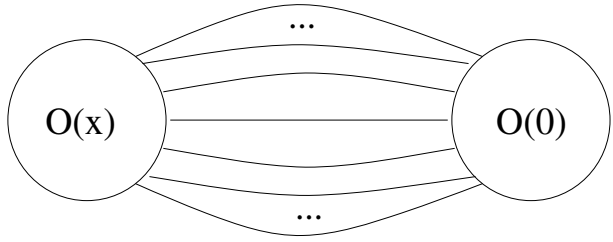


Fig. 20.7 *Spectral representation of the two-point correlation functions.*

procedure does is to implement the change from the bare to the physical parameters (such as the physical value of the mass of the particle). But the form factors employ *ab initio* all the physical parameters of the theory and for this reason the divergences of the perturbative series are absent.

- If the S-matrix depends on a coupling constant, as it happens in the Sinh–Gordon model or in other Toda field theories, each matrix element provides the exact resummation of all terms of perturbation theory.
- If the correlation functions do not have particularly violent ultraviolet singularities (this is the case, for instance, of the correlation functions of the relevant fields), the corresponding spectral series has an extremely fast convergent behavior for all values of mr. In the infrared region, that is for large values of mr, this is pretty evident from the nature of the series, because its natural parameter of expansion is e^{-mr}. The reason of the fast convergent behavior also in the ultraviolet region $mr \to 0$ is twofold: the peculiar behavior of the n-particle phase space in two-dimensional theories (see Appendix C of Chapter 17) and a further enhancement of the convergence provided by the form factors. To better understand this aspect, consider the Fourier transform of the correlator

$$G(x) \;=\; \langle \mathcal{O}(x)\, \mathcal{O}(0) \rangle \;=\; \int \frac{d^2p}{(2\pi)^2}\, e^{ip\cdot x}\, \hat{G}(p). \tag{20.5.3}$$

The function $\hat{G}(p)$ can be written as

$$\hat{G}(p) \;=\; \int_0^\infty d\mu^2\, \rho(\mu^2)\, \frac{1}{p^2+\mu^2}, \tag{20.5.4}$$

where $\rho(k^2)$ is a relativistically invariant function called the *spectral density*

$$\rho(k^2) \;=\; 2\pi \sum_{n=0}^{\infty} \int d\Omega_1 \ldots d\Omega_n\, \delta^2(k - P_n)\, |\langle 0\, |\mathcal{O}(0)\, |\theta_1, \ldots, \theta_n\rangle|^2$$

$$d\Omega \;=\; \frac{dp}{2\pi E} \;=\; \frac{d\theta}{2\pi}, \quad P_n^{(0)} = \sum_{k=0}^{n} \cosh\theta_k, \quad P_n^{(1)} = \sum_{k=0}^{n} \sinh\theta_k.$$

Since $1/(p^2+\mu^2)$ is the two-point correlation function of the euclidean free theory with mass μ, i.e. the propagator, eqn (20.5.4) shows that the two-point correlation function can be regarded as a linear superposition of the free propagators weighted

with the spectral density $\rho(\mu^2)$. Notice that the contribution given by the single-particle state of mass m in the spectral density is given by

$$\rho_{1part}(k^2) \;=\; \frac{1}{2\pi}\delta(k^2 - m^2). \tag{20.5.5}$$

To analyze the behavior of $\rho(k^2)$ by varying k^2, let's make the initial approximation to take equal to 1 all the matrix elements. In this way, each term of the spectral series coincides with the n-particle phase space

$$\Phi_n(k^2) \equiv \int \prod_{k=1}^{n} d\Omega_k \, \delta^2(k - P_n). \tag{20.5.6}$$

As shown in Appendix C of Chapter 17, in two dimensions the space goes to zero when $k^2 \to \infty$ as

$$\Phi_n(k^2) \simeq \frac{1}{(2\pi)^{n-2}} \, \frac{1}{(n-2)!} \, \frac{1}{k^2} \, \left(\log \frac{k^2}{m^2}\right)^{n-2}, \tag{20.5.7}$$

whereas for $d > 2$ it diverges as

$$\Phi_n(k^2) \;\sim\; k^{\frac{n(d-2)-d}{2}}. \tag{20.5.8}$$

On the other hand, $\Phi_n(k^2) = 0$ if $k^2 < (n\,m)^2$ and near the threshold values we have

$$\Phi(k^2) \;\simeq\; A_n \left(\sqrt{k^2} - n\,m\right)^{\frac{n-3}{2}}. \tag{20.5.9}$$

Hence, we see that for pure reasons related to the phase space we have two different scenarios for the quantum field theories in two dimensions and in higher dimensions: while in $d > 2$ surpassing the various thresholds the spectral density receives contributions that are more divergent, in $d = 2$ they are all of the same order and all go to zero at large values of the energy. Hence, for $d > 2$ it is practically impossible to approximate the spectral density for large values of k^2 by using the first terms of the series, relative to the states with few particles, whereas in $d = 2$ this is perfectly plausible. If we now include in the discussion also the form factors, one realizes that the situation is even better in $d = 2$! In fact, from the general expression (20.2.3) and for the vanishing of $F_{\min}(\theta_{ij})$ at the origin (eqn 20.2.5), the form factors vanish at the n-particle thresholds as

$$|\langle 0\,|\mathcal{O}(0)\,|\theta_1, \ldots, \theta_n\rangle|^2 \;\simeq\; \left(\sqrt{k^2} - n\,m\right)^{n(n-1)}, \qquad \theta_1 \simeq \ldots \simeq \theta_n \simeq 0 \tag{20.5.10}$$

while, for large values of their rapidities, they typically tend to a constant.[4] This scenario implies that the spectral density of the correlation functions of the two-dimensional integrable models usually flatten more at the thresholds and therefore becomes a very smooth function varying as k^2 (see Fig. 20.8). For all these reasons, the spectral density can be approximated with great accuracy just by taking the first terms of the series, even for large values of k^2, therefore leading to fast convergent behavior also in the ultraviolet region.

[4]This is what usually happens for the form factors of the strongly relevant operators.

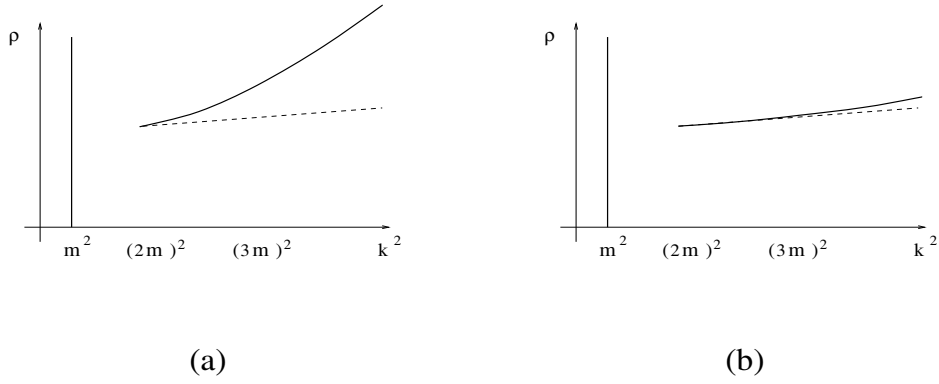


Fig. 20.8 *Plot of the spectral series in a model in $d = 4$ (a) and in $d = 2$ (b). The contribution of the two-particle state is given by the dashed line. In $d = 4$ this does not provide a good approximation of $\rho(k^2)$ for large values of k^2 while in $d = 2$ it very often gives an excellent approximation of this quantity.*

20.6 Form Factors of the Stress–Energy Tensor

The stress–energy tensor is an operator that plays an important role in quantum field theory and its form factors have special properties. From its conservation law $\partial_\mu T^{\mu\nu}(x) = 0$, this operator can be written in terms of an auxiliary scalar field $A(x)$ as

$$T_{\mu\nu}(x) = (\partial_\mu\partial_\nu - g_{\mu\nu}\Box)\, A(x). \tag{20.6.1}$$

In light-cone coordinates, $x^\pm = x^0 \pm x^1$, its components are

$$T_{++} = \partial_+^2\, A, T_{--} = \partial_-^2\, A,$$

$$\Theta = T_\mu^\mu = -\Box\, A = -4\,\partial_+\partial_-\, A.$$

Introducing the variables $x_j = e^{\theta_j}$ and the elementary symmetric polynomials $\sigma_i^{(n)}$, it is easy to see that

$$\begin{aligned} F_n^{T_{++}}(\theta_1,\ldots,\theta_n) &= -\frac{1}{4}\, m^2 \left(\frac{\sigma_{n-1}^{(n)}}{\sigma_n^{(n)}}\right)^2 F_n^A(\theta_1,\ldots,\theta_n), \\ F_n^{T_{--}}(\theta_1,\ldots,\theta_n) &= -\frac{1}{4}\, m^2 \left(\sigma_1^{(n)}\right)^2 F_n^A(\theta_1,\ldots,\theta_n), \\ F_n^{\Theta}(\theta_1,\ldots,\theta_n) &= m^2\, \frac{\sigma_1^{(n)}\sigma_{n-1}^{(n)}}{\sigma_k}\, F_n^A(\theta_1,\ldots,\theta_n). \end{aligned} \tag{20.6.2}$$

Solving for F_n^A, we have

$$F_n^{T_{++}}(\theta_1,\ldots,\theta_n) \;=\; -\,\frac{1}{4}\,\frac{\sigma_{n-1}^{(n)}}{\sigma_1^{(n)}\sigma_n^{(n)}}\,F_n^{\Theta}(\theta_1,\ldots,\theta_n),$$

$$F_n^{T_{--}}(\theta_1,\ldots,\theta_n) \;=\; -\,\frac{1}{4}\,\frac{\sigma_1^{(n)}\sigma_n^{(n)}}{\sigma_{n-1}^{(n)}}\,F_n^{\Theta}(\theta_1,\ldots,\theta_n). \tag{20.6.3}$$

Hence, the whole set of form factors of $T_{\mu\nu}$ can be recovered by the form factors of the trace Θ. This is a scalar operator and therefore its form factors depend on the differences of the rapidities $\theta_{ij} = \theta_i - \theta_j$. Moreover, since the form factors of T_{--} and T_{++} must have the same singularities as those of Θ, $F_n^{\Theta}(\theta_1,\ldots,\theta_n)$ (for $n>2$) has to be proportional to the combination $\sigma_1^{(n)}\sigma_{n-1}^{(n)}$ of the elementary symmetric polynomials. This combination corresponds to the relativistic invariant given by the total energy and momentum of the system.

For the normalization of these matrix elements, the recursive structure reduces the problem of finding the normalization of the form factors of $\Theta(x)$ on the one and two-particle states, i.e. $F_1^{\Theta}(\theta)$ and $F_2^{\Theta}(\theta_{12})$. The normalization of $F_2^{\Theta}(\theta_{12})$ can be determined by using the total energy of the system

$$E \;=\; \frac{1}{2\pi}\int_{-\infty}^{+\infty} dx^1\, T^{00}(x). \tag{20.6.4}$$

Computing the matrix element of both terms of this equation on the asymptotic states $\langle\theta'|$ and $|\theta\rangle$, for the left-hand side we have

$$\langle\theta'|\,E\,|\theta\rangle \;=\; 2\pi\, m\,\cosh\theta\,\delta(\theta'-\theta).$$

On the other hand, taking into account that $T^{00} = \partial_1^2 A$ and using the relation

$$\langle\theta'|\,\mathcal{O}(x)|\theta\rangle \;=\; e^{i\left(p^\mu(\theta')\,-\,p^\mu(\theta)\right)x_\mu}\,F_2^{\mathcal{O}}(\theta,\,\theta'\,-\,i\pi)$$

which holds for any hermitian operator $\mathcal{O}$, we obtain

$$F_2^{\partial_1^2 A}(\theta_1,\theta_2) \;=\; -\,m^2\,(\sinh\theta_1\,+\sinh\theta_2)^2\,F_2^A(\theta_{12}).$$

From eqns (20.6.2) and (20.6.4) it follows that the normalization of F_2^{Θ} is given by

$$F_2^{\Theta}(i\pi) \;=\; 2\,\pi\,m^2. \tag{20.6.5}$$

However, there is no particular constraint on the one-particle form factor of $\Theta(x)$ coming from general considerations

$$F_1^{\Theta} \;=\; \langle 0\mid\Theta(0)\mid\theta\rangle. \tag{20.6.6}$$

This is a free parameter of the theory, related to the intrinsic ambiguity of $T^{\mu\nu}(x)$, since this tensor can always be modified by adding a total divergence (see Problem 1).

20.7 Vacuum Expectation Values

The recursive equations enable us to determine the form factors $F_n^{\mathcal{O}}$ in terms of the previous $F_{n-1}^{\mathcal{O}}$ or $F_{n-2}^{\mathcal{O}}$. At the bottom of this iterative structure there are, as its initial seeds, the lowest quantities $F_0^{\mathcal{O}}$, i.e. the vacuum expectation value of the operator $\mathcal{O}$ and F_1, i.e. its matrix elements on one-particle states. Presently it is not known how to determine in general all the one-particle matrix elements. However, the situation is much better for the vacuum expectation values: they can be computed exactly for several operators both of the Sine–Gordon and Bullogh–Dodd models, as well as of RSOS restrictions thereof. The theoretical steps that lead to these results are quite technical but well described in the series of papers quoted at the end of the chapter and will not be reviewed here. In this section we will simply present the most relevant formulas, in particular, the exact vacuum expectation values of primary fields in integrable perturbed conformal field theories, with respect to the deformations $\Phi_{1,3}$, $\Phi_{1,2}$, and $\Phi_{2,1}$. In the following to denote such theories we use the notation

$$\mathcal{S}_m^{(k,l)\pm} = \mathcal{S}_m^{(CFT)} \pm \lambda \int d^2x\, \Phi_{k,l}(x), \tag{20.7.1}$$

where $\mathcal{S}_m$ is the action of the m-th conformal minimal model, $\Phi_{r,s}$ is the relevant primary field that leads to an integrable model, and $\lambda > 0$ is its dimensional coupling constant setting the scale of the quantum field theory (the sign of the coupling only makes sense after fixing the normalization of the fields $\Phi_{r,s}$). Hereafter

$$x \equiv (m+1)k - ml.$$

Integrable theory $\mathcal{S}_m^{(1,3)-}$. For $\lambda > 0$, $\Phi_{1,3}$ induces a massless flow between the minimal models $\mathcal{M}_m \to \mathcal{M}_{m-1}$ (see Section 15.6). For $\lambda < 0$, $\Phi_{1,3}$ drives instead the conformal model into a massive phase where there are kinks interpolating the $(m-1)$ RSOS degenerate vacua labeled as

$$\mathbf{a} = 0,\ \frac{1}{2}, \ldots, \frac{(m-2)}{2}.$$

For the vacuum expectation values of the primary fields on the various vacua we have

$$\langle \mathbf{a}|\Phi_{k,l}|\mathbf{a}\rangle^{(1,3)-} = \frac{\sin\left(\frac{\pi(2a+1)}{m}((m+1)k-ml)\right)}{\sin\frac{\pi(2a+1)}{m}} F_{k,l}^m(x) \tag{20.7.2}$$

where

$$F_{k,l}^m(x) = \left(M \frac{\sqrt{\pi}\Gamma\left(\frac{m+3}{2}\right)}{2\Gamma\left(\frac{m}{2}\right)}\right)^{2\Delta_{k,l}} \mathcal{Q}_{1,3}(x)$$

and

$$\mathcal{Q}_{1,3}(\eta) = \exp\left\{\int_0^\infty \frac{dt}{t}\left[\frac{\cosh(2t)\sinh((\eta-1)t)\sinh((\eta+1)t)}{2\cosh(t)\sinh(mt)\sinh((1+m)t)} - \frac{\eta^2-1}{2m(m+1)}\mathrm{e}^{-4t}\right]\right\}.$$

In the formula above

$$M = \frac{2\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m+1}{2}\right)}\left[\frac{\pi\lambda(1-m)(2m-1)}{(1+m)^2}\sqrt{\frac{\Gamma\left(\frac{1}{m+1}\right)\Gamma\left(\frac{1-2m}{m+1}\right)}{\Gamma\left(\frac{m}{m+1}\right)\Gamma\left(\frac{3m}{m+1}\right)}}\right]^{\frac{1+m}{4}}$$

is the common mass of the kinks expressed in term of the coupling constant λ.

Integrable theory $\mathcal{S}_m^{(1,2)}$. For the integrable model $\mathcal{S}_m^{(1,2)}$, the vacuum structure of the theory depends on whether m is odd or even.

- m **even.** When m is even, the field $\Phi_{1,2}$ is even under the Z_2 spin symmetry and the two theories $\mathcal{S}_m^{(1,2)\pm}$ are different although related by duality. The number of RSOS vacua of $\mathcal{S}_m^{(1,2)+}$ is equal to $(m-2)/2$, while the number of vacua of $\mathcal{S}_m^{(1,2)-}$ is equal to $m/2$. Their label is

$$\mathbf{a} = \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-3}{2}, \quad \lambda > 0$$
$$\mathbf{a} = 0, 1, \ldots, \frac{m-2}{2}, \quad \lambda < 0.$$

- m **odd**. In this case the field $\Phi_{1,2}$ is odd under the Z_2 symmetry and the two theories $\mathcal{S}_m^{(1,2)\pm}$ are equal. There are $(m-1)/2$ degenerate vacua in both theories that we label as

$$a = \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-2}{2}, \quad \lambda > 0$$
$$a = 0, 1, \ldots, \frac{m-3}{2}, \quad \lambda < 0.$$

The vacuum expectation values of the primary fields on the various vacua are:

$$\langle \mathbf{a}|\Phi_{k,l}|\mathbf{a}\rangle^{(1,2)} = \frac{\sin\left(\frac{\pi(2a+1)}{m}((m+1)k - ml)\right)}{\sin\frac{\pi(2a+1)}{m}} G_{k,l}^m(x) \tag{20.7.3}$$

where

$$G_{k,l}^m(x) = \left(M\frac{\pi(m+1)\Gamma\left(\frac{2m+2}{3m+6}\right)}{2^{\frac{2}{3}}\sqrt{3}\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{m}{3m+6}\right)}\right)^{2\Delta_{k,l}} \mathcal{Q}_{1,2}(x)$$

and

$$\mathcal{Q}_{1,2}(\eta) = \exp\left\{\int_0^\infty \frac{dt}{t}\Big[\frac{\sinh((m+2)t)\sinh((\eta-1)t)\sinh((\eta+1)t)}{\sinh(3(m+2)t)\sinh(2(m+1)t)\sinh(mt)}\right.$$
$$\times(\cosh(3(m+2)t) + \cosh((m+4)t) - \cosh((3m+4)t) + \cosh(mt) + 1)$$
$$\left.-\frac{\eta^2-1}{2m(m+1)}\mathrm{e}^{-4t}\Big]\right\}.$$

In the formula above

$$M = \frac{2^{\frac{m+5}{3m+6}}\sqrt{3}\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{m}{3m+6}\right)}{\pi\Gamma\left(\frac{2m+2}{3m+6}\right)}\left[\frac{\pi^2\lambda^2\Gamma\left(\frac{3m+4}{4m+4}\right)\Gamma\left(\frac{1}{2}+\frac{1}{m+1}\right)}{\Gamma\left(\frac{m}{4m+4}\right)\Gamma\left(\frac{1}{2}-\frac{1}{m+1}\right)}\right]^{\frac{m+1}{3m+6}}$$

is the mass of the kinks expressed in terms of the coupling constant λ.

Integrable theory $\mathcal{S}_m^{(2,1)}$. For this theory the situation is reversed with respect to the previous one: $\Phi_{2,1}$ is odd under the Z_2 symmetry when m is even (and the theory is independent of the sign of its coupling), while it is a Z_2 even field when m is odd (and the theories with $\lambda > 0$ and $\lambda < 0$ are related by duality). For the RSOS degenerate vacua, in this case we have the following labels:

- when m is even, both for $\lambda > 0$ and $\lambda < 0$, their number is $m/2$ and

$$\begin{aligned} a &= \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-1}{2}, \quad \lambda > 0 \\ a &= 0, 1, \ldots, \frac{m-2}{2}, \quad \lambda < 0; \end{aligned}$$

- when m is odd, their number is $(m-1)/2$ for $\lambda > 0$, and $(m+1)/2$ for $\lambda < 0$, with

$$\begin{aligned} a &= \frac{1}{2}, \frac{3}{2}, \ldots, \frac{m-2}{2}, \quad \lambda > 0 \\ a &= 0, 1, \ldots, \frac{m-1}{2}, \quad \lambda < 0. \end{aligned}$$

The vacuum expectation values of the primary fields on the various vacua are the expectation values

$$\langle \mathbf{a}|\Phi_{k,l}|\mathbf{a}\rangle^{(2,1)} = \frac{\sin\left(\frac{\pi(2a+1)}{m+1}((m+1)k - ml)\right)}{\sin\frac{\pi(2a+1)}{m+1}} H_{k,l}^m(x) \tag{20.7.4}$$

where

$$H_{k,l}^m(x) = \left(M\frac{\pi m\Gamma\left(\frac{2m}{3m-3}\right)}{2^{\frac{2}{3}}\sqrt{3}\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{m+1}{3m-3}\right)}\right)^{2\Delta_{k,l}} \mathcal{Q}_{2,1}(x)$$

and

$$\begin{aligned} \mathcal{Q}_{2,1}(\eta) = \exp\Bigg\{ \int_0^\infty \frac{dt}{t}\Big[& \frac{\sinh((m-1)t)\sinh((\eta-1)t)\sinh((\eta+1)t)}{\sinh(3(m-1)t)\sinh(2mt)\sinh((m+1)t)} \\ & \times(\cosh(3(m-1)t) + \cosh((m-3)t) - \cosh((3m-1)t) \\ & + \cosh((m+1)t) + 1) - \frac{\eta^2-1}{2m(m+1)}\mathrm{e}^{-4t}\Big]\Bigg\}. \end{aligned} \tag{20.7.5}$$

The mass of the kinks, as a function of the coupling constant λ, is expressed by

$$M = \frac{2^{\frac{m-4}{3m-3}}\sqrt{3}\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{m+1}{3m-3}\right)}{\pi\Gamma\left(\frac{2m}{3m-3}\right)}\left[\frac{\pi^2\lambda^2\Gamma\left(\frac{3m-1}{4m}\right)\Gamma\left(\frac{1}{2}-\frac{1}{m}\right)}{\Gamma\left(\frac{m+1}{4m}\right)\Gamma\left(\frac{1}{2}+\frac{1}{m}\right)}\right]^{\frac{m}{3m-3}}.$$

20.8 Ultraviolet Limit

In the ultraviolet limit, the correlation functions of the scaling operators has a power law behavior, dictated by the conformal weight of the operator

$$G(r) = \langle\mathcal{O}(r)\,\mathcal{O}(0)\rangle \simeq \frac{1}{r^{4\Delta}}, \quad r \to 0. \tag{20.8.1}$$

One may wonder how the spectral series (20.5.2), which is based on the exponential terms $e^{-k\,mr}$, is able to reproduce a power law in the limit $r \to 0$. The answer to this question comes from an interesting analogy.

Feynman gas. Note that the formula (20.5.2) is formally similar to the expression of the grand-canonical partition function of a fictitious one-dimensional gas

$$\mathcal{Z}(mr) = \sum_{n=0}^{\infty} z^n\, Z_n. \tag{20.8.2}$$

To set up the vocabulary of this analogy, let's identify the coordinates of the gas particles with the rapidities θ_i, with the Boltzmann weight relative to the interactive potential of the gas with the modulus squared of the form factors

$$e^{-V(\theta_1,\ldots,\theta_n)} \equiv |\langle 0\,|\mathcal{O}(0)|\theta_1,\ldots,\theta_n\rangle|^2. \tag{20.8.3}$$

Finally, let's identify the fugacity of the gas with

$$z(\theta) = \frac{1}{2\pi}\,e^{-mr\cosh\theta}. \tag{20.8.4}$$

We have defined in this way the *Feynman gas* that was analyzed at the end of Chapter 2. The only difference with respect to the standard case is the coordinate dependence of the fugacity of this gas. Although the coordinates of the particles of this gas span the infinite real axis, the effective volume of the system is however determined by the region in which the fugacity (20.8.4) is significantly different from zero, as shown in Fig. 20.9. Note that $z(\theta)$ is a function that rapidly goes to zero outside a finite interval and, in the limit $mr \to 0$, presents a plateau of height $z_c = 1/(2\pi)$ and width

$$L \simeq 2\,\log\frac{1}{mr}.$$

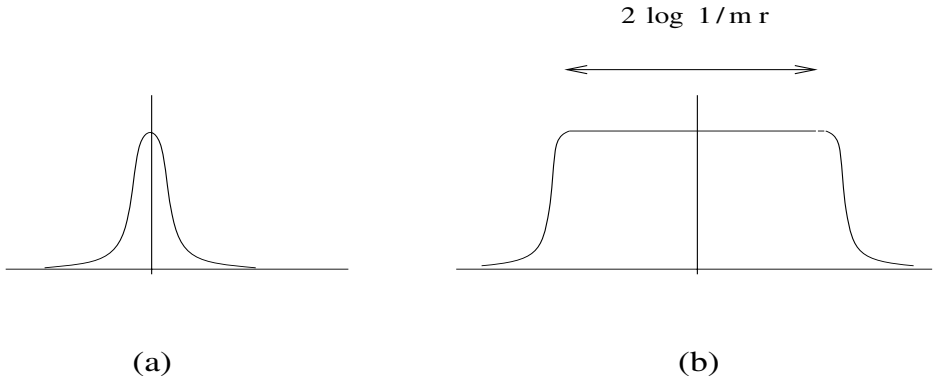


Fig. 20.9 *Plot of the fugacity as a function of θ: (a) for finite values of* (mr)*; (b) in the limit* $(mr) \to 0$.

The equation of state of a one-dimensional gas is given by

$$\mathcal{Z} \,=\, e^{p(z)L},$$

where $p(z)$ is the pressure as a function of the fugacity. Following this analogy, for the two-point correlation function in the limit $(mr) \to 0$, we have

$$G(r) \,=\, \mathcal{Z} \,=\, e^{p(z_c)\,L} \,\simeq\, e^{2p(z_c)\log 1/(mr)} \,=\, \left(\frac{1}{mr}\right)^{2p(z_c)}, \tag{20.8.5}$$

i.e. a power law behavior! Moreover, comparing with the short-distance behavior of the correlator given in eqn (20.8.1), there is an interesting result: the conformal weight can be expressed in terms of the pressure of this fictitious one-dimensional gas, evaluated at the plateau value of the fugacity

$$2\,\Delta \,=\, p(1/2\pi). \tag{20.8.6}$$

Besides the thermodynamics of the Feynman gas, the conformal weight of the operators can also be extracted by applying the sum rule given by the Δ-theorem (see Chapter 15)

$$\Delta \,=\, -\frac{1}{2\langle\mathcal{O}\rangle}\int_0^\infty dr\, r\,\langle\Theta(r)\mathcal{O}(0)\rangle. \tag{20.8.7}$$

To compute this quantity, it is necessary to know the form factors of the operator $\mathcal{O}(x)$ and the trace of the stress–energy tensor $\Theta(x)$.

c-theorem sum rule. Additional control of the ultraviolet limit of the theory is provided by the sum-rule of the c-theorem: it gives the central charge of conformal field theory associated to the ultraviolet limit of the massive theory through the integral

$$c \,=\, \frac{3}{2}\int_0^\infty dr\, r^3\,\langle\Theta(r)\Theta(0)\,\rangle_c.$$

Using the spectral representation of this correlator we have

$$c = \sum_{n=1}^{\infty} c_n, \tag{20.8.8}$$

where the n-particle contribution is

$$\begin{aligned} c_n \;=\; & \frac{12}{n!}\int_0^\infty \frac{d\mu}{\mu^3}\int_{-\infty}^{\infty}\frac{d\theta_1}{2\pi}\cdots\frac{d\theta_n}{2\pi} \\ & \times\delta\left(\sum_{i=1}^{n}\sinh\theta_i\right)\,\delta\left(\sum_{i=1}^{n}\cosh\theta_i-\mu\right)\,|\langle 0|\Theta(0)|\theta_1,\ldots,\theta_n\rangle|^2. \end{aligned} \tag{20.8.9}$$

Usually this series presents very fast behavior. This permits us to obtain rather accurate estimations of the central charge c, with an explicit check of the entire formalism of the S-matrix and form factors. It is easy to understand the reason for this fast convergence by studying the integrand, shown in Fig. 20.10: the term r^3 kills the singularity of the correlator at short distance (therefore the integrand vanishes at the origin), while it weights the correlator more at large distances. But this is just the region where a few terms of the spectral series are very efficient in approximating the correlation function with high accuracy.

Asymptotic behavior. Finally, let's discuss the upper bound on the asymptotic behavior of the form factors dictated by the ultraviolet behavior of the correlator (20.8.1). To establish this bound, let's start by noting that in a massive theory we have

$$M_p \equiv \int d^2x\,|x|^p\,\langle\,\mathcal{O}(x)\mathcal{O}(0)\,\rangle_c \qquad < +\infty \qquad \text{if} \qquad p > 4\Delta_{\mathcal{O}} - 2. \tag{20.8.10}$$

Employing now the spectral representation of the correlator (20.5.3) and integrating over p, μ, and x, we get

$$M_p \sim \sum_{n=1}^{\infty}\int_{\theta_1>\ldots>\theta_n} d\theta_1\ldots d\theta_n\,\frac{|F_n^{\mathcal{O}}(\theta_1,\ldots,\theta_n)|^2}{(\sum_{k=1}^{n} m_k\cosh\theta_k)^{p+1}}\;\delta\left(\sum_{k=1}^{n} m_k\sinh\theta_k\right). \tag{20.8.11}$$

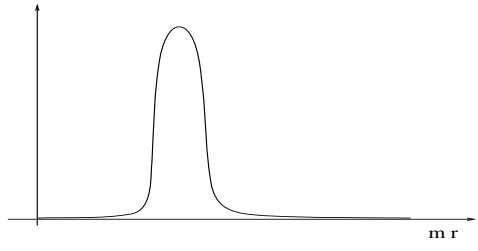


Fig. 20.10 *Plot of the integrand $r^3\langle\Theta(r)\Theta(0)\rangle$ in the C-theorem sum rule.*

Equation (20.8.10) can now be used to find an upper limit on the real quantity y_Φ, defined by

$$\lim_{|\theta_i|\to\infty} F_n^{\mathcal{O}}(\theta_1,\ldots,\theta_n) \sim e^{y_\Phi|\theta_i|}. \tag{20.8.12}$$

In fact, taking the limit $\theta_i \to +\infty$ in the integrand of (20.8.11), the delta-function forces some other rapidities to move at $-\infty$ as $-\theta_i$. Because the matrix element $F_n^{\mathcal{O}}(\theta_1,\ldots,\theta_n)$ depends on the differences of the rapidities, it contributes to the integrand with the factor $e^{2y_\Phi|\theta_i|}$ in the limit $|\theta_i| \to \infty$. Hence, eqn (20.8.10) leads to the condition

$$y_{\mathcal{O}} \leq \Delta_{\mathcal{O}}. \tag{20.8.13}$$

This equation provides information on the partial degree of the polynomial $Q_n^{\mathcal{O}}$. Note, however, that this conclusion may not apply for non-unitary theories because not all terms of the expansion on the intermediate states are necessarily positive in this case.

20.9 The Ising Model at $T \neq T_c$

In this section we present the form factors and the correlation functions of the relevant operators $\epsilon(x)$, $\sigma(x)$, and $\mu(x)$ of the two-dimensional Ising model when the temperature T is away from its critical value. From the duality of the model, we can discuss equivalently the case $T > T_c$ or $T < T_c$. Suppose the system is in the high-temperature phase where the scattering theory of the off-critical model involves only one particle with an S-matrix $S = -1$. There are no bound states. The particle A can be considered as being created by the magnetization operator $\sigma(x)$, so that it is odd under the Z_2 symmetry of the Ising model, with its mass given by $m = |T - T_c|$.

Let's now employ the form factor equations to find the matrix elements of the various operators on the multiparticle states. The first step is the determination of the function $F_{min}(\theta)$ which satisfies

$$\begin{aligned} F_{\min}(\theta) &= -F_{\min}(-\theta) \\ F_{\min}(i\pi-\theta) &= F_{\min}(i\pi+\theta). \end{aligned} \tag{20.9.1}$$

The minimal solution is

$$F_{\min}(\theta) = \sinh\frac{\theta}{2}. \tag{20.9.2}$$

20.9.1 The Energy Operator

Let's initially discuss the form factors of the energy operator $\epsilon(x)$ or, equivalently, those of the trace of the stress–energy tensor, since the two operators are related by

$$\Theta(x) = 2\pi m\,\epsilon(x). \tag{20.9.3}$$

This is an even operator under the Z_2 symmetry and therefore it has matrix elements only on states with an even number of particles, F_{2n}^{Θ}. The recursive equations of the kinematical poles are particularly simple

$$-i\lim_{\tilde\theta\to\theta}(\tilde\theta-\theta)F_{2n+2}^{\Theta}(\tilde\theta+i\pi,\theta,\theta_1,\theta_2,\ldots,\theta_{2n}) = \left(1-(-1)^{2n}\right)F_{2n}^{\Theta}(\theta_1,\ldots,\theta_{2n}) = 0. \tag{20.9.4}$$

Taking into account the normalization of the trace operator $F_2^{\Theta}(i\pi) = 2\pi m^2$, the simplest solution of all these equations is

$$F_{2n}^{\Theta}(\theta_1, \ldots, \theta_{2n}) = \begin{cases} -2\pi i\, m^2 \, \sinh\frac{\theta_1-\theta_2}{2} & , n = 2 \\ 0 & , \text{otherwise.} \end{cases} \tag{20.9.5}$$

In the light of the discussion in Section 20.4, note that the identification of the operator Θ with this specific sequence of form factors is equivalent to putting equal to zero all coefficients of the kernel solutions $F_{2n}^{(i)}$ at all the higher levels.

We have an explicit check that (20.9.5) is the correct sequence of the form factors of the trace operator which comes from its two-point correlation function and from the c-theorem. For the correlator we get

$$\begin{aligned} G^{\Theta}(r) = \langle\Theta(r)\Theta(0)\rangle &= \frac{1}{2}\int \frac{d\theta_1}{2\pi}\frac{d\theta_2}{2\pi}|F_2^{\Theta}(\theta_{12})|^2\, e^{-mr(\cosh\theta_1+\cosh\theta_1)} \\ &= \frac{m^4}{2}\int d\theta_1\, d\theta_2\, \sinh^2\frac{\theta_1-\theta_2}{2}\, e^{-mr(\cosh\theta_1+\cosh\theta_2)} \\ &= \frac{m^4}{4}\int d\theta_1\, d\theta_2\, [\cosh(\theta_1-\theta_2)-1]\, e^{-mr(\cosh\theta_1-\cosh\theta_2)} \\ &= m^4\left(\left[\int d\theta\cosh\theta\, e^{-mr\cosh\theta}\right]^2 - \left[\int d\theta\, e^{-mr\cosh\theta}\right]^2\right) \\ &= m^4\left(K_1^2(mr) - K_0^2(mr)\right) \end{aligned} \tag{20.9.6}$$

where, in the last line, we used the integral representation of the modified Bessel functions

$$K_\nu(z) = \int_0^\infty dt\, \cosh\nu t\, e^{-z\cosh t}.$$

Hence, we have

$$G^{\Theta}(r) = \langle\Theta(r)\Theta(0)\rangle = m^4\left[K_1^2(mr) - K_0^2(mr)\right]. \tag{20.9.7}$$

whose plot is in Fig. 20.11. This function has the correct ultraviolet behavior associated to the energy operator

$$G^{\Theta}(r) \to \frac{m^2}{|x|^2}, \qquad |x| \to 0. \tag{20.9.8}$$

Substituting the expression above in the c-theorem, we get the correct value of the central charge of the Ising model

$$c = \frac{3}{2}\int_0^\infty dr\, r^3\langle\Theta(r)\Theta(0)\rangle = \frac{1}{2}. \tag{20.9.9}$$

20.9.2 Magnetization Operators

In the high-temperature phase, the order parameter $\sigma(x)$ is odd under the Z_2 symmetry while the disorder operator $\mu(x)$ is even. Hence, $\sigma(x)$ has matrix elements on states

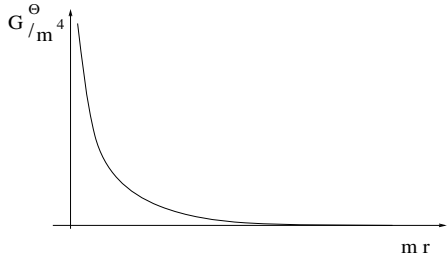


Fig. 20.11 *Plot of the two-point correlation function of the trace of the stress–energy tensor for the thermal Ising model.*

with an odd number of particles, F^{σ}_{2n+1}, whereas $\mu(x)$ is on an even number, F^{μ}_{2n}. In writing down the residue equations relative to the kinematical poles, we have to take into account that the operator μ has a semilocal index equal to $1/2$ with respect to the operator $\sigma(x)$ that creates the asymptotic states. Denoting by F_n the form factors of these operators (for n even they refer to $\mu(x)$ while for n odd to $\sigma(x)$), we have the recursive equation

$$-i \lim_{\tilde{\theta}\to\theta} (\tilde{\theta} - \theta) F_{n+2}(\tilde{\theta} + i\pi, \theta, \theta_1, \theta_2, \ldots, \theta_{2n}) = 2\, F_n(\theta_1, \ldots, \theta_{2n}). \tag{20.9.10}$$

As for any form factor equation, these equations admit an infinite number of solutions that can be obtained by adding all possible kernel solutions at each level. The minimal solution is the one chosen to identify the form factors of the order and disorder operators

$$F_n(\theta_1, \ldots, \theta_n) \;=\; H_n \prod_{i<j}^{n} \tanh \frac{\theta_i - \theta_j}{2}. \tag{20.9.11}$$

The normalization coefficients satisfy the recursive equation

$$H_{n+2} \;=\; i\, H_n.$$

The solutions with n even are therefore fixed by choosing $F_0 = H_0$, namely with a non-zero value of the vacuum expectation of the disorder operator

$$F_0 \;=\; \langle 0|\mu(0)|0\rangle \;=\; \langle \mu \rangle, \tag{20.9.12}$$

while those with n odd are determined by the real constant F_1 relative to the one-particle matrix element of $\sigma(x)$

$$F_1 \;=\; \langle 0|\sigma(0)|A\rangle. \tag{20.9.13}$$

Adopting the conformal normalization of both operators

$$\langle \sigma(x)\sigma(0)\rangle \;=\; \langle \mu(x)\mu(0)\rangle \simeq \frac{1}{|x|^{1/4}}, \qquad |x| \to 0 \tag{20.9.14}$$

it is possible to show that $F_0 = F_1$ and the vacuum expectation value F_0 can be computed using eqn (20.7.2)

$$F_0 \,=\, F_1 \,=\, 2^{1/3}\, e^{-1/4}\, A^3\, m^{1/4}, \tag{20.9.15}$$

where $A = 1.282427\ldots$ is called the Glasher constant. Vice versa, if we choose $F_0 = F_1 = 1$ (as we do hereafter), for the ultraviolet behavior of the correlation functions we have

$$\langle \sigma(x)\sigma(0)\rangle \,=\, \langle \mu(x)\mu(0)\rangle \simeq \frac{2^{-1/3} e^{1/4}\, A^{-3}}{|x|^{1/4}} \,=\, \frac{0.5423804\ldots}{|x|^{1/4}}, \qquad |x| \to 0. \tag{20.9.16}$$

There are several ways to check the correct identification of the form factors of the order/disorder operators. A direct way is to employ the Δ-theorem. In fact, using the matrix elements of $\mu(x)$ and $\Theta(x)$, we can compute their correlator, following the same procedure as in eqn (20.9.6)

$$\begin{aligned} \langle \Theta(r)\mu(0)\rangle &= \frac{1}{2}\int \frac{d\theta_1}{2\pi}\frac{d\theta_2}{2\pi} F^{\Theta}(\theta_{12})\, \bar{F}^{\mu}(\theta_{12})\, e^{-mr(\cosh\theta_1+\cosh\theta_2)} \\ &= -m^2\,\langle\mu\rangle\,\left[\frac{e^{-2mr}}{2mr} + Ei(-2mr)\right] \end{aligned} \tag{20.9.17}$$

where

$$Ei(-x) \,=\, -\int_x^{\infty} \frac{dt}{t}\, e^{-t}.$$

Substituting this correlator in the formula of the Δ-theorem, one obtains the correct value of the conformal weight of the disorder operator

$$\Delta \,=\, -\frac{1}{2\langle\mu\rangle}\int_0^{\infty} dr\, r\langle\Theta(r)\mu(0)\rangle \,=\, \frac{1}{4\pi}\int_0^{\infty} d\theta\, \frac{\sinh^2\theta}{\cosh^3\theta} \,=\, \frac{1}{16}. \tag{20.9.18}$$

Another way to determine the conformal weight of the magnetization operators consists of solving the thermodynamics of the Feynman gas associated to the form factors. Using the nearest-neighbor approximation discussed in Chapter 2, the pressure of this gas satisfies the integral equation (Problem 2)

$$z_c^{-1} \,=\, 2\pi \,=\, \int_0^{\infty} dx\, \tanh^2\frac{x}{2}\, e^{-px}, \tag{20.9.19}$$

whose numerical solution is

$$p \simeq 0.12529\ldots \tag{20.9.20}$$

Comparing with the exact value

$$p = 2\Delta \,=\, \frac{1}{8} \,=\, 0.125, \tag{20.9.21}$$

we see that the relative precision is less than one part in a thousand! This result confirms the validity of the form factor solution for the magnetization operators and, furthemore, it explicitly shows the convergence property of the spectral series.

20.9.3 The Painlevé Equation

The two-point correlation functions of the magnetization operators are given by

$$\begin{aligned} \langle \mu(r)\mu(0)\rangle &= \textstyle\sum_{n=0}^{\infty} g_{2n}(r) \\ \langle \sigma(r)\sigma(0)\rangle &= \textstyle\sum_{n=0}^{\infty} g_{2n+1}(r) \end{aligned}$$

where

$$g_n(r) = \frac{1}{n!}\int \left[\prod_{k=1}^{n} \frac{d\theta_k}{2\pi}\, e^{-mr\cosh\theta_k}\right] \prod_{i<j} \tanh^2 \frac{\theta_{ij}}{2}.$$

These expressions can be further elaborated: imposing $u_i = e^{\theta_i}$ and using

$$\tanh^2 \frac{\theta_i - \theta_j}{2} = \left(\frac{u_i - u_j}{u_i + u_j}\right)^2,$$

we get

$$\prod_{i<j} \tanh^2 \frac{\theta_{ij}}{2} = \prod_{i<j} \left(\frac{u_i - u_j}{u_i + u_j}\right)^2 = \det W, \tag{20.9.22}$$

where the matrix elements of the operator W are

$$W_{ij} = \frac{2\sqrt{u_i\, u_j}}{u_i + u_j}.$$

Combining the two correlators

$$G^{(\pm)}(r) = \langle \mu(r)\mu(0)\rangle \pm \langle \sigma(r)\sigma(0)\rangle = \sum_{n=0}^{\infty} \lambda^n\, g_n(r) \tag{20.9.23}$$

(with $\lambda = \pm 1$) and using (20.9.22) we obtain

$$G^{(\pm)}(r) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int \left[\prod_{k=1}^{n} \frac{d\theta_k}{2\pi}\, e^{-mr\cosh\theta_k}\right] \det W. \tag{20.9.24}$$

The last expression is nothing else but the Fredholm determinant of an integral operator V, whose kernel is

$$\begin{aligned} V(\theta_i, \theta_j, r) &= \frac{E(\theta_i, r)\, E(\theta_j, r)}{u_i + u_j} \\ E(\theta_i, r) &= (2u_i\, e^{-mr\cosh\theta_i})^{1/2}. \end{aligned}$$

Hence

$$G^{(\pm)}(r) = \mathrm{Det}\ (1 + \lambda V). \tag{20.9.25}$$

The remarkable circumstance that the correlation functions are expressed in terms of the Fredholm determinant of an integral operator is crucial for studying their properties. The detailed discussion is beyond the scope of this book and here we simply present the main conclusions.

First of all, the expression given in eqn (20.9.25) permits us to solve *exactly* the thermodynamics of the Feynman gas associated to the form factors of the correlation function $G^{(+)}(r)$. The exact expression of the pressure of the Feynman gas is given by

$$\begin{aligned} p(z) &= \frac{1}{4}\int \frac{dp}{2\pi}\,\log\left[1+\left(\frac{2\pi z}{\sinh \pi p}\right)^2\right] \\ &= \frac{1}{4\pi}\arcsin(2\pi z) - \frac{1}{4\pi^2}\arcsin^2(2\pi z). \end{aligned}$$

Substituting in this formula the plateau value of the fugacity, $z = z_c = 1/(2\pi)$, one obtains the exact value of the conformal weight of the magnetization operators, $p = 2\Delta = 1/8$.

Secondly, using the Fredholm determinant (20.9.25), it is possible to show that the correlators can be concisely written as

$$\begin{pmatrix} \langle\mu(r)\mu(0)\rangle \\ \langle\sigma(r)\sigma(0)\rangle \end{pmatrix} = \begin{pmatrix} \cosh\frac{\Psi(s)}{2} \\ \sinh\frac{\Psi(s)}{2} \end{pmatrix} \exp\left[-\frac{1}{4}\int_s^\infty dt\, t\left[\left(\frac{d\Psi}{dt}\right)^2 - \sinh^2\Psi\right]\right] \tag{20.9.26}$$

($s = mr$), where $\Psi(s)$ is a function solution of the differential equation

$$\frac{d^2\Psi}{ds^2} + \frac{1}{s}\frac{d\Psi}{ds} = 2\sinh(2\Psi), \tag{20.9.27}$$

with boundary conditions

$$\begin{aligned} \Psi(s) &\simeq -\log s + \text{constant}, \quad s \to 0 \\ \Psi(2) &\simeq 2/\pi\, K_0(2s), \quad s \to \infty. \end{aligned} \tag{20.9.28}$$

With the substitution $\eta = e^{-\Psi}$, the differential equation becomes the celebrated Painlevé differential equation of the third kind

$$\frac{\eta''}{\eta} = \left(\frac{\eta'}{\eta}\right)^2 - \frac{1}{s}\left(\frac{\eta'}{\eta}\right) + \eta^2 - \frac{1}{\eta^2}. \tag{20.9.29}$$

This equation was originally obtained by T.T. Wu, B. McCoy, C. Tracy and E. Barouch by studying the scaling limit of the lattice Ising model. It has also been derived by M. Jimbo, T. Miwa, and K. Ueno by using the monodromy theory of differential equations.

20.10 Form Factors of the Sinh–Gordon Model

In this section we study the form factors of an integrable lagrangian theory, the one defined by the Sinh–Gordon model. The action is

$$\mathcal{S} = \int d^2x \left[\frac{1}{2}(\partial_\mu\phi)^2 - \frac{m^2}{g^2}\cosh\, g\phi(x)\right], \tag{20.10.1}$$

and it possesses the Z_2 symmetry $\phi \to -\phi$. The exact S-matrix relative to the particle created by the field $\phi(x)$ is given by

$$S(\theta, B) \;=\; \frac{\tanh \frac{1}{2}(\theta - i\frac{\pi B}{2})}{\tanh \frac{1}{2}(\theta + i\frac{\pi B}{2})}, \tag{20.10.2}$$

where B is a function of the coupling constant g:

$$B(g) \;=\; \frac{2g^2}{8\pi + g^2}. \tag{20.10.3}$$

The theory does not have bound states, therefore the form factors satisfy the recursive equations coming from the kinematic poles only. As we already discussed in Chapter 18, the S-matrix is invariant under the transformation

$$B \to 2 - B \tag{20.10.4}$$

namely, under the weak–strong duality

$$g \to \frac{8\pi}{g}. \tag{20.10.5}$$

The Z_2 symmetry implies that the even (odd) operators have form factors different from zero only on asymptotic states with an even (odd) number of particles. The simplest odd field is just $\phi(x)$, with the normalization given by

$$F_1^{\phi}(\theta) \;=\; \langle 0 \mid \phi(0) \mid \theta \rangle_{\rm in} \;=\; \frac{1}{\sqrt{2}}. \tag{20.10.6}$$

One of the most important fields is the stress–energy tensor

$$T_{\mu\nu}(x) \;=\; 2\pi \,(: \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu} \mathcal{L}(x) \;:) \tag{20.10.7}$$

where $::$ denotes the normal order of the composite operators. Its trace $T^{\mu}_{\mu}(x) = \Theta(x)$ is normalized as

$$F_2^{\Theta}(\theta_{12} = i\pi) \;=\; {}_{\rm out}\langle \theta_1 \mid \Theta(0) \mid \theta_2 \rangle_{\rm in} \;=\; 2\pi m^2, \tag{20.10.8}$$

while F_1^{Θ} is a free parameter. In the following we will only discuss the case $F_1^{\Theta} = 0$: this is equivalent to regarding the Sinh–Gordon model as a deformation of the conformal field theory with central charge $c = 1$ (see Chapter 16 and Problem 1 at the end of the chapter).

20.10.1 Minimal Form Factor

The first step to the solution of the form factor equation consists of finding the minimal two-particle form factor. Expressing the S-matrix as

$$S(\theta) \;=\; \exp\left[8 \int_0^{\infty} \frac{dx}{x} \sinh\left(\frac{xB}{4}\right) \sinh\left(\frac{x}{2}(1 - \frac{B}{2})\right) \sinh\frac{x}{2} \sinh\left(\frac{x\theta}{i\pi}\right)\right].$$

we have

$$F_{\min}(\theta,B) \;=\; \mathcal{N}\,\exp\left[8\int_0^\infty \frac{dx}{x}\,\frac{\sinh\left(\frac{xB}{4}\right)\sinh\left(\frac{x}{2}(1-\frac{B}{2})\right)\,\sinh\frac{x}{2}}{\sinh x}\,\sin^2\left(\frac{x\hat\theta}{2\pi}\right)\right] \tag{20.10.9}$$

($\hat\theta \equiv i\pi - \theta$), with the normalization given by

$$\mathcal{N} \;=\; \exp\left[-4\int_0^\infty \frac{dx}{x}\,\frac{\sinh\left(\frac{xB}{4}\right)\sinh\left(\frac{x}{2}(1-\frac{B}{2})\right)\,\sinh\frac{x}{2}}{\sinh^2 x}\right].$$

The analytic structure of this function can be studied using its representation in terms of an infinite product of Γ functions (see Problem 3)

$$F_{\min}(\theta,B) \;=\; \prod_{k=0}^{\infty}\left|\frac{\Gamma\left(k+\frac{3}{2}+\frac{i\hat\theta}{2\pi}\right)\Gamma\left(k+\frac{1}{2}+\frac{B}{4}+\frac{i\hat\theta}{2\pi}\right)\Gamma\left(k+1-\frac{B}{4}+\frac{i\hat\theta}{2\pi}\right)}{\Gamma\left(k+\frac{1}{2}+\frac{i\hat\theta}{2\pi}\right)\Gamma\left(k+\frac{3}{2}-\frac{B}{4}+\frac{i\hat\theta}{2\pi}\right)\Gamma\left(k+1+\frac{B}{4}+\frac{i\hat\theta}{2\pi}\right)}\right|^2. \tag{20.10.10}$$

$F_{\min}(\theta,B)$ has a simple zero at $\theta=0$ since $S(0)=-1$ and its asymptotic behavior is

$$\lim_{\theta\to\infty} F_{\min}(\theta,B) = 1.$$

It satisfies the functional equation

$$F_{\min}(i\pi+\theta,B)F_{\min}(\theta,B) \;=\; \frac{\sinh\theta}{\sinh\theta+\sinh\frac{i\pi B}{2}} \tag{20.10.11}$$

which can be proved by employing its representation (20.10.10). For the numerical evaluation of this function it is useful to use the mixed representation given by

$$F_{\min}(\theta,B) \;=\; \mathcal{N}\prod_{k=0}^{N-1}\left[\frac{\left(1+\left(\frac{\hat\theta/2\pi}{k+\frac{1}{2}}\right)^2\right)\left(1+\left(\frac{\hat\theta/2\pi}{k+\frac{3}{2}-\frac{B}{4}}\right)^2\right)\left(1+\left(\frac{\hat\theta/2\pi}{k+1+\frac{B}{4}}\right)^2\right)}{\left(1+\left(\frac{\hat\theta/2\pi}{k+\frac{3}{2}}\right)^2\right)\left(1+\left(\frac{\hat\theta/2\pi}{k+\frac{1}{2}+\frac{B}{4}}\right)^2\right)\left(1+\left(\frac{\hat\theta/2\pi}{k+1-\frac{B}{4}}\right)^2\right)}\right]^{k+1}$$
$$\times\exp\left[8\int_0^\infty \frac{dx}{x}\,\frac{\sinh\left(\frac{xB}{4}\right)\sinh\left(\frac{x}{2}(1-\frac{B}{2})\right)\,\sinh\frac{x}{2}}{\sinh^2 x}\,(N+1-N\,e^{-2x})\,e^{-2Nx}\,\sin^2\left(\frac{x\hat\theta}{2\pi}\right)\right].$$

The convergence of the integral in this formula can be improved by increasing the value of N.

20.10.2 Recursive Equations

The Sinh–Gordon model does not have bound states. Hence the only recursive equations come from the kinematical poles relative to the three-particle clusters. Using the identity

$$(p_1+p_2+p_3)^2 - m^2 \;=\; 8m^2\cosh\frac{1}{2}\theta_{12}\cosh\frac{1}{2}\theta_{13}\cosh\frac{1}{2}\theta_{23},$$

all possible poles in these channels are taken into account using the parameterization

$$F_n(\theta_1,\ldots,\theta_n) \;=\; H_n\, Q_n(x_1,\ldots,x_n) \prod_{i<j} \frac{F_{\min}(\theta_{ij})}{x_i + x_j} \tag{20.10.12}$$

where $x_i = e^{\theta_i}$ and H_n are normalization factors. The expression above has simple poles each time the difference of two rapidities θ_{ij} is equal to $i\pi$. The functions $Q_n(x_1,\ldots,x_n)$ are symmetric polynomials in x_i. For the form factors of the scalar operators, the total degree of these polynomials must be equal to that of the denominator, given by $n(n-1)/2$. The partial degree of Q_n depends instead on the asymptotic behavior of the operator $\mathcal{O}$. With the parameterization above, the recursive equations can be expressed as recursive equations for the polynomials Q_n

$$(-)^n\, Q_{n+2}(-x,x,x_1,\ldots,x_n) \;=\; x\mathcal{C}_n(x,x_1,x_2,\ldots,x_n)\, Q_n(x_1,x_2,\ldots,x_n) \tag{20.10.13}$$

where we have introduced the function

$$\mathcal{C}_n = \frac{-i}{4\sin(\pi B/2)} \left(\prod_{i=1}^{n} \left[(x+\omega x_i)(x-\omega^{-1}x_i)\right] - \prod_{i=1}^{n} \left[(x-\omega x_i)(x+\omega^{-1}x_i)\right] \right)$$

with $\omega = \exp(i\pi B/2)$. The normalization constants H_n in (20.10.12) satisfy the recursive equations

$$H_{2n+1} = H_1 \mu^{2n}, \quad H_{2n} = H_2 \mu^{2n-2},$$

with

$$\mu \equiv \left(\frac{4\sin(\pi B/2)}{F_{min}(i\pi, B)} \right)^{\frac{1}{2}}$$

where H_1 and H_2 are the initial conditions, fixed by the operator. Using the generating function of the elementary symmetric polynomials, the function $\mathcal{C}_n$ can be written as

$$\mathcal{C}_n(x,x_1,\ldots,x_n) \;=\; \sum_{k=1}^{n} \sum_{m=1,odd}^{k} [m]\, x^{2(n-k)+m} \sigma_k^{(n)} \sigma_{k-m}^{(n)} (-1)^{k+1} \tag{20.10.14}$$

where we have introduced the symbol $[n]$ defined by

$$[n] \;\equiv\; \frac{\sin(n\frac{B}{2})}{\sin\frac{B}{2}}.$$

Note that the elementary symmetric polynomials satsify the recursive equation

$$\sigma_k^{(n+2)}(-x,x,x_1,\ldots,x_n) \;=\; \sigma_k^{(n)}(x_1,x_2,\ldots,x_n) - x^2 \sigma_{k-2}^{(n)}(x_1,x_2,\ldots,x_n). \tag{20.10.15}$$

20.10.3 General Properties of the Q_n Solutions

The form factors of the derivative operators present a factorized form: for instance, for the operator $\partial\bar{\partial}\phi$ we have $Q_n = \sigma_{n-1}\sigma_1\tilde{Q}_n$. For this reason, it is convenient to focus attention on the so-called *irreducible operators*, whose form factors cannot be factorized, and use them as building blocks for the form factors of all other operators. The polynomials Q_n of the irreducible operators satisfy a series of interesting results coming from the recursive equations (20.10.13). Let's initially show that the partial degree of Q_n satisfies the inequality

$$\deg(Q_n) \leq n - 1. \tag{20.10.16}$$

It is easy to see that this result holds for Q_1 and Q_2. To show that it also holds for the higher polynomials, let us consider the two cases (a) $Q_n \neq 0$ and (b) $Q_n = 0$ separately.

- In case (a) the proof is by induction. Assume $\deg(Q_n) \leq n - 1$. Since $\mathcal{C}_n$ is bilinear in $\sigma^{(n)}$ (see eqn 20.10.14), the partial degree of $Q_{n+2}(-x, x, x_1, \dots x_n)$ in the variables $x_1, \dots x_n$ is less than or equal to $n + 1$. But the partial degree of $Q_{n+2}(x_1, x_2, \dots x_{n+2})$ is equal to the partial degree of $Q_{n+2}(-x, x, x_1 \dots, x_n)$, hence the partial degree of Q_{n+2} must be less than or equal to $n + 1$.
- In case (b), the space of the solutions is given by the kernel of the operator $\mathcal{C}$, namely

  $$Q_{n+2}(-x, x, \dots, x_{n+2}) = 0.$$

  In the space of the polynomials $\mathcal{P}$ of total degree $(n + 2)(n + 1)/2$, there is only one solution of this equation, given by

  $$Q_{n+2} = \prod_{i<j}^{n+2}(x_i + x_j). \tag{20.10.17}$$

  This polynomial has partial degree $n + 1$ and coincides with the polynomial of the denominator of eqn (20.10.12).

We have thus shown that the partial degree of Q_n must be less than or equal to $(n - 1)$ for any irreducible scalar operator. The first consequence is that the form factors of these operators cannot diverge when $\theta_i \to \infty$. The second consequence is the presence of an additional parameter at each step of the iterative procedure. This comes from a simple argument: the dimension of the space of the polynomials Q_n is given by the dimension of the space of the polynomials Q_{n-2} *plus* the dimension of the kernel. Since the kernel is one dimensional, the dimension of the space of the solutions increases exactly by one at each iterative step. With the initial conditions $\dim(Q_1) = \dim(Q_2) = 1$, we finally get

$$\dim(Q_{2n-1}) = \dim(Q_{2n}) = n. \tag{20.10.18}$$

Hence the most general form factor of an irreducible scalar operator belongs to a linear space that can be spanned by a basis Q_n^k:

$$Q_{2n}(A_1^{(2n)},\dots,A_n^{(2n)}) = \sum_{p=1}^{n} A_p^{(2n)} Q_{2n}^p \tag{20.10.19}$$

$$Q_{2n-1}(A_1^{(2n-1)},\dots,A_n^{(2n-1)}) = \sum_{p=1}^{n} A_p^{(2n-1)} Q_{2n-1}^p.$$

Each polynomial above defines a matrix element of an operator of the Sinh–Gordon model. Note that the dimension of this linear space grows exactly as the number of powers ϕ^k $(k < n)$ of the elementary field. This means that the matrix elements of the composite operators ϕ^k can be obtained as linear combinations of the above functions.

20.10.4 The Elementary Solutions

A remarkable class of solutions of the recursive equations (20.10.13) is given by[5]

$$Q_n(k) \;=\; ||M_{ij}(k)||, \tag{20.10.20}$$

where $M_{ij}(k)$ is the $(n-1)\times(n-1)$ matrix

$$M_{ij}(k) \;=\; \sigma_{2i-j}\,[i-j+k]. \tag{20.10.21}$$

and $||M||$ denotes the determinant of the matrix M. These polynomials are called *elementary solutions*: they depend on an arbitrary integer k and satisfy

$$Q_n(k) \;=\; (-1)^{n+1} Q_n(-k). \tag{20.10.22}$$

Although all $Q_n(k)$ are solutions of (20.10.13), not all of them are linearly independent. The simplest reason is that the dimension of the space of the solutions at the level $N = 2n$ (or $N = 2n-1$) is at most n. Among the first representatives we have

$$Q_3(k) \;=\; \left|\left| \begin{matrix} [k]\sigma_1 & [k+1]\sigma_3 \\ \text{[k-1]} & [k]\sigma_2 \end{matrix} \right|\right|.$$

Using the trigonometric identity $[n]^2 - [n-1][n+1] = 1$, it is easy to see that this expression satisfies eqn (20.10.13) (with $A_0^1 = 1$) for any integer k. These solutions allow us to express at once all the form factors of the elementary field $\phi(x)$ and the trace $\Theta(x)$ of the stress–energy tensor. In fact, it is possible to prove that the matrix elements of $\phi(x)$ are given by $Q_n(0)$. Note that the form factors relative to an even number of particles are automatically zero, in agreement with the Z_2 symmetry of the model. Those with an odd number of asymptotic particles vanish when $\theta_i \to \infty$, in agreement with the perturbative evaluation of these matrix elements given by the Feynman diagrams. The form factors of $\Theta(x)$ are instead given by the even polynomials $Q_{2n}(1)$, which go to a finite limit when $\theta_i \to \infty$, once again in agreement with their

[5]For simplicity we have suppressed the dependence of $Q_n(k)$ on the variables x_i.

Table 20.1: *Approximate values of the central charge of the Sinh–Gordon model obtained by using only the two-particle form factor of $\Theta(x)$ in the c-theorem.*

B	$\frac{g^2}{4\pi}$	$\Delta\, c^{(2)}$
$\frac{1}{10}$	$\frac{2}{19}$	0.9989538
$\frac{3}{10}$	$\frac{6}{17}$	0.9931954
$\frac{2}{5}$	$\frac{1}{2}$	0.9897087
$\frac{1}{2}$	$\frac{2}{3}$	0.9863354
$\frac{2}{3}$	1	0.9815944
$\frac{7}{10}$	$\frac{14}{13}$	0.9808312
$\frac{4}{5}$	$\frac{4}{3}$	0.9789824
1	2	0.9774634

perturbative computation. A further confirmation of the validity of this identification can be obtained by using the c-theorem. Employing just the two-particle form factor, we have the following approximated value of the ultraviolet central charge

$$c^{(2)} \;=\; \frac{3}{2F^2_{min}(i\pi)} \int_0^\infty \frac{d\theta}{\cosh^4\theta}\, |F_{min}(2\theta)|^2. \tag{20.10.23}$$

The numerical values for different values of the coupling constant $g^2/4\pi$ are collected in Table 20.1. From this table one can see that the sum rule is saturared by the two-particle form factor even for large values of the coupling constant: this proves once again the fast convergent behavior of the spectral series.

It is interesting to understand which are the operators $\Psi_k(x)$ associated to the elementary solutions $Q_n(k)$ $(k \neq 0)$. For the sequence of form factors related to $Q_n(k)$, let's choose the normalization as follows

$$H^k_1 = \mu[k], \quad H^k_2 = \mu^2\,[k]. \tag{20.10.24}$$

The present conjecture is that the operators Ψ_k correspond to the vertex operators $e^{kg\phi}$. A non-trivial check of this conjecture is provided by the computation of the conformal weights $\Delta_k(g)$ that emerge in their ultraviolet limits. These quantities can be computed by analyzing the limit $x \to 0$ of the correlation function

$$\begin{aligned} G_{k,m}(x) \;&=\; \langle \Psi_k(x)\, \Psi_m(0)\rangle \\ &= \sum_{n=0}^{\infty} \int \frac{d\beta_1 \dots d\beta_n}{n!(2\pi)^n} F_n^{\Psi_k}(\beta_1 \dots \beta_n) F_n^{\Psi_m}(\beta_n \dots \beta_1) \exp\left(-mr \sum_{i=1}^n \cosh\beta_i\right). \end{aligned}$$

At first order in g, we have $\Delta_k(g) = -k^2 g^2/8\pi$ which coincides with the conformal weight of the vertex operators $e^{k\,g\,\phi(x)}$, computed using the gaussian conformal theory.

20.11 The Ising Model in a Magnetic Field

The Ising model in a magnetic field has quite a rich S-matrix: it has eight massive exitations and 36 elastic scattering amplitudes, some of them with higher order poles.

In addition to the functional and recursive equations, the form factors of this theory also satisfy other recursive equations related to the higher poles of the S-matrix. The relative formulas can be found in the papers by G. Delfino, G. Mussardo, and P. Simonetti listed at the end of the chapter. Here we only report the main results about the form factors of the energy operator $\epsilon(x)$ and of the magnetization operator $\sigma(x)$. In this theory, the latter operator is proportional to the trace

$$\Theta(x) = 2\pi h(2 - 2\Delta_\sigma)\,\sigma(x). \tag{20.11.1}$$

Relying on the fast convergence of the spectral series, for the correlation functions of these operators we can focus our attention on the one and two-particle form factors. To begin with, let's fix some notation. For the S-matrix of the particles A_a and A_b we have

$$S_{ab}(\theta) = \prod_{\alpha \in \mathcal{A}_{ab}} (f_\alpha(\theta))^{p_\alpha} \tag{20.11.2}$$

where

$$f_\alpha(\theta) \equiv \frac{\tanh\frac{1}{2}(\theta + i\pi\alpha)}{\tanh\frac{1}{2}(\theta - i\pi\alpha)}. \tag{20.11.3}$$

The set of the numbers $\mathcal{A}_{ab}$ and their multiplicity p_α can be found in Table 18.3 of Chapter 18. It is convenient to parameterize the two-particle form factors of this theory as

$$F_{ab}^{\mathcal{O}}(\theta) \;=\; \frac{Q_{ab}^{\Phi}(\theta)}{D_{ab}(\theta)} F_{ab}^{min}(\theta), \tag{20.11.4}$$

where $D_{ab}(\theta)$ and $Q_{ab}^{\mathcal{O}}(\theta)$ are polynomials in $\cosh\theta$: the latter is fixed by the singularities of the S-matrix, the former depends on the operator $\mathcal{O}(x)$. The minimal form factors can be written as

$$F_{ab}^{min}(\theta) \;=\; \left(-i\sinh\frac{\theta}{2}\right)^{\delta_{ab}} \prod_{\alpha \in \mathcal{A}_{ab}} (G_\alpha(\theta))^{p_\alpha}\,, \tag{20.11.5}$$

where

$$G_\alpha(\theta) \;=\; \exp\left\{2\int_0^\infty \frac{dt}{t}\frac{\cosh\left(\alpha - \frac{1}{2}\right)t}{\cosh\frac{t}{2}\sinh t}\sin^2\frac{(i\pi - \theta)t}{2\pi}\right\}. \tag{20.11.6}$$

For large values of the rapidity, we have

$$G_\alpha(\theta) \;\sim\; \exp(|\theta|/2), |\theta| \;\to\; \infty, \tag{20.11.7}$$

independently of the index α.

From the analysis of the singularities of the form factors, one can arrive at the following expression for the denominator

$$D_{ab}(\theta) \;=\; \prod_{\alpha \in \mathcal{A}_{ab}} (\mathcal{P}_\alpha(\theta))^{i_\alpha} (\mathcal{P}_{1-\alpha}(\theta))^{j_\alpha}\,, \tag{20.11.8}$$

where

$$\begin{array}{lll} i_\alpha = n+1,\, j_\alpha = n, \text{ se} & & p_\alpha = 2n+1; \\ i_\alpha = n, & j_\alpha = n, \text{ se} & p_\alpha = 2n, \end{array} \tag{20.11.9}$$

having introduced the notation

$$\mathcal{P}_\alpha(\theta) \equiv \frac{\cos \pi\alpha - \cosh\theta}{2\cos^2 \frac{\pi\alpha}{2}}. \tag{20.11.10}$$

Both quantities $F_{ab}^{min}(\theta)$ and $D_{ab}(\theta)$ are normalized to be equal to 1 when $\theta = i\pi$.

The polynomials of the numerator can be expressed as

$$Q_{ab}^{\mathcal{O}}(\theta) = \sum_{k=0}^{N_{ab}^{\mathcal{O}}} c_{ab,\mathcal{O}}^{(k)} \cosh^k \theta. \tag{20.11.11}$$

The condition $\left[F_{ab}^{\mathcal{O}}(\theta)\right]^* = F_{ab}^{\mathcal{O}}(-\theta)$ follows from the monodromy condition satisfied by the form factors and from the property $S_{ab}^*(\theta) = S_{ab}(-\theta)$. This means that the coefficients $c_{ab,\mathcal{O}}^{(k)}$ are real numbers and their values identify the different operators.

The degrees of the polynomials are fixed by the conformal weight of the operators and, for both $\sigma(x)$ and $\epsilon(x)$, we have in particular $N_{11}^{\Phi} \leq 1$. Therefore the initial conditions of the recursive equation for the form factors of the two relevant operators consists of *two* free parameters, i.e. the coefficients $c_{11,\mathcal{O}}^{(0)}$ and $c_{11,\mathcal{O}}^{(1)}$. Furthemore, it can be checked that the number of free parameters does not increase implementation the bootstrap equations. Consider, for instance the condition $N_{12}^{\mathcal{O}} \leq 2$, which seems to imply three new coefficients $c_{12,\mathcal{O}}^{(k)}$ $(k = 1, 2, 3)$ for $F_{12}^{\mathcal{O}}(\theta)$. However, the amplitudes $S_{11}(\theta)$ and $S_{12}(\theta)$ have three common bound states. This circumstance gives rise to three equations

$$\frac{1}{\Gamma_{11}^c} \text{Res}_{\theta = iu_{11}^c} F_{11}^{\Phi}(\theta) = \frac{1}{\Gamma_{12}^c} \text{Res}_{\theta = iu_{12}^c} F_{12}^{\Phi}(\theta), \qquad c = 1, 2, 3$$

that permit us to fix the three coefficients $c_{12,\mathcal{O}}^{(k)}$ in terms of the two coefficients $c_{11,\mathcal{O}}^{(k)}$.

Table 20.2: *Central charge given by the partial sum of the form factors entering the c-theorem. $c_{ab..}$ denotes the contribution of the state $A_a A_b ..$. The exact result is $c = 1/2$.*

c_1	0.472038282
c_2	0.019231268
c_3	0.002557246
c_{11}	0.003919717
c_4	0.000700348
c_{12}	0.000974265
c_5	0.000054754
c_{13}	0.000154186
c_{partial}	0.499630066

Table 20.3: *Conformal weights $\Delta_{\mathcal{O}}$ given by the partial sum of the form factors of the correlation functions entering the Δ-theorem. $\Delta_{ab..}$ denotes the contribution of the state $A_a A_b ..$. The exact values are $\Delta_\sigma = 1/16 = 0.0625$ and $\Delta_\varepsilon = 1/2$.*

	σ	ϵ
Δ_1	0.0507107	0.2932796
Δ_2	0.0054088	0.0546562
Δ_3	0.0010868	0.0138858
Δ_{11}	0.0025274	0.0425125
Δ_4	0.0004351	0.0069134
Δ_{12}	0.0010446	0.0245129
Δ_5	0.0000514	0.0010340
Δ_{13}	0.0002283	0.0065067
Δ_{partial}	0.0614934	0.4433015

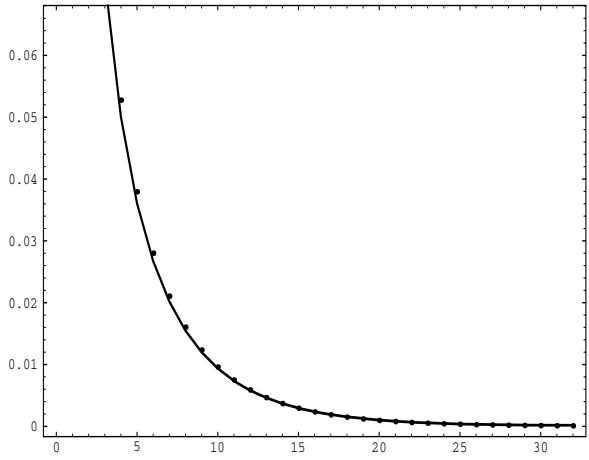


Fig. 20.12 *Plot of the correlation function $\langle\sigma(r)\sigma(0)\rangle$ for the Ising model in a magnetic field. The continuous line is the determination obtained with the first eight form factors, while the dots are the numerical determination of the correlators obtained by a Monte Carlo simulation.*

There is additional information about the numerator Q_{ab} of the operator $\Theta(x)$. In fact, from the conservation law $\partial_\mu T^{\mu\nu} = 0$ it follows that the polynomials Q^{Θ}_{ab} contain the factor

$$\left(\cosh\theta + \frac{m_a^2 + m_b^2}{2m_a m_b}\right)^{1-\delta_{ab}} . \tag{20.11.12}$$

The determination of the coefficients $c^{(k)}_{ab}$ and the one-particle form factors of the two operators $\sigma \sim \Theta$ and ϵ has been done in the papers cited at the end of the chapter and their values can be found there.

Employing these lowest form factors one can compute the correlation functions and perform some non-trivial checks by applying the sum rules of the c-theorem and Δ-theorem. The relative results are given in Tables 20.2 and 20.3. A successful check of the correlation function $\langle\sigma(r)\sigma(0)\rangle$ has also been done versus the numerical determination of this function, as shown in Fig. 20.12.

References and Further Reading

The form factor program has been proposed by:

M. Karowski, P. Weisz, *Exact form-factors in (1+1)-dimensional field theoretic models with soliton behavior*, Nucl. Phys. B 139 (1978), 455.

F. Smirnov, *Form Factors in Completely Integrable Models of Quantum Field Theory*, World Scientific, Singapore, 1992.

The form factors of the thermal Ising model have been studied in detail in the article:

V.P. Yurov, Al.B. Zamolodchikov, *Correlation functions of integrable 2-D models of relativistic field theory. Ising model*, Int. J. Mod. Phys. A6 (1991), 3419

The determination of the correlators of the Ising model on the lattice and their relation with the Painlevé equation can be found in the articles:

T.T. Wu, B. McCoy, C. Tracy, E. Barouch, *Spin-spin correlation functions for the two-dimensional Ising model: Exact theory in the scaling region*, Phys. Rev. B 13 (1976), 316.

B. McCoy, C. Tracy, T.T. Wu, *Painlevé functions of the third kind*, J. Math. Phys. 18 (1977), 1058.

M. Jimbo, T. Miwa, K. Ueno, *Monodromy preserving deformation of linear ordinary differential equations with rational coefficients: I. General theory and τ-function*, Physica D2, (1981), 306.

The classification of the operators obtained in terms of the different solutions of the form factor equations has been proposed in the paper:

J. Cardy, G. Mussardo *Form-factors of descendent operators in perturbed conformal field theories* Nucl. Phys. B 340 (1990), 387.

The form factors of the Sinh–Gordon model have been determined in the papers:

A. Fring, G. Mussardo, P. Simonetti, *Form-factors for integrable Lagrangian field theories, the sinh–Gordon theory*, Nucl. Phys. B 393 (1993), 413.

A. Koubek, G. Mussardo, *On the operator content of the Sinh–Gordon model*, Phys. Lett. B 311 (1993), 193.

The matrix elements of the Ising model in a magnetic field have been computed in:

G. Delfino, G. Mussardo, *The spin spin correlation function in the two-dimensional Ising model in a magnetic field at $T = T_c$*, Nucl. Phys. B 455 (1995), 724.

G. Delfino, P. Simonetti, *Correlation functions in the two-dimensional Ising model in a magnetic field at $T = T_c$*, Phys. Lett. B 383 (1996), 450.

G. Delfino, *Integrable field theory and critical phenomena. The Ising model in a magnetic field*, J. Phys. A 37 (2004), R45.

The convergence property of the spectral series has been analyzed by several authors. It is useful to consult the papers:

J. Cardy, G. Mussardo, *Universal properties of self-avoiding walks from two-dimensional field theory*, Nucl. Phys. B 410 (1993), 451.

Al.B. Zamolodchikov, *Two point correlation function in scaling Lee–Yang model*, Nucl. Phys. B 348 (1991), 619.

The vacuum expectation values of several operators of integrable field theories have been determined in a remarkable series of papers:

S. Lukyanov, A.B. Zamolodchikov, *Exact expectation values of local field in quantum sine–Gordon model*, Nucl. Phys. B 493 (1997), 571.

V. Fateev, S. Lukyanov, A.B. Zamolodchikov, Al.B. Zamolodchikov, *Expectation values of local fields in Bullogh–Dodd model and integrable perturbed conformal theories*, Nucl. Phys. B 516 (1998), 652.

V. Fateev, S. Lukyanov, A.B. Zamolodchikov, Al.B. Zamolodchikov, *Expectation values of descendent fields in the Sine–Gordon model*, Nucl. Phys. B 540 (1999), 587.

Form factors have found interesting applications in problems of condensed matter physics, see:

F. Essler, R. Konik, *Applications of massive integrable quantum field theories to problems in condensed matter physics*, in "From Fields to Strings: Circumnavigating Theoretical Physics", World Scientific, Singapore, 2005.

Problems

1. Form factors of a free theory

Consider the theory of a free bosonic field $\phi(x)$ associated to a particle A of mass m.

a Compute the form factors of $\phi(x)$ and prove that $\langle 0|\phi(0)|A\rangle = 1/\sqrt{2}$. Show that the euclidean correlation function is given by

$$\langle \phi(x)\phi(0)\rangle \;=\; \frac{1}{\pi}\, K_0(mr).$$

b Show that the arbitrariness of the one-particle form factor of the trace of the stress–energy tensor

$$F_1^{\Theta} = \langle 0|\Theta(0)|A\rangle \equiv -\sqrt{2\pi}\, m^2\, Q$$

corresponds to the possibility of redefining the stress–energy tensor by adding a total divergence

$$\Theta(x) = 2\pi \left(m^2\phi^2 + \frac{Q}{\sqrt{\pi}} \Box\phi \right).$$

c Use the c-theorem and the form factors of $\Theta(x)$ to show that the central charge in the ultraviolet region is given by

$$c = 1 + 12Q^2.$$

2. Feynman gas

a Derive the equation of state of the Feynman gas associated to the form factors of the magnetization operators in the nearest neighbor approximation. Prove that the pressure $p(z)$ satisfies the integral equation (20.9.19).

b Justify the accuracy of the approximation of the conformal weights computing the average number of particles per unit length by means of the formula

$$\frac{\langle N\rangle}{L} = z\,\frac{\partial p}{\partial z}$$

and checking the very dilute nature of the gas.

3. Infinite products

Using the integral

$$\int \frac{dt}{t} e^{-\beta t} \sin^2 \frac{\alpha t}{2} = \frac{1}{4} \log \frac{\alpha^2+\beta^2}{\beta^2},$$

and the identity satisfied by the Γ functions

$$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\gamma)\Gamma(\beta-\gamma)} = \prod_{k=0}^{\infty} \left[\left(1+\frac{\gamma}{\alpha+k}\right) \left(1-\frac{\gamma}{\beta+k}\right) \right],$$

to derive the expression for $F_{\min}(\theta)$ of the Sinh–Gordon model.

4. Cluster properties

Consider the form factors of a scattering theory based on the functions

$$f_x(\theta) = \frac{\tanh \frac{1}{2}(\theta + i\pi x)}{\tanh \frac{1}{2}(\theta - i\pi x)}$$

that have the property $\lim_{\theta\to\infty} f_x(\theta) = 1$.

a Using the Watson equation satisfied by the form factors $F_n^{\mathcal{O}_a}(\theta_1, \ldots, \theta_n)$ of an operator $\mathcal{O}_a$, prove that taking the limit

$$\lim_{\Lambda \to \infty} F_n^{\mathcal{O}_a}(\beta_1 + \Delta, \ldots, \beta_m + \Delta, \beta_{m+1}, \ldots, \beta_n) = F_m^{\mathcal{O}_b}(\beta_1, \ldots, \beta_m) F_{n-m}^{\mathcal{O}_c}(\beta_{m+1}, \ldots, \beta_n)$$

the form factor factorizes in terms of two functions both satisfying the Watson equations. Hence they can be considered the form factors of the operators $\mathcal{O}_b$ and $\mathcal{O}_c$. This expresses the cluster property of the form factors.

b Prove that the form factors of the elementary solutions of the Sinh–Gordon model are self-clustering quantities.

5. Correlation functions of the Ising model

Use the fermionic representation of the energy operator of the Ising model, $\epsilon = i\bar{\psi}\psi$, and the mode expansion of the fermionic field in terms of the creation and annihilation operators, to compute the matrix elements of $\epsilon(x)$ and its two-point correlation function.

6. Form factors of the Yang–Lee model

Using the form factors of the Sinh–Gordon model, obtain the form factors of the Yang–Lee model by using the analytic continuation $B \to -\frac{2}{3}$.

21
Non-Integrable Aspects

Nobody is perfect!

Billy Wilder

The integrable quantum field theories analyzed in the previous chapters provide the exact solution of many statistical models away from the critical point. Despite the elegance and the undeniable success of these methods, the generic situation that occurs in statistical physics is that of non-integrable dynamics: many interesting statistical models fall within this class and therefore it would be highly desirable to develop an appropriate formalism to deal with the lack of integrability. This task is notoriously difficult for the rich phenomenology that arises when the dynamics are not integrable: there are decays and production scattering processes, confinement phenomena and nucleation of false vacua, resonance peaks in the cross-sections, etc. All these physical aspects are usually accompanied by great mathematical complexity. To see this, it is sufficient to consider the analytic structure of the S-matrix of such theories: once one has given up the integrability condition, the infinite number of thresholds of the production processes gives rise to nested patterns of branch cut singularities, in addition to the pole structure associated to the bound states or the resonances (see Fig. 21.1).

Although the detailed analysis of all physical aspects of non-integrable models goes beyond the scope of this book, in this chapter we will discuss a series of results particularly helpful to understand the class of universality of some important models. Our study takes advantage of a perturbative approach based on the exact form factors of the integrable models discussed in the previous chapter.

21.1 Multiple Deformations of the Conformal Field Theories

We focus our attention on a particular class of non-integrable models. They can be defined in terms of a conformal action deformed by two relevant operators, each of them giving rise *individually* to an integrable model

$$\mathcal{S} = \mathcal{S}_{CFT} + \lambda_1 \int d^2x \varphi_1(x) + \lambda_2 \int d^2x \varphi_2(x). \tag{21.1.1}$$

There are many interesting physical systems that belong to this class of non-integrable model. Let's briefly discuss two of them.

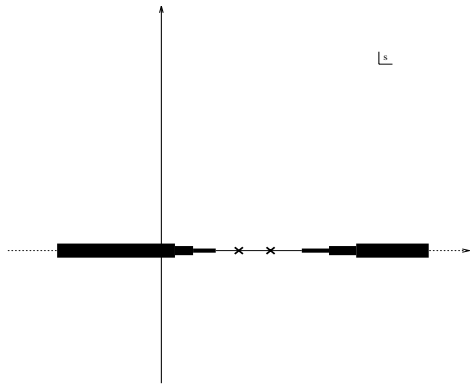


Fig. 21.1 *Analytic structure of the S-matrix of non-integrable models. At the production thresholds, the scattering amplitudes develop new branch cuts.*

- The first is the Ising model at temperature T different from the critical value T_c and in an external magnetic field h. Its action is given by

$$\mathcal{S} = \mathcal{S}_{CFT} + \tau \int d^2x \varepsilon(x) + h \int d^2x \sigma(x), \tag{21.1.2}$$

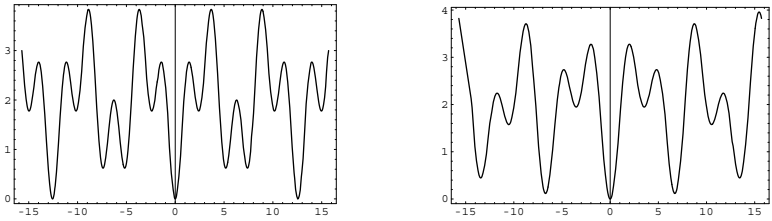
  with $\tau = T - T_c$. When $h = 0$, the above action corresponds to the integrable theory of the thermal deformation, which has only one particle excitation and an elastic S-matrix equal to $S = -1$. On the contrary, when $T = T_c$ one recovers the integrable theory of the Ising in a magnetic field: its spectrum consists of eight massive particles, whose S-matrix amplitudes were discussed in Chapter 18.
- The second example is provided by the multifrequency Sine–Gordon model. The action is given by

$$\mathcal{S} = \int d^2x \left[\frac{1}{2} (\partial_\mu \varphi)^2 + \lambda_1 \cos \beta\varphi + \lambda_2 \cos \alpha\varphi \right]. \tag{21.1.3}$$

  When $\lambda_1 = 0$, this action gives rise to the integrable theory of the Sine–Gordon model with frequency α. In addition to the soliton states, such a theory has a number of neutral bound states given by[1] $N_2 = \left[\frac{\pi}{\xi_\alpha} \right]$, where $\xi_\alpha = \frac{\alpha^2}{8} / (1 - \alpha^2/8\pi)$. Viceversa, if $\lambda_2 = 0$, we again have a Sine–Gordon model but of frequency β and a different number of neutral bound states, $N_1 = \left[\frac{\pi}{\xi_\beta} \right]$. When the ratio of α and β is a rational number, the potential of the theory has an infinite number of periodic and degenerate vacua. On the contrary, when the ratio of the frequencies is an irrational number, the potential has only one vacuum that can always be placed at the origin (see Fig. 21.2).

To study the field theories associated to an action such as (21.1.1), it is convenient to regard it as *a deformation of an integrable action* rather than as a multiple deformation

[1] $[x]$ denotes the integer part of the real number x.



(a) **(b)**

Fig. 21.2 *Potential of the multiple-frequency Sine–Gordon model: rational (a) and irrational (b) ratio of the frequencies.*

of a conformal theory. By taking this point of view and grouping the terms differently, the action (21.1.1) can be written as

$$\mathcal{S} = \mathcal{S}^i_{\rm int} + \lambda_j \int d^2x \varphi_j(x). \tag{21.1.4}$$

$(i = 1, 2,\ j \neq i)$. There are several advantages in doing so.

1. The first convenience becomes evident by going to Minkoswki space. In fact, the non-integrable theory can be analyzed starting from the basis of the Hilbert space provided by the particle excitations associated to the integrable model $\mathcal{S}^i_{\rm int}$. Although the spectra of $\mathcal{S}$ and $\mathcal{S}^i_{int}$ may be different, the basis provided by the particles of the integrable model will certainly be more appropriate than the conformal one, as far as the infrared properties of the non-integrable model are concerned.
2. The second advantage consists of the exact solvability of the integrable models, in particular, the possibility of computing exactly all the matrix elements (form factors) of local and non-local operators of such theories. Hence, in complete analogy with ordinary quantum mechanics, one can set up a perturbative approach based on the form factors of the integrable models. As we will see in the following, this perturbative approach will enable us to reach a remarkable series of predictions about the mass correction, the decay processes, or the correction to the scattering amplitudes.
3. When each deformation is individually integrable, there is the obvious freedom of using any of them as a starting point. By this choice we select a particular basis of the particles and bound states thereof. However, since the actual dynamics of the model should be insensitive to such a choice, there should be a series of mathematical identities that links one perturbative series to the other.

21.2 Form Factor Perturbation Theory

Let us consider a quantum field theory in Minkowski space, defined by the action

$$\mathcal{S} = \mathcal{S}_0 + \mathcal{S}_I = \mathcal{S}_0 - \lambda \int d^2x \Psi(x), \tag{21.2.1}$$

where $\mathcal{S}_0$ denotes here the minkowskian action of the unperturbed theory and Ψ one of its operators. We suppose that the QFT associated to the action $\mathcal{S}_0$ is exactly solvable (although not necessarily free), so that the spectrum of its particles, their scattering amplitudes, and the matrix elements of the its operators (in particular those of Ψ) are assumed to be all known. For the sake of simplicity, in this section we consider the case of an isospectral perturbation of a solvable theory. By this we mean that the spectrum of the total action $\mathcal{S}$ is made of the same number of particles of the unperturbed one $\mathcal{S}_0$: the new interaction $\mathcal{S}_I$ is going to change the values of the masses of the physical particles but not their stability properties.[2]

Let us now describe the properties of the theory (21.2.1). Under the hypothesis that the interaction term is turned off at $t \to \pm\infty$, it is possible to adopt the formalism of the asymptotic "*in*" and "*out*" states. We are interested in computing the scattering amplitude

$$\begin{aligned} S\{q_1, \dots, q_n \to q'_1, \dots, q'_m\} &= {}^{\text{out}}\langle q'_1, \dots, q'_m | q_1, \dots, q_n \rangle^{\text{in}} \\ &= {}^{\text{in}}\langle q'_1, \dots, q'_m | S | q_1, \dots, q_n \rangle^{\text{in}}, \end{aligned} \tag{21.2.2}$$

where q_i and q'_j label the momenta of the *in*-going and *out*-going set of particles. Since in the remote past $t \to -\infty$ the interaction is not present yet, the asymptotic *in* states coincide with the unperturbed ones:

$$|q_1, \dots, q_n\rangle^{\text{in}} = |q_1, \dots, q_n\rangle_0^{\text{in}}. \tag{21.2.3}$$

As usual, the scattering operator S in eqn (21.2.2) can be obtained as the limit

$$S = \lim_{t \to +\infty} U(t, -t) \tag{21.2.4}$$

of the time evolution operator $U(t, t_0)$, the solution of the equation

$$i\frac{d}{dt}U(t, t_0) = HU(t, t_0), \qquad U(t_0, t_0) = 1, \tag{21.2.5}$$

where $H = H_0 + H_I$ denotes the Hamiltonian of the theory (21.2.1).

Interactive representation. Following the standard quantum mechanical procedure,[3] the operator U can be factorized as $U = U_0 U_I$, where U_0 and U_I are the solutions of eqn (21.2.5) with H replaced by H_0 and $\tilde{H}_I(t) = U_0^{-1} H_I U_0$, respectively. Then, we can write the scattering operator of the theory (21.2.1) as $S = S_0 S_I$, where $S_0 = \lim_{t \to +\infty} U_0(t, -t)$ is the unperturbed and exactly known scattering matrix

$$\begin{aligned} S_0\{q_1, \dots, q_n \to q'_1, \dots, q'_m\} &= {}_0^{\text{out}}\langle q'_1, \dots, q'_m | q_1, \dots, q_n \rangle_0 \\ &= {}_0^{\text{in}}\langle q'_1, \dots, q'_m | S_0 | q_1, \dots, q_n \rangle_0^{\text{in}}, \end{aligned} \tag{21.2.6}$$

[2] We will comment later on the more general case.

[3] See, for instance, C. Cohen Tannoudji, B. Liu, F. Laloe, *Quantum Mechanics*, John Wiley, New York, 1977.

while S_I has the usual formal representation

$$S_I = \lim_{t\to+\infty} U_I(t,-t) = T\exp\left(iA_I[\Psi]\right). \tag{21.2.7}$$

The scattering amplitude is therefore given by

$$\begin{aligned}
{}^{\rm out}\langle q'_1,\dots,q'_m|q_1,\dots,q_n\rangle^{\rm in} &= {}^{\rm in}_0\langle q'_1,\dots,q'_m|T\exp\left(-i\lambda\int d^2x\Psi(x)\right)|q_1,\dots,q_n\rangle^{\rm in}_0 \\
&= \sum_{k=0}^{+\infty}\frac{(-i\lambda)^k}{k!}\int d^2x_1\dots d^2x_k\,{}^{\rm in}_0 \\
&\quad\times\langle q'_1,\dots,q'_m|T\left(\Psi(x_1)\dots\Psi(x_k)\right)|q_1,\dots q_n\rangle^{\rm in}_0,
\end{aligned} \tag{21.2.8}$$

where (21.2.6) has been used in order to absorb the factor S_0.

In the ordinary lagrangian perturbation scheme based on free theories, the computation of scattering amplitudes would now proceed through the use of creation/annihilation operators and their employment in Wick's theorem, finally leading to the diagrammatic expansion that is characteristic of the Feynman covariant perturbation theory. This approach, however, cannot be generally followed here because we might not know, in general, a local lagrangian formulation of the theory associated to $\mathcal{S}_0$. However, the exact solution of its dynamics, in the form specified at the beginning of this section, naturally suggests computing the scattering amplitudes (21.2.8) within the same framework used to deal with ordinary time-dependent perturbation theory in quantum mechanics. In other words, let's initially insert between the operators $\Psi(x_l)$ and $\Psi(x_{l+1})$ $(l=1,\dots,k-1)$ in the second line of (21.2.8) a sum over a complete set of asymptotic states of the unperturbed theory

$$\sum_n |n\rangle_{\rm out\,out}\langle n| = 1 = \sum_n |n\rangle_{\rm in\,in}\langle n|, \tag{21.2.9}$$

with $|n\rangle$ denoting an asymptotic state containing n on–shell particles. The integrations over the space coordinates in (21.2.8) can be immediately performed: they lead to delta functions that constraint the total momentum of the intermediate states to coincide with that of the initial and final states. In doing the integrations over the time variables, the time ordering prescription gives rise in this case to the appearence of energy denominators. The final expression is

$$\begin{aligned}
{}^{\rm out}\langle q'_1,\dots,q'_m|q_1,\dots,q_n\rangle^{\rm in} &= {}^{\rm in}_0\langle q'_1,\dots,q'_m|q_1,\dots,q_n\rangle^{\rm in}_0 \\
&\quad+(2\pi)^2\delta^{(2)}\left(\sum_{j=1}^m q'_j-\sum_{j=1}^n q_j\right) \\
&\quad\times\left\{-i\lambda\,{}^{\rm in}_0\langle q'_1,\dots,q'_m|\Psi(0)|q_1,\dots,q_n\rangle^{\rm in}_0\right. \\
&\quad\left.+\frac{1}{2\pi i}\sum_{k=2}^{+\infty}(2\pi\lambda)^k\sum_{n_1}\dots\right.
\end{aligned} \tag{21.2.10}$$

$$\sum_{n_{k-1}} \left[\frac{\delta(Q-P_1)\dots\delta(Q-P_{k-1})}{(E-E_1+i\epsilon)\dots(E-E_{k-1}+i\epsilon)} \right.$$
$$\left. \times \, {}^{\text{in}}_{0}\langle q'_1,\dots,q'_m|\Psi(0)|n_1\rangle_0 \dots {}_0\langle n_{k-1}|\Psi(0)|q_1,\dots,q_n\rangle^{\text{in}}_0 \right] \Big\} ,$$

where E and E_i (Q and P_i) denote the total energy (momentum) of the initial state and of the i-th intermediate state, respectively. Each intermediate sum can equivalently be taken either on the basis of the *in* states or on that of *out* states. Since the matrix elements between asymptotic states of the perturbing operator $\Psi(x)$ are supposed to be known, the scattering amplitudes (21.2.10) are in principle computable quantities, order by order in the coupling constant λ. The above expansion over intermediate states must be contrasted with the usual formalism of covariant perturbation theory in which both energy and momentum are conserved in the internal lines of Feynman diagrams but the corresponding particles are off-shell.

Refinements: normalization conditions. The above formula (21.2.10), however, is not completely correct because the new interaction changes both the vacuum energy density and the mass of the particles. We have to refine the action $\mathcal{S}_I$ with the introduction of some counterterms to take properly into account the correct normalization of the states. We impose the validity of the following conditions for any value of the coupling constant: the normalization of the vacuum state

$$\langle 0|0\rangle = {}_0\langle 0|0\rangle_0 = 1, \tag{21.2.11}$$

and the normalization of the one-particle states

$${}^{\text{out}}\langle q'|q\rangle^{\text{in}} = {}^{\text{out}}_{0}\langle q'|q\rangle^{\text{in}}_{0} = 2\pi E \delta(q'^1 - q^1). \tag{21.2.12}$$

The two conditions given above should be enforced order by order in perturbation theory when one uses eqn (21.2.10) to compute the vacuum-to-vacuum transition and the one-particle amplitudes.

The condition (21.2.11) leads to a subtraction of a constant term $\delta\mathcal{E}_{vac}(\lambda)$ from the interaction density. This extra term obviously measures the variation of the vacuum energy density under the effect of the perturbation. This effect is usually ignored in lagrangian perturbation theory with the prescription of disregarding the disconnected vacuum bubble diagrams. We keep track of this term here because, for the class of models we are considering, it is a measurable quantity.

To enforce the correct one-particle normalization we need to introduce a "mass" term operator in the interaction density. This operator, denoted here by $O^{(2)}(x)$, can be defined in terms of its (unperturbed) form factors given by

$$F_n^{O^{(2)}} = {}_0\langle 0|O^{(2)}(0)|q_1,\dots,q_n\rangle = \delta_{n,2}. \tag{21.2.13}$$

With this definition, the coefficient in front of the operator $O^{(2)}(x)$ in the interaction density plays the role of a mass counterterm $\delta m^2(\lambda)$ and has to be determined by imposing eqn (21.2.12) order by order in the coupling λ.

In summary, the correct formula for the scattering amplitude is given by

$$
\begin{aligned}
&{}^{\rm out}\langle q'_1,\dots,q'_m|q_1,\dots,q_n\rangle^{\rm in} \\
&\quad = {}^{\rm in}_0\langle q'_1,\dots,q'_m|q_1,\dots,q_n\rangle^{\rm in}_0 \\
&\qquad -i(2\pi)^2\delta^{(2)}\left(\sum_{j=1}^m q'_j-\sum_{j=1}^n q_j\right) \\
&\qquad \times\left\{{}^{\rm in}_0\langle q'_1,\dots,q'_m|\left(\lambda\Psi(0)-\frac{1}{2}\delta m^2 O^{(2)}(0)-\delta\mathcal{E}_{\rm vac}\right)|q_1,\dots,q_n\rangle^{\rm in}_0\right. \\
&\qquad +\sum_{n_1}\frac{2\pi\delta(Q-P_1)}{(E-E_1+i\epsilon)}{}^{\rm in}_0\langle q'_1,\dots,q'_m|\left(\lambda\Psi(0)-\frac{1}{2}\delta m^2 O^{(2)}(0)-\delta\mathcal{E}_{\rm vac}\right)|n_1\rangle_0 \\
&\qquad \left._0\langle n_1|\left(\lambda\Psi(0)-\frac{1}{2}\delta m^2 O^{(2)}(0)-\delta\mathcal{E}_{\rm vac}\right)|q_1,\dots,q_n\rangle^{\rm in}_0+\cdots\right\}.
\end{aligned}
\tag{21.2.14}
$$

Let us remark that the above expansion appears as the most physical one since it deals with the true physical degrees of freedom of the problem. But, as in any quantum field theory, divergent contributions are expected to pop up when the above formula is applied beyond the first perturbative order. A general discussion of such divergences and of the renormalization procedure which must be adopted to deal with the infinities, seems to be an interesting open problem in the unconventional setting we are considering: notice that the perturbing operator $\Psi(x)$ has in general non-vanishing matrix elements on *all* the asymptotic states and therefore a sum of an infinite number of terms is required at any perturbative order beyond the first. However the general aspects of this problem are beyond the scope of this book and will not be investigated further here. We rather concentrate our attention only on the first-order approximation because this is enough to catch the leading effects induced by a small perturbation which breaks the integrability.

21.3 First-order Perturbation Theory

Let us now apply the results of the previous section to study the action (21.1.1), where the integrable theory is defined by the action (21.1.1) with $\lambda_2 = 0$, whereas its perturbation is given by the relevant scalar operator $\varphi_2(x)$. Let x_1 and x_2 be the scaling dimensions of the two operators φ_1 and φ_2. The theory depends in this case on the two dimensionful coupling constants[4] λ_1 and λ_2. Since $\lambda_1 \sim M^{2-x_1}$ and $\lambda_2 \sim M^{2-x_2}$ (where M is a mass scale), we can decide to use λ_1 as the dimensionful parameter of the theory and the dimensionless combination

$$
\chi \equiv \lambda_2 \lambda_1^{-\frac{2-x_2}{2-x_1}} \tag{21.3.1}
$$

as a label of the different renormalization group trajectories which originate from the fixed point at $\lambda_1 = \lambda_2 = 0$. For example, if $N(\chi)$ denotes the number of stable particles

[4]For the sake of simplicity of notation, we assume that no other terms are generated by renormalization effects, as it happens for the models that we will discuss later. The first-order corrections do not depend on this assumption.

in the spectrum of the theory, their masses can be expressed as

$$m_a(\lambda_1,\chi), = \mathcal{C}_a(\chi)\lambda_1^{\frac{1}{2-x_1}}, \qquad a = 1,2,\ldots,N(\chi), \tag{21.3.2}$$

where $\mathcal{C}_a(\chi)$ is an amplitude which characterizes the whole trajectory. Similarly, the vacuum energy density can be written as

$$\mathcal{E}_{\rm vac}(\lambda_1,\chi) = \mathcal{E}(\chi)\lambda_1^{\frac{2}{2-x_1}}. \tag{21.3.3}$$

Dimensionless quantities, as for instance mass ratios, only depend on χ and therefore they do not vary along the trajectories of the renormalization group.

Once the new interaction $\lambda_2 \int d^2x\varphi_2(x)$ is switched on in the action, the integrability of the unperturbed theory is generally lost and the S-matrix amplitudes become complicated quantities. Inelastic processes of particle production are no longer forbidden and, as a consequence, the analytic structure of the scattering amplitudes present additional cuts due to the higher thresholds. In particular, their expression is no longer factorized into the sequence of two-body scattering amplitudes and, even in elastic channels, the only surviving restriction on the final momenta comes from energy–momentum conservation.

Knowledge of the matrix elements of the perturbing field $\varphi_2(x)$ ensures the possibility of computing perturbatively both the amplitudes of the inelastic processes and the corrections to the elastic ones. To first order in λ_2, with an obvious extension of the notation, eqn (21.2.14) reads

$$\begin{aligned}
&{}^{out}\langle b_1(q_1^1)\ldots b_m(q_m^1)|a_1(p_1^1)\ldots a_n(p_n^1)\rangle^{in} \\
&\quad\simeq \delta_{mn}{}_0^{out}\langle b_1(q_1^1)\ldots b_n(q_n^1)|a_1(p_1^1)\ldots a_n(p_n^1)\rangle_0^{in} - i\delta^2\left(\sum_{k=1}^{n} p_k^\mu - \sum_{k=1}^{m} q_k^\mu\right) \\
&\qquad\times {}_0^{out}\langle b_1(q_1^1)\ldots b_n(q_m^1)|\left(\lambda_2\varphi_2(0) - \frac{1}{2}\sum_{a,b=1}^{N}\delta M_{ab}^2 O_{ab}^{(2)}(0) - \delta\mathcal{E}_{\rm vac}\right) \\
&\qquad |a_1(p_1^1)\ldots a_n(p_n^1)\rangle_0^{in}.
\end{aligned} \tag{21.3.4}$$

The "mass operator" $O_{ab}^{(2)}(x)$ is defined by assigning its form factors. With an obvious generalization of eqn (21.2.13), they are given by

$$F^{O_{ab}^{(2)}}_{a_1\ldots a_n}(\theta 1,\ldots,\theta_n) = \delta_{n2}\delta_{aa_1}\delta_{ba_2}.$$

The first-order corrections to the masses of the particles and to the vacuum energy density are obtained by imposing the conditions (21.2.12) and (21.2.11). The result is

$$\delta M_{ba}^2 \simeq 2\lambda_2 F_{ba}^{\varphi_2}(i\pi,0)\delta_{m_a m_b}, \tag{21.3.5}$$

$$\delta\mathcal{E}_{\rm vac} \simeq \lambda_2\left[{}_0\langle 0|\varphi_2|0\rangle_0\right]. \tag{21.3.6}$$

Role of the rapidity. However, one must be careful in using the rapidity parameterization in eqn (21.3.4). To illustrate this point let's consider the first-order correction to

some elastic process $ab \to cd$. In the unperturbed theory, this process is characterized by the scattering amplitude $S_{ab}^{cd}(\theta)$, where $\theta = \theta_1 - \theta_2$ denotes the rapidity difference of the colliding particles. In two dimensions, the momenta of the particles in a two-body elastic collision are individually conserved even in the absence of integrability, so that the general elastic amplitude $S_{ab}^{cd}(\theta, \chi)$ can be introduced through the relation

$$^{out}\langle c(\theta_1)d(\theta_2)|a(\theta_3)b(\theta_4)\rangle^{in} = (2\pi)^2\delta(\theta_1 - \theta_3)\delta(\theta_2 - \theta_4)S_{ab}^{cd}(\theta_1 - \theta_2, \chi). \qquad (21.3.7)$$

However, note that away from the integrable direction (i.e. $\chi = 0$), the scattering amplitude $S_{ab}^{cd}(\theta, \chi)$ is no longer a meromorphic function of θ for the opening of inelastic channels. In computing the correction to $S_{ab}^{cd}(\theta, \chi)$ around $\chi = 0$, we must take into account that the total energy of the colliding system is fixed and therefore the variation in the masses given by eqn (21.3.5) induces a corresponding change in the rapidity difference, expressed by

$$\delta\theta \simeq -\frac{m_a\delta m_a + m_b\delta m_b + (m_b\delta m_a + m_a\delta m_b)\cosh\theta}{m_a m_b \sinh\theta}. \qquad (21.3.8)$$

Then the correction to the amplitude can be decomposed as

$$\delta S_{ab}^{cd}(\theta, \chi) = \frac{\partial S_{ab}^{cd}(\theta)}{\partial\theta}\delta\theta + \left.\frac{\partial S_{ab}^{cd}(\theta, \chi)}{\partial\chi}\right|_{\chi=0}\delta\chi. \qquad (21.3.9)$$

The first-order result for this quantity is obtained by using formula (21.3.4). Taking into account the cancellation occurring between the disconnected parts of the form factors and the contributions of the counterterms, one finally obtains

$$\delta S_{ab}^{cd}(\theta, \chi) \simeq -i\lambda_2 \frac{F_{\bar{c}\bar{d}ab}^{\varphi_2}(\theta)}{m_a m_b \sinh\theta}, \qquad (21.3.10)$$

where

$$F_{\bar{c}\bar{d}ab}^{\varphi_2}(\theta_1 - \theta_2) \equiv F_{\bar{c}\bar{d}ab}^{\varphi_2}(\theta_1 + i\pi, \theta_2 + i\pi, \theta_1, \theta_2). \qquad (21.3.11)$$

Cancellation of $i\pi$ singularities. The right-hand side of (21.3.10) employs the expression of the form factor at very special values of the rapidity variables. According to eqn (21.3.5), the form factors present pole singularities whenever the rapidities of a particle–antiparticle pair differ by $i\pi$ and these kinematical poles are often explicitly inserted into the denominator of their parameterization. Apart from a term encoding the monodromy properties, in Chapter 20 we have showed that this parameterization may be written as $\mathcal{Q}/\mathcal{D}$ where both $\mathcal{Q}$ and $\mathcal{D}$ are polynomials in the variables $\cosh\theta_{ij}$: the denominator is uniquely fixed by the pole structure of the S-matrix whereas the numerator is determined by means of the residue equations, as for instance those of eqn (21.3.5). From the finiteness of the left-hand side of eqn (21.3.10), we expect therefore that the "$i\pi$ singularities" of the denominator of the form factors $F_{\bar{a}\bar{b}ab}^{\varphi_2}(\theta_1, \theta_2, \theta_3, \theta_4)$ should be cancelled by the polynomial $\mathcal{Q}$, once evaluated at the specific rapidity configuration of eqn (21.3.11). This should hold in general whenever

the perturbing operator is local with respect to the fields that create the particles in the unperturbed theory.[5]

Universal ratios. Equations (21.3.5), (21.3.6), and (21.3.10) are the main results of this section. The best use of these formulas is to get rid of the explicit dependence on the normalization of the perturbing operator by defining universal quantities, as for instance ratios of the mass shifts. Hence, under the validity of the linear approximation, all the universal quantities of non-integrable field theories can be entirely expressed in terms of form factors of the integrable ones. A comparison of the theoretical predictions with their numerical determinations will be presented in the sequel.

Trivial deformation and its consequences. It is particularly instructive to focus the above discussion on the "trivial" case in which the perturbing operator $\varphi_2(x)$ coincides with the operator $\varphi_1(x)$ that defines the initial integrable theory. In this case, of course, the physics should be invariant, since the result of the additional perturbation simply corresponds to a shift of the coupling constant of the original integrable model by an amount $\delta\lambda_1 = \lambda_2$. The variations of the masses of the particles and the vacuum energy density corresponding to such a shift can be directly computed from eqns (21.3.2), and (21.3.3), respectively. But we can also apply our general formulas (21.3.5) and (21.3.6) to estimate the first order corrections. The two different routes coincide as long as the following identities are valid

$$\begin{aligned} &F^{\Theta}_{\bar{a}a}(i\pi, 0) = 2\pi m_a^2, \\ &\mathcal{E}_{vac} = \tfrac{1}{4\pi}\langle 0|\Theta|0\rangle, \end{aligned} \tag{21.3.12}$$

where $\Theta(x) = 2\pi\lambda_1(2 - x_1)\varphi_1(x)$ is the trace of the energy–momentum tensor for the trajectory $\chi = 0$. The two relationships above are indeed true and can be easily derived by other means, as we know from Chapter 20: it is interesting to notice that, in this context, their validity emerges as consistency equations. By the same token, considering higher multiparticle scattering processes, we can generate an infinite number of identities involving the form factors of the original integrable field theory. For instance, next to (21.3.12), a new identity is obtained by comparing eqn (21.3.9) with eqn (21.3.10): since χ is constant in the case we are considering, we have

$$\frac{\partial S^{cd}_{ab}(\theta)}{\partial\theta} = -\frac{1}{2\pi i}\frac{F^{\Theta}_{\bar{c}\bar{d}ab}(\theta)}{s_{ab}(\theta)}. \tag{21.3.13}$$

This identity provides a simple and unique way to normalize the four-particle form factors of the stress–energy tensor. It may then be particularly useful in the study of massless field theories where the first relationship in eqn (21.3.12) cannot be used for this purpose.

It is also obvious that the first-order inelastic amplitudes computable by formula (21.3.4) must vanish identically when we choose $\varphi_2(x) = \varphi_1(x)$. This is ensured by the fact that the form factors of the stress–energy tensor $F^{\Theta}_{a_1\dots a_n}(\theta_1, \dots, \theta_n)$ factorize the term $P_\mu P^\mu$, with $P^\mu = \sum_{i=1}^n p_i^\mu$ denoting the total energy–momentum of the set

[5] We refer to the section devoted to the Ising model for the discussion of the case in which this condition is not fulfilled.

of particles. Since $p_i^\mu \to -p_i^\mu$ when the i-th particle is crossed from the initial to the final state, P^μ is zero for a set of particles entering a physical scattering process. Only in the case of elastic scattering, the zeros coming from the factor $P_\mu P^\mu$ cancel the kinematical poles and relations analogous to eqn (21.3.13) are obtained.

21.4 Non-locality and Confinement

Let us consider in more detail the mass correction, given by

$$\delta m_a^2 \simeq 2\lambda F_{a\bar{a}}^{\varphi_2}(i\pi), \tag{21.4.1}$$

where the form factor of the operator $\varphi_2(x)$ is defined by the matrix element

$$F_{a\bar{a}}^{\varphi_2}(\theta) \equiv \langle 0|\varphi_2(0)|a(\theta_1)\bar{a}(\theta_2)\rangle. \tag{21.4.2}$$

Let us recall that the two-particle form factor of an integrable theory satisfies the equations

$$F_{a\bar{a}}^{\mathcal{O}}(\theta) = S_{a\bar{a}}^{b\bar{b}}(\theta) F_{\bar{b}b}^{\mathcal{O}}(-\theta), \tag{21.4.3}$$

$$F_{a\bar{a}}^{\mathcal{O}}(\theta + 2i\pi) = e^{-2i\pi\gamma_{\mathcal{O},a}} F_{\bar{a}a}^{\mathcal{O}}(-\theta). \tag{21.4.4}$$

In the second equation the explicit phase factor $e^{-2i\pi\gamma_{\mathcal{O},a}}$ is inserted to take into account a possible semilocality of the operator which interpolates the particles and the operator $\mathcal{O}(x)$. If $\gamma \neq 0$, the two-particle form factor presents a pole at $\theta = \pm i\pi$, with the residue given by

$$-i\mathrm{R}es_{\theta=\pm i\pi} F_{a\bar{a}}^{\mathcal{O}}(\theta) = (1 - e^{\mp 2i\pi\gamma_{\mathcal{O},a}})\langle 0|\mathcal{O}|0\rangle. \tag{21.4.5}$$

According to whether the perturbing field is local or non-local with respect to the asymptotic particles, there are two different scenarios. If the field that breaks integrability is a local operator, the mass correction of the particles is finite. Vice versa, if the perturbing field is non-local, the mass correction of the particles is divergent. The last case implies the confinement of the particles that occurs as soon as the non-integrable perturbation is turned on.

There are several ways to show the confinement phenomena. One consists of computing the propagator $\langle A(p)A(-p)\rangle$ of the particle A in the perturbed theory. At the tree level approximation, shown in Fig. 21.3, this consists of a geometric series that can be explicitly summed and for the propagator of the perturbed theory we have

$$\langle A(p)A(-p)\rangle \simeq \frac{1}{p^2 - m^2 - \delta m^2}. \tag{21.4.6}$$

If $\delta m^2 = \infty$, the propagator obviously vanishes, i.e. the particle disappears from the spectrum. A more intuitive explanation of the confinement phenomenon comes from the analysis of the Ising model.

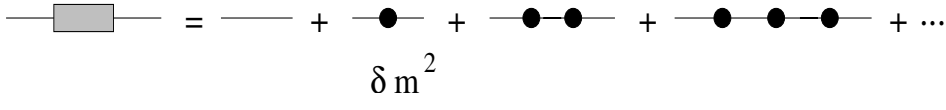


Fig. 21.3 *Perturbative series of the propagator.*

21.5 The Scaling Region of the Ising Model

In this section we study the evolution of the mass spectrum of the Ising model by moving its couplings along the path C shown in Fig. 21.4 in the plane (τ, h): this curve starts from the low-temperature phase of the model and ends at its high-temperature phase, here represented by the points (a) and (f), respectively.

The action of the model is given by

$$\mathcal{A} = \mathcal{A}_{CFT} + \tau \int d^2x \varepsilon(x) + h \int d^2x \sigma(x). \tag{21.5.1}$$

As shown in the previous chapters, this theory gives rise to a massive integrable model when one of the two coupling constants is switched off. The action (21.5.1) defines a family of field theories identified by $\chi \equiv \tau |h|^{-8/15} \in (-\infty, +\infty)$, a dimensionless RG invariant quantity. The spectrum of the theory changes in a significant way moving χ.

In the low-temperature phase (corresponding to $\chi = -\infty$ and to the point (a) of the curve C), the model has two degenerate vacua and therefore its excitations consist of the topological kink and antikink that are interpolated between the two ground states. Along the magnetic axes ($\chi = 0$, corresponding to the point (d) of the curve C), the spectrum of the model consists instead of eight particles, with different masses. Finally, in the high-temperature phase (i.e. $\chi = +\infty$), the system has a unique vacuum and only a massive excitation above it. Let's see how this scenario can be recovered by form factor perturbation theory.

Let's start our analysis from the point (a), where the massive excitations are the kink/antikinks that interpolate between the two degenerate vacua. By switching on the magnetic field, the model moves to the point (b) of the curve C. The form factor of the perturbing field σ on the two-particle kink/antikink state is given by

$$F^{\sigma}(\theta_1 - \theta_2) = \langle 0|\sigma(0)|A(\theta_1)A(\theta_2)\rangle = \tanh \frac{\theta_1 - \theta_2}{2}.$$

Therefore eqn (21.3.5) leads to an infinite correction to the mass of the kinks, i.e. the kinks get confined as soon as the magnetic field is switched on. Looking at the effective potential of the theory, it is not difficult to see that this is the correct conclusion: no

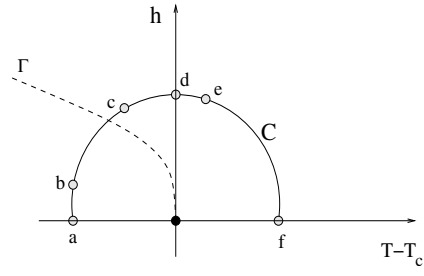


Fig. 21.4 *Interpolation curve in the (τ, h) plane between the low- and high-temperature phases of the Ising model. Γ is a renormalization group trajectory, identified by the dimensionless parameter χ.*

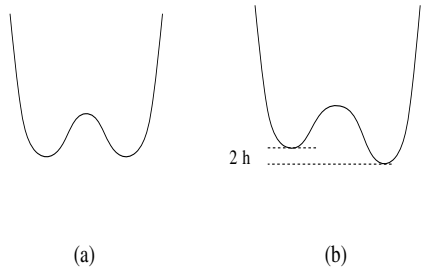


Fig. 21.5 *Effective potential in the low-temperature phase (a) and in the presence of an infinitesimal magnetic field (b). In the last case, the two minima are no longer degenerate and the kink/antikink disappears from the spectrum of the asymptotic states.*

matter how small the magnetic field may be, it lifts the degeneracy of the two vacua, as shown in Fig. 21.5, and consequently there is no longer the possibility of having topological configurations.

Consider now the effect of the magnetic field on a state made of a kink and an antikink separated by a distance R. When the magnetic field is absent, the energy of this state is essentially equal to $2M$, i.e. the sum of the masses of the kink and the antikink. The energy of this state depends very weakly on the distance R because this field configuration takes values on the zeros of the effective potential and, no matter how large the distance R could be, there is no change in the energy of this state. This situation changes by switching on the magnetic field since, in this case, at every point of the space there is an energy gap equal to $2h$ and the energy U of this state becomes a linear function of R, $U(R) = 2M + 2hR$. This attractive interaction between the kink and the antikink gives rise to a discrete spectrum of bound states. Regarding the kinks as very massive and quasi-static, the energy of the bound states can be obtained by solving the quantum mechanical problem of the bound states for a linear potential, well-known in quantum mechanics. The result is simply

$$E_k \equiv m_k = (2 + h^{2/3}\gamma_k^{2/3})M, \tag{21.5.2}$$

where γ_k are the positive roots of the equation

$$\mathcal{J}(\gamma_k) = J_{\frac{1}{3}}\left(\frac{1}{3}\gamma_k\right) + J_{-\frac{1}{3}}\left(\frac{1}{3}\gamma_k\right) = 0$$

($J_\nu(x)$ is the Bessel function of order ν) (Fig. 21.6). The structure of the bound states is shown in Fig. 21.7. Obviously not all these states are stable: the stable ones are identified by the condition $m_n < 2m_1$, while all particles with a mass higher than the threshold $2m_1$ decay into particles of lower masses. When χ increases, i.e. when the system moves clockwise along the curve C of Fig. 21.4, the number of bound states monotonically decreases. At the point (d), there are the eight stable particles of Zamolodchikov's solution of the Ising model in a magnetic field, and the value of their mass can be found

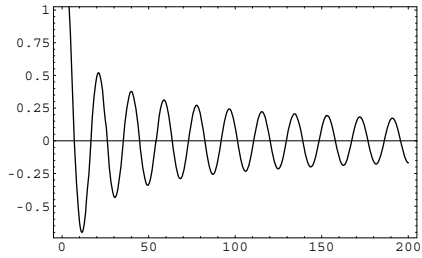


Fig. 21.6 *Plot of the function $\mathcal{J}(x)$. The zeros of this function determine the energies of the bound states of the Ising model coming from the original kink–antikink state.*

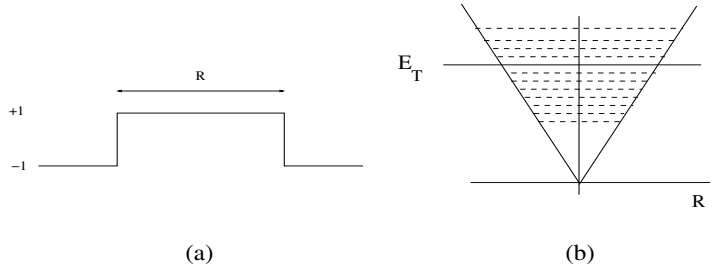


Fig. 21.7 *(a) Kink–antikink state separated by a distance R; (b) Kink–antikink potential in the presence of a magnetic field h and its bound states. The stable bound states are identified by the condition $E_n < E_T$.*

in Section 18.4.2. It is worth stressing that in this case the five particles with mass higher than the threshold are stable just for the integrability of the model. Moving away from the magnetic axes by means of the operator $\varepsilon(x)$, the first three particles change the value of their masses, while the remaining five particles decay into the low-energy channels. To estimate both effects, we need the form factors of the energy operator but in the integrable theory of the Ising model in a magnetic field. Here we simply report their expressions

$$\begin{aligned}
&\langle 0|\varepsilon(0)|0\rangle = m_1,\\
&F_{11}^{\varepsilon}(i\pi) = \langle 0|\varepsilon(0)|A_1(\theta+i\pi)A_1(\theta)\rangle = -17.8933\ldots m_1,\\
&F_{22}^{\varepsilon}(i\pi) = \langle 0|\varepsilon(0)|A_2(\theta+i\pi)A_2(\theta)\rangle = -24.9467\ldots m_1,\\
&F_{33}^{\varepsilon}(i\pi) = \langle 0|\varepsilon(0)|A_3(\theta+i\pi)A_3(\theta)\rangle = -53.6799\ldots m_1,\\
&F_{44}^{\varepsilon}(i\pi) = \langle 0|\varepsilon(0)|A_4(\theta+i\pi)A_4(\theta)\rangle = -49.3169\ldots m_1.
\end{aligned}$$

The first equation may be regarded as the normalization condition of the energy operator $\varepsilon(x)$. The corrections of the universal ratios are given by

$$\begin{aligned}
\frac{\delta \mathcal{E}_{vac}}{\delta m_1} &= \frac{\langle 0|\varepsilon|0\rangle}{F^{\varepsilon}_{11}(i\pi)} m_1^0 = -0.0558\ldots m_1^0, \\
\frac{\delta m_2}{\delta m_1} &= \frac{F^{\varepsilon}_{22}(i\pi)}{F^{\varepsilon}_{11}(i\pi)} \frac{m_1^0}{m_2^0} = 0.8616\ldots, \\
\frac{\delta m_3}{\delta m_1} &= \frac{F^{\varepsilon}_{33}(i\pi)}{F^{\varepsilon}_{11}(i\pi)} \frac{m_1^0}{m_3^0} = 1.5082\ldots.
\end{aligned} \tag{21.5.3}$$

In turn, these quantities can be independently determined by a numerical solution of the model and the values determined in this way are

$$\begin{aligned}
\frac{\delta \mathcal{E}_{vac}}{\delta m_1} &\simeq -0.05 m_1^0, \\
\frac{\delta m_2}{\delta m_1} &\simeq 0.87, \\
\frac{\delta m_3}{\delta m_1} &\simeq 1.50.
\end{aligned} \tag{21.5.4}$$

As can be seen from the expressions above, there is satisfactory agreement between the theoretical and numerical estimates.

Breaking the integrability of the Ising model in a magnetic field has a more dramatic effect on the five particles with a mass above threshold. Their stability is only due to integrability and, in its absence, they decay. In the perturbative approach, the decay process is associated to the presence of a negative imaginary part in the mass that is a second-order perturbative effect in τ, as shown in Fig. 21.8.

$$\mathrm{Im} m_c^2 = -\sum_{a\le b, m_a+m_b\le m_c} m_c \Gamma_{c\to ab} \simeq -\tau^2 \sum_{a\le b, m_a+m_b\le m_c} 2^{1-\delta_{ab}} \frac{|f_{cab}|^2}{m_c m_a \left|\sinh\theta_a^{(cab)}\right|},$$

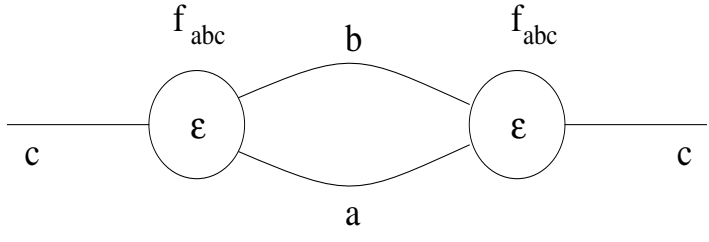


Fig. 21.8 *Perturbative diagram at second order in τ relative to the imaginary part of the mass of the particle c. The intermediate particles a and b satisfy the on-shell conditions $p_a^2 = m_a^2$ e $p_b^2 = m_b^2$. When $c > 5$, there are additional diagrams with more intermediate particles.*

where $\Gamma_{c\to ab}$ is the decay amplitude of the particle A_c into the two particles A_aA_b, whereas

$$f_{cab} = F^{\varepsilon}_{cab}(i\pi, \theta_a^{(cab)}, \theta_b^{(cab)})\Big|_{\tau=0}.$$

The rapidities $\theta_a^{(cab)}$ and $\theta_b^{(cab)}$ are fixed by the conservation of energy and momentum in the decay process $A_c \to A_aA_b$ in the rest frame of the particle A_c. In the above equations all masses are the unperturbed values at $\tau = 0$. When $c > 5$ the sum must be completed including the contribution of the decay channels with more than two particles in the final state. Once the decay amplitudes $\Gamma_{c\to ab}$ are known, one can determine the lifetime t_c of the unstable particle A_c given by

$$t_c = \frac{1}{\Gamma_c}, \qquad \Gamma_c = \sum_{a\le b} \Gamma_{c\to ab}. \tag{21.5.5}$$

For the Ising model, the relevant matrix elements are

$$\begin{aligned} |f_{411}| &= (36.73044\ldots)\,|\langle\varepsilon\rangle|_{\tau=0} \\ |f_{511}| &= (19.16275\ldots)\,|\langle\varepsilon\rangle|_{\tau=0} \\ |f_{512}| &= (11.2183\ldots)\,|\langle\varepsilon\rangle|_{\tau=0} \end{aligned}$$

where the normalization of the operator ε is fixed by its vacuum expectation value

$$\langle\varepsilon\rangle_{\tau=0} = (2.00314\ldots)|h|^{8/15}.$$

The imaginary part of the mass of the first two particles, which are over threshold, is given by

$$\begin{aligned} \mathrm{Im}m_4^2 &\simeq (-840.172\ldots)\left(\frac{\tau\langle\varepsilon\rangle_{\tau=0}}{m_1}\right)^2 = (-173.747\ldots)\tau^2 \\ \mathrm{Im}m_5^2 &\simeq (-240.918\ldots)\left(\frac{\tau\langle\varepsilon\rangle_{\tau=0}}{m_1}\right)^2 = (-49.8217\ldots)\tau^2. \end{aligned}$$

The ratio of their lifetime is universal

$$\lim_{\tau\to 0}\frac{t_4}{t_5} = \lim_{\tau\to 0}\frac{m_4\mathrm{Im}\ m_5^2}{m_5\mathrm{Im}\ m_4^2} = 0.23326\ldots \tag{21.5.6}$$

While the particle A_4 can only decay into A_1A_1, the particle A_5 can also decay into the channel A_1A_2. The ratios of the amplitudes of these decays

$$b_{c\to ab} = \frac{m_c|_{\tau=0}\Gamma_{c\to ab}}{|\mathrm{Im}m_c^2|}$$

are given by

$$\lim_{\tau\to 0} b_{5\to 11} = 0.47364\ldots, \qquad \lim_{\tau\to 0} b_{5\to 12} = 0.52635\ldots.$$

Notice that eqn (21.5.6) predicts that the lifetime of the particle A_5 is almost four times longer than the lifetime of the particle A_4. This paradoxical result, in contradiction

with the intuitive idea that a heavy particle should decay faster than a light one, finds its explanation once again in the peculiar behavior of the phase space in two dimensions. For the decay process $A_c \to A_a A_b$ the phase space in d dimensions is given by

$$\int \frac{d^{d-1}\vec{p}_a}{p_a^0} \frac{d^{d-1}\vec{p}_b}{p_b^0} \delta^d(p_a - p_b) \sim \frac{p^{d-3}}{m_c}, \tag{21.5.7}$$

where $p = |\vec{p}_a| = |\vec{p}_b|$ is the value in the rest frame. For fixed decay products, p grows with m_c: in $d = 2$, this term joins the factor m_c in the denominator and leads to a suppression of the phase space. In the Ising model, eqn (21.5.6) shows that this suppression is further enhanced by the dynamics (i.e. by the vaues of the matrix elements) in a way that is not compensated by the additional decay channels.

If we keep moving along the curve C, one first meets a value χ_1 at which the mass of the particle A_3 becomes larger than $2m_1$ and, later on, a second value χ_2 at which also the mass of the particle A_2 becomes larger than $2m_1$. When $\chi > \chi_2$ the spectrum of the stable particles of the theory consists of only one excitation. In the limit $\chi \to +\infty$, this is nothing but the particle of the integrable theory of the high-temperature phase of the model.

References and Further Reading

The form factor perturbation theory of non-integrable models has been introduced in the articles:

G. Delfino, G. Mussardo, P. Simonetti, *Nonintegrable quantum field theories as perturbations of certain integrable models*, Nucl. Phys. B 473 (1996), 469.

G. Delfino, G. Mussardo, *Nonintegrable aspects of the multifrequency Sine–Gordon model*, Nucl. Phys. B 516 (1988), 675.

The decay processes of the heavier particles of the Ising model in a magnetic field has been computed in:

G. Delfino, P. Grinza, G. Mussardo, *Decay of particles above threshold in the Ising field theory with magnetic field*, Nucl. Phys. B 737 (2006), 291.

The decay processes in the Ising model have also been studied in:

S. B. Rutkevich, *Large-n excitations in the ferromagnetic ising field theory in a weak magnetic field: Mass spectrum and decay widths*, Phys. Rev. Lett. 95, (2005), 250601.

The infinite tower of bound states in the Ising model was first identified in:

B.M. McCoy, T.T. Wu, *Two-dimensional Ising field theory in a magnetic field: Breakup of the cut in the two-point function*, Phys. Rev. D 18, 1259 (1978).

The analytic properties of the free energy of the Ising model by varying T and h can be found in:

P. Fonseca, A. Zamolodchikov, *Ising field theory in a magnetic field: Analytic properties of the free energy*, J. Stat. Phys. 110 (2003), 527.

Other references which may be useful to understand non-integrable effects are:

D. Controzzi, G. Mussardo, *On the mass spectrum of the two-dimensional $O(3)$ sigma model with θ term*, Phys. Rev. Lett. 92 (2004), 021601.

G. Mussardo, V. Riva, G. Sotkov, G. Delfino, *Kink scaling functions in 2-D non-integrable quantum field theories*, Nucl. Phys. B 736 (2006), 259.

G. Mussardo, *Neutral bound states in kink-like theories*, Nucl. Phys. B 779 (2007), 101.

G. Mussardo, *Kink confinement and supersymmetry*, J. High Energy Phys., JHEP 08 (2007), 003.

L. Lepori, G. Mussardo, G. Toth, *The particle spectrum of the tricritical Ising model with spin reversal symmetric perturbations*, JSTAT 0809, (2008), P09004.

G. Delfino, P. Grinza, *Confinement in the q-state Potts field theory*, Nucl. Phys. B 799 (2008), 364.

Problems

1. Two-frequency Sine–Gordon model

Consider the Sine–Gordon model for the bosonic scalar field $\varphi(x)$ with two frequencies. The most general lagrangian is given by

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)^2 + \mu \cos \beta\varphi + \lambda \cos(\alpha\varphi + \delta).$$

When $\lambda = 0$ the theory is integrable and it is possible to determine the semilocal index γ of the vertex operator $e^{i\alpha\varphi}$ with respect to the solitons of the unperturbed theory. It is given by

$$\gamma = \frac{\alpha}{\beta}.$$

a Prove that the form factor of the operator $\Psi(x) = \cos(\alpha\varphi + \delta)$ on the soliton–antisoliton state has a pole at $\theta = \pm i\pi$ with residue equal to

$$-i\mathrm{Res}_{\theta=\pm i\pi} F^{\Psi}_{s\bar{s}}(\theta) = [\cos\delta - \cos(\delta \mp 2\pi\alpha/\beta)]\, \langle 0|e^{i\alpha\varphi}|0\rangle.$$

b Use the above result to prove that in the perturbed theory ($\lambda \neq 0$) with generic values of α, β, and δ, the soliton and the antisoliton of the original integrable theory are all confined.

c Consider now the case $\alpha/\beta = m/n$, with m, n two integer numbers. If

$$| \delta | = \frac{\pi}{n}$$

and if there exists an integer number k such that

$$| kn - m | = 1$$

prove that in the perturbed theory either the soliton or the antisoliton is confined but not both.
(Hint. To solve the exercise it is useful to plot the potential of the theory for these particular vales of α, β and δ.)

Index