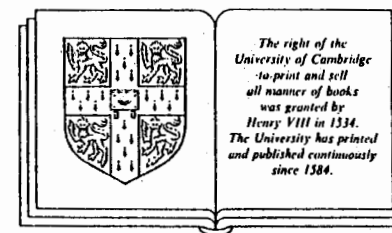


Volume 1
From Brownian motion to renormalization
and lattice gauge theory

CLAUDE ITZYKSON
JEAN-MICHEL DROUFFE

*Commissariat à l'Energie Atomique
Centre d'Etudes Nucléaires de Saclay
Service de Physique Théorique*



CAMBRIDGE UNIVERSITY PRESS

Cambridge

New York Port Chester Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1989

First published 1989

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Itzykson, Claude.

Statistical field theory / Claude Itzykson, Jean-Michel Drouffe.

p. cm. — (Cambridge monographs on mathematical physics)

Includes bibliographies and indexes.

Contents: v. 1. From Brownian motion to renormalization and lattice gauge theory — v. 2. Strong coupling, Monte Carlo methods, conformal field theory, and random systems.

ISBN 0-521-34058-6 (v. 1). ISBN 0-521-37012-4 (v. 2)

1. Statistical physics. 2. Field theory (Physics) I. Drouffe, Jean-Michel. II. Title. III. Series.

QC174.8.189 1989

530.1'3—dc19 88-30162

CIP

British Library Cataloging in Publication applied for.

Contents

<i>Contents of Volume 2</i>	<i>ix</i>
<i>Preface</i>	<i>xi</i>
1 From Brownian motion to Euclidean fields	1
1.1 Brownian motion	1
1.1.1 Random walks	1
1.1.2 The sum over paths	9
1.1.3 The dimension two of Brownian curves	12
1.2 Euclidean fields	21
1.2.1 Free fields	22
1.2.2 Interacting fields and random walks	25
1.2.3 Self-avoiding walks and the limit $n \rightarrow 0$	31
1.2.4 Comparison with the high temperature expansion	33
1.2.5 The one-dimensional case	37
1.A Lattices	43
Notes	47
2 Grassmannian integrals and the two-dimensional Ising model	48
2.1 Grassmannian integrals	48
2.1.1 Anticommuting variables	48
2.1.2 Integrals	52
2.2 The two-dimensional Ising model	58
2.2.1 Duality	59
2.2.2 Transfer matrix	62
2.2.3 Fermionic representation	63
2.2.4 Free energy	67
2.2.5 Spontaneous magnetization	72
2.2.6 Correlation function in the high temperature phase	78
2.2.7 Surface tension	85

vi	<i>Contents</i>	
2.3	Critical continuous theory	91
2.3.1	Effective action	91
2.3.2	Correlation functions	95
2.A	Quadratic differences and Painlevé equations	98
	Notes	104
3	Spontaneous symmetry breaking, mean field	107
3.1	Mean field approximation	108
3.1.1	Dielectric constant of a polarizable medium	108
3.1.2	Classical spin model with a finite symmetry group	112
3.1.3	Continuous symmetry group	118
3.1.4	The Bethe approximation	121
3.1.5	Critical exponents	125
3.2	Lee-Yang zeroes	131
3.2.1	The Lee-Yang theorem	132
3.2.2	The one-dimensional case	135
3.2.3	General properties	136
3.2.4	Zeroes in the temperature plane	139
3.3	Large n limit	140
3.3.1	Saddle point method	141
3.3.2	Factorization	145
3.3.3	Coupling to an external field	148
3.4	Corrections to mean field	151
3.4.1	Laplace transform	153
	Notes	159
4	Scaling transformations and the XY-model	162
4.1	Scaling laws. Real space renormalization	162
4.1.1	Homogeneity and scale invariance	162
4.1.2	Recurrence relations in real space	168
4.1.3	Examples and approximations	176
4.2	The XY-model	193
4.2.1	High temperature behaviour	195
4.2.2	Low temperature expansion. Vortices	197
4.2.3	The Villain action	204
4.2.4	Correlations	207
4.2.5	Renormalization flow	213
4.A	Two-dimensional systems with continuous symmetry	219
4.A.1	Magnetization inequality	219
4.A.2	Correlation inequality	222

	<i>Contents</i>	vii
4.B	Phenomenological renormalization	224
	Notes	230
5	Continuous field theory and the renormalization group	233
5.1	The Lagrangian and dimensional analysis	233
5.1.1	Introduction	233
5.1.2	Generating functionals and dimensional analysis	236
5.2	The perturbative method	242
5.2.1	Diagrammatic series	242
5.2.2	Loop expansion	246
5.2.3	Evaluation of integrals and dimensional continuation	250
5.2.4	Group theoretical factors	259
5.2.5	Power counting	261
5.2.6	Perturbative renormalization	264
5.3	The renormalization group	270
5.3.1	Renormalization flow	270
5.3.2	Critical exponents	280
5.3.3	From the Gaussian ultraviolet fixed point to the infrared critical point in dimension less than four	282
5.3.4	Correlation functions at the critical point	284
5.3.5	Expansion near the critical point	290
5.3.6	Scaling laws below the critical temperature	296
5.4	Corrections to scaling laws	301
5.4.1	Deviation from the critical point in dimension lower than four	301
5.4.2	Logarithmic corrections in dimension four	303
5.4.3	Irrelevant operators	307
5.5	Numerical results	310
5.5.1	ϵ -expansion of critical exponents	311
5.5.2	Equation of state	313
5.5.3	Amplitude ratios	314
5.5.4	Three-dimensional results	317
5.A	Multicritical points	317
	Notes	326
6	Lattice gauge fields	328
6.1	Generalities	328

6.1.1	Presentation	328
6.1.2	The continuous limit	332
6.1.3	Order parameter and Elitzur's theorem	341
6.1.4	Duality	345
6.2	Structure of the phase diagram	352
6.2.1	Mean field approximation	352
6.2.2	Corrections to mean field and restoration of gauge invariance	359
6.2.3	Discrete groups: $1/d$ expansion	362
6.2.4	Continuous groups: computation of corrections	365
6.3	Strong coupling expansions	371
6.3.1	Convergence	371
6.3.2	Character expansions	374
6.3.3	Free energy	380
6.3.4	String tension and roughening transition	385
6.3.5	Mass spectrum	390
6.4	Lattice fermions	393
6.4.1	The doubling problem	393
6.4.2	The Nielsen-Ninomiya theorem	396
6.4.3	Staggered fermions	399
	Notes	401
	<i>Index</i>	

Contents of Volume 2

7	Diagrammatic methods	405
8	Numerical simulations	456
9	Conformal invariance	501
10	Disordered systems and fermionic methods	646
11	Random geometry	738

Preface

Some ten years ago, when completing with J.-B. Zuber a previous text on *Quantum Field Theory*, the senior author was painfully aware that little mention was made that methods in statistical physics and Euclidean field theory were coming closer and closer, with common tools based on the use of path integrals and the renormalization group giving insights on global structures. It was partly to fill this gap that the present book was undertaken. Alas, over the five years that it took to come to life, both subjects have undergone a new evolution. Disordered media, growth patterns, complex dynamical systems or spin glasses are among the new important topics in statistical mechanics, while superstring theory has turned to the study of extended systems, Kaluza-Klein theories in higher dimensions, anticommuting coordinates ... in an attempt to formulate a unified model including all known interactions. New and sophisticated techniques have invaded statistical physics, ranging from algebraic methods in integrable systems to fractal sets or random surfaces. Powerful computers or special devices provide "experimental" means for a new brand of theoretical physicists. In quantum field theory, applications of differential topology, geometry, Riemannian manifolds, operator theory ... require a deeper background in mathematics and a knowledge of some of its most recent developments. As a result, when surveying what has been included in the present volume in an attempt to uncover the basic unity of these subjects, the authors have the same unsatisfactory feeling of not being able to bring the reader really up to date. It is presumably the fate of such endeavours to always come short of accomplishing their purpose.

With these shortcomings fully admitted, we have tried to present to the reader an overview of the main themes which justify the title "Statistical field theory." This interpretation of Euclidean field theory offers a new language, effective computing means, as

well as a natural and consistent short-distance cutoff. In other words, it allows one to give a global meaning to path integrals, to discover possible anomalies arising from integration measures, or to understand in simple terms systems with redundant variables such as gauge models. The theory of continuous phase transitions provides a bridge between probabilistic mechanics and continuous field theory, using the renormalization group to filter out relevant operators and interactions. Many authors contributed to these views, culminating in the work of K. Wilson and his collaborators and followers, which promoted the renormalization group as a universal tool to analyse the large-distance behaviour. It still retains its value, while new developments take place, particularly with conformal, or local scale invariance coming to prominence in the study of two-dimensional systems.

The content of this book is naturally divided into two parts. The first six chapters describe in succession Brownian motion, its anti-commutative counterpart in the guise of Onsager's solution to the two-dimensional Ising model, the mean-field or Landau approximation, scaling ideas exemplified by the Kosterlitz-Thouless theory for the XY-transition, the continuous renormalization group applied to the standard φ^4 theory, the simplest typical case, and lattice gauge theory as an attempt to understand quark confinement in chromodynamics.

The next five chapters (in volume 2) cover more diverse subjects. We give an introduction to strong coupling expansions and various means of analyzing them. We then briefly introduce Monte Carlo simulations with an emphasis on the applications to gauge theories. Next we turn to the significant advances in two-dimensional conformal field theory, with a lengthy presentation of the methods as well as early results. A chapter on simple disordered systems includes sample applications of fermionic techniques with no pretence at completeness. The final chapter is devoted to random geometry and an introduction to the Polyakov model of random surfaces which illustrates the relations between string theory and statistical physics.

At the price of being perhaps a bit repetitive, we have tried in the first part to introduce the subject in an elementary fashion. It is, however, assumed that the reader has some familiarity with thermodynamics as well as with quantum field theory. We often switch from one to the other interpretation, assuming that it will

not be disturbing once it is realized that the exponential of the action plays the role of the Boltzmann-Gibbs statistical weight. The last chapters cover subjects still in fast evolution.

Many important subjects could unfortunately not be covered. In random order they include dynamical critical phenomena, renormalization of σ -models or non-Abelian gauge fields except for a mention of lowest order results, topological aspects, classical solutions, instantons, monopoles, anomalies (except for the conformal one). Integrable systems are missing apart from the two-dimensional Ising model. Quantum gravity *à la Regge* is only mentioned. The list could, of course, be made much longer. Our involvement in some of the topics has certainly produced obvious biases and overemphases in certain instances. We have tried, as much as possible, to correct for these defects as well as for the numerous omissions by including at the end of each chapter a section entitled "Notes." Here we quote our sources, original articles, reviews, books and complementary material. These notes are purposely scattered through the volume, as we are sure that our quotations are very incomplete. A fair bibliography in such a large domain is beyond human capacities. Should any one feel that his or her work has not been reported or not properly mentioned, he or she is certainly right and we present our most sincere apologies. On the other hand we did not hesitate to use and sometimes follow very closely some articles or reviews which served our purpose. For instance chapter 5 is built around the definitive contributions of E. Brézin, J.-C. Le Guillou, J. Zinn-Justin and G. Parisi. Except for some further elaboration by the authors themselves, it was futile to try to improve on their work. Further examples are mentioned in the notes. It is the very nature of a survey such as this one to be inspired largely by other people's works. We hope that we did not distort or caricature them.

A book might give the illusion, especially to students, that some knowledge has become definitive and that the authors understand every part of it. This is a completely false view. No one can really fully master even his own subject, and this is luckily a source of progress. It is in the process of learning, of objecting, of finding misprints and errors, in rediscovering for oneself, that one gets the real benefits. It is very likely that, in spite of our care, many errors have crept in here and there. We welcome gladly comments and criticisms.

It was very hard to keep uniform notation throughout the text, even sometimes in the same chapter. This is a standard difficulty, especially when traditional notation in a given domain comes into conflict with those used in another one, and a compromise is necessary. We hope that this will not be a source of confusion for the reader.

We have added appendices which generally gather material in very concise form. They should be supplemented by further reading. For instance appendix C of chapter 9 is obviously insufficient to describe finite and infinite Lie algebras and their representations. This appendix is, rather, meant to induce the interested reader to study the subject further. This is also the nature of several sections where the degree of mathematical sophistication seems to increase beyond the standard background, reflecting recent trends. It was felt difficult to omit these developments but also impossible to give a proper complete introduction.

Included in small type here and there are comments, exercises and short complements ... It was felt inappropriate to develop a scholarly set of problems. In this respect the whole text can be read as a problem book.

One of the "heroes" of the whole subject of statistical physics, in one guise or another, is still to this day our old friend the Ising model. We keep a few bottles of good old French wine for the lucky person who solves it in three dimensions. It would seem appropriate to create in the theoretical physics community a prize for its solution, analogous to the one founded at the beginning of the century for the proof of Fermat's theorem. Both subjects have a similar flavour, being elementary to formulate. While it is to be presumed that the answer itself is to a large extent inessential, they motivated creative efforts (and still do) which go largely beyond the goal of solving the problem itself.

Among the many books which either overlap or amply complement the present one, the foremost are of course those in the series edited by C. Domb and M.S. Green and now J. Lebowitz, entitled *Phase transitions and critical phenomena* and published through the years by Academic Press (New York). We freely refer to this series in the notes. Let us also quote here a few others, again without pretence at exhaustivity. On the statistical side, K. Huang, *Statistical mechanics*, Wiley, New York (1963); H.E. Stanley, *In-*

troduction to phase transitions and critical phenomena, Oxford University Press (1971); S.K. Ma, *Modern theory of critical phenomena*, Benjamin, New York (1976) and *Statistical mechanics*, World Scientific, Singapore (1985); D.J. Amit, *Field theory, the renormalization group and critical phenomena*, 2nd edition, World Scientific, Singapore (1984).

Books on the path integral approach to field theory are by now numerous. Among them, the classical one is R.P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw Hill, New York (1965). Further aspects are covered in C. Itzykson and J.-B. Zuber, *Quantum field theory*, McGraw Hill, New York (1980); P. Ramond, *Field theory, a modern primer*, Benjamin/Cummings, Reading, Mass. (1981); J. Glimm and A. Jaffe, *Quantum physics*, Springer, New York (1981). To fill some gaps on other developments in field theory, see S. Coleman, *Aspects of symmetries*, Cambridge University Press (1985); S. Treiman, R. Jackiw, B. Zumino, E. Witten *Current algebra and anomalies*, World Scientific, Singapore (1985), and to learn about integrable systems, R. Baxter *Exactly solved models in statistical mechanics*, Academic Press, New York (1982), and M. Gaudin *La fonction d'onde de Bethe*, Masson, Paris (1983). Of course, many more books are mentioned in the notes. We are also aware that several important texts are either in preparation or will appear in the near future.

Our knowledge of English remains to this day very primitive and we apologize for our cumbersome use of a foreign language. This lack of fluency has prevented us of any attempt at humour which would have been sometimes more than welcome.

We would have never undertaken writing, were it not for the teaching opportunities that we were given by several universities and schools. One of the authors (C.I.) is grateful to his colleagues from the "Troisième cycle de Suisse Romande" in Lausanne and Geneva, from the "Département de Physique de l'Université de Louvain La Neuve" and from the "Troisième cycle de physique théorique" in Marseille for giving him the possibility to teach what became parts of this text, as well as to the staff of these institutions for providing secretarial help in preparing a French unpublished manuscript. The other author (J.M.D.) acknowledges similar opportunities afforded by the "Troisième cycle de physique théorique" in Paris.

The final and certainly most pleasant duty is, of course, to thank all those, friends, colleagues, collaborators, students and secretaries who have helped us through the years. A complete list should include all the members of the Saclay "Service de physique théorique", together with its numerous visitors and the members of the many departments, institutions and meetings which offered us generous hospitality and stimulation.

Particular thanks go to our very long time friends and colleagues R. Balian, M. Bander, M. Bauer, D. Bessis, E. Brézin, A. Cappelli, A. Coste, F. David, J. des Cloizeaux, C. De Dominicis, E. Gardner, M. Gaudin, B. Derrida, J.-M. Luck, A. Morel, P. Moussa, H. Orland, G. Parisi, Y. Pomeau, R. Lacaze, H. Saleur, R. Stora, J. Zinn-Justin, and J.-B. Zuber for friendly collaborations, endless discussions and generous advice. The final form of the manuscript owes a great deal to Dany Bunel. Let her receive here our warmest thanks for her tireless help. We are also very grateful to M. Porneuf and to the documentation staff, M. Féron, J. Delouvrier and F. Chétivaux.

Last but not least, we thank the Commissariat à l'Energie Atomique, the Institut de Recherche Fondamentale and the Service de Physique Théorique for their support.

Saclay, 1988

1

FROM BROWNIAN MOTION TO EUCLIDEAN FIELDS

It may seem surprising to start our study with a description of Brownian motion. However, this offers an interesting introduction to the concept of Euclidean quantum field, and an intuitive understanding of the role of dimensionality. The effective (or Hausdorff) dimension two of Brownian curves is particularly significant. It means that two such curves fail to intersect, hence to interact, in dimension higher than four. This is illustrated in the first section of this chapter, which also discusses the transition from a discrete to a continuous walk. A similar analysis for interacting fields, pioneered by K. Symanzik, is presented in the second section. It is related to strong coupling, or high temperature, expansions, to be studied later, in particular in chapter 6 of this volume and chapter 7 of volume 2. The introduction of n -component fields provides the means to incorporate "self-avoiding" walks in the limit $n \rightarrow 0$. We conclude this chapter with an analysis of elementary one-dimensional systems. This enables us to introduce the useful concept of transfer matrix.

1.1 Brownian motion

1.1.1 Random walks

We begin with a discussion of random walks on a regular, infinite lattice in d -dimensional Euclidean space. Each site has q neighbours, where q is called the *coordination number* of the lattice. At regular time intervals, separated by an amount $\Delta t = 1$, a walker jumps from one site towards a neighbouring one, chosen at random. The probability of landing on any particular adjacent site is $1/q$. Successive jumps are considered to be statistically independent events. We choose a (hyper)-cubic lattice generated

The final and certainly most pleasant duty is, of course, to thank all those, friends, colleagues, collaborators, students and secretaries who have helped us through the years. A complete list should include all the members of the Saclay "Service de physique théorique", together with its numerous visitors and the members of the many departments, institutions and meetings which offered us generous hospitality and stimulation.

Particular thanks go to our very long time friends and colleagues R. Balian, M. Bander, M. Bauer, D. Bessis, E. Brézin, A. Cappelli, A. Coste, F. David, J. des Cloizeaux, C. De Dominicis, E. Gardner, M. Gaudin, B. Derrida, J.-M. Luck, A. Morel, P. Moussa, H. Orland, G. Parisi, Y. Pomeau, R. Lacaze, H. Saleur, R. Stora, J. Zinn-Justin, and J.-B. Zuber for friendly collaborations, endless discussions and generous advice. The final form of the manuscript owes a great deal to Dany Bunel. Let her receive here our warmest thanks for her tireless help. We are also very grateful to M. Porneuf and to the documentation staff, M. Féron, J. Delouvrier and F. Chétivaux.

Last but not least, we thank the Commissariat à l'Energie Atomique, the Institut de Recherche Fondamentale and the Service de Physique Théorique for their support.

Saclay, 1988

1

FROM BROWNIAN MOTION TO EUCLIDEAN FIELDS

It may seem surprising to start our study with a description of Brownian motion. However, this offers an interesting introduction to the concept of Euclidean quantum field, and an intuitive understanding of the role of dimensionality. The effective (or Hausdorff) dimension two of Brownian curves is particularly significant. It means that two such curves fail to intersect, hence to interact, in dimension higher than four. This is illustrated in the first section of this chapter, which also discusses the transition from a discrete to a continuous walk. A similar analysis for interacting fields, pioneered by K. Symanzik, is presented in the second section. It is related to strong coupling, or high temperature, expansions, to be studied later, in particular in chapter 6 of this volume and chapter 7 of volume 2. The introduction of n -component fields provides the means to incorporate "self-avoiding" walks in the limit $n \rightarrow 0$. We conclude this chapter with an analysis of elementary one-dimensional systems. This enables us to introduce the useful concept of transfer matrix.

1.1 Brownian motion

1.1.1 Random walks

We begin with a discussion of random walks on a regular, infinite lattice in d -dimensional Euclidean space. Each site has q neighbours, where q is called the *coordination number* of the lattice. At regular time intervals, separated by an amount $\Delta t = 1$, a walker jumps from one site towards a neighbouring one, chosen at random. The probability of landing on any particular adjacent site is $1/q$. Successive jumps are considered to be statistically independent events. We choose a (hyper)-cubic lattice generated

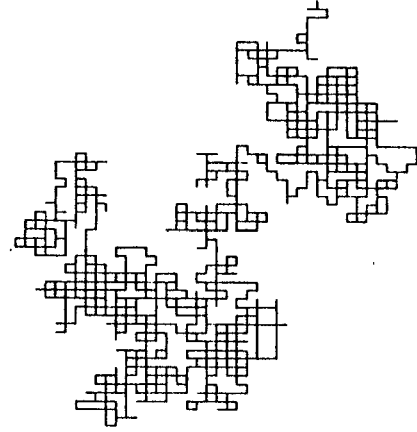


Fig. 1 A two-dimensional random walk of 2000 steps.

by d orthonormal vectors $e_{(1)}, \dots, e_{(d)}$ such that $e_{(\mu)} \cdot e_{(\nu)} = \delta_{\mu\nu}$. Sites are located at $x = x^\mu e_{(\mu)}$, where the coordinates x^μ are integers. For this lattice, $q = 2d$. Figure 1 shows such a two-dimensional random walk of 2000 steps. Note the presence of large patches where the walker visited all sites.

Let us determine the conditional probability $P(x_1, t_1; x_0, t_0)$ for the walker to be on the site x_1 at time t_1 , knowing that his initial position was x_0 at time t_0 . For $t_1 = t_0$, we have

$$P(x_1, t_0; x_0, t_0) = \delta_{x_1, x_0} \quad (1)$$

The notation δ_{x_1, x_0} is shorthand for the product $\prod_{\mu=1}^d \delta_{x_1^\mu, x_0^\mu}$. The probability P is defined for $t_1 \geq t_0$ and depends only on the differences $t_1 - t_0$, $x_1 - x_0$, because of the translational invariance in both time and space. At fixed t_1 , the normalization condition reads

$$\sum_{x_1} P(x_1, t_1; x_0, t_0) = 1 \quad (2)$$

and each probability P is either positive or zero.

The discrete formulation depends strongly on the choice of lattice. For instance, if $t_1 - t_0$ is even (odd), then so is the sum of coordinates of $x_1 - x_0$. However, we are only interested in those

$t = 0$	$t = 1$	$t = 2$	$t = 3$
			1
	1	1	3 · 3
1	1 · 1	2 · 2	3 · 9 · 3
	1	1 · 4 · 1	1 · 9 · 9 · 1
		2 · 2	3 · 9 · 3
		1	3 · 3
			1
1	4	16	64

Fig. 2 Relative probabilities for a two-dimensional random walk on a square lattice. The normalization factor is 4^{-t} .

asymptotic properties which are independent of the particular lattice structure chosen.

A recurrence relation between the probabilities P at successive times t and $t+1$ follows from the fact that the walker can reach the point x only if, one unit of time before, he was on a neighbouring site $x' = x \pm e_{(\mu)}$

$$P(x, t+1; x_0, t_0) = \frac{1}{2d} \sum_{x' \text{ neighbour of } x} P(x', t; x_0, t_0) \quad (3)$$

This relation generalizes Pascal's construction for the binomial coefficients, which corresponds to $d = 1$. Now P is completely determined by equations (1) to (3). Figure 2 illustrates the two-dimensional case.

Equation (3) can be rewritten, using a discretized version, Δ_r , of the Laplacian operator Δ ,

$$\Delta_r f(x) = \frac{1}{2d} \sum_{\mu=1}^d [f(x + e_{(\mu)}) + f(x - e_{(\mu)}) - 2f(x)] \quad (4)$$

as

$$P(x, t+1; x_0, t_0) - P(x, t; x_0, t_0) = \Delta_r P(x, t; x_0, t_0) \quad (5)$$

This is a finite difference approximation to the diffusion equation in continuous space

$$\left(\frac{\partial}{\partial t} - \Delta \right) P = 0 \quad (6)$$

A Fourier transform allows one to solve equations (3) or (5). Writing

$$P(\mathbf{x}, t; \mathbf{x}_0, t_0) = \int_{-\pi}^{\pi} \frac{d^d \mathbf{k}}{(2\pi)^d} e^{i\mathbf{k} \cdot \mathbf{x}} \tilde{P}(\mathbf{k}, t) \quad (7)$$

we deduce

$$\tilde{P}(\mathbf{k}, t + i) = \frac{1}{d} \sum_{\mu=1}^d \cos k_{\mu} \tilde{P}(\mathbf{k}, t), \quad \text{with } \tilde{P}(\mathbf{k}, t_0) = e^{-i\mathbf{k} \cdot \mathbf{x}_0}$$

as follows from the initial condition at $t = t_0$. Hence the solution is

$$P(\mathbf{x}_1, t_1; \mathbf{x}_0, t_0) = \int_{-\pi}^{\pi} \frac{d^d \mathbf{k}}{(2\pi)^d} e^{i\mathbf{k} \cdot (\mathbf{x}_1 - \mathbf{x}_0)} \left(\frac{1}{d} \sum_{\mu} \cos k_{\mu} \right)^{t_1 - t_0} \quad (8)$$

We want to obtain the asymptotic properties of this solution at large distance with respect to the lattice spacing and for long times. It is convenient to rescale distances and time using a lattice spacing a rather than 1, and a time interval τ rather than 1. Now $\mathbf{e}_{(\mu)} \cdot \mathbf{e}_{(\nu)} = a^2 \delta_{\mu\nu}$. Performing the substitutions $t \rightarrow t/\tau$, $\mathbf{x} \rightarrow \mathbf{x}/a$ and $\mathbf{k} \rightarrow a\mathbf{k}$, equation (8) becomes

$$P(\mathbf{x}_1 - \mathbf{x}_0, t_1 - t_0) = a^d \int_{-\pi/a}^{\pi/a} \frac{d^d \mathbf{k}}{(2\pi)^d} e^{i\mathbf{k} \cdot (\mathbf{x}_1 - \mathbf{x}_0)} \left(\frac{1}{d} \sum_{\mu} \cos ak_{\mu} \right)^{(t_1 - t_0)/\tau} \quad (9)$$

We now take the limit as a and τ go to zero, keeping the distances and time intervals fixed. We consider a volume Δx around $\mathbf{x}$ which is large with respect to the elementary lattice volume a^d , but which is also sufficiently small to ensure that P remains nearly constant within Δx ; this last requirement is fulfilled if $(t_1 - t_0)/\tau$ is also large. This permits a probability density $p = P/a^d$ to be defined as

$$\begin{aligned} p(\mathbf{x}_1 - \mathbf{x}_0, t_1 - t_0) \Delta x &= \sum_{\mathbf{x}'_1 \in \mathbf{x}_1 + \Delta x} P(\mathbf{x}'_1 - \mathbf{x}_0; t_1 - t_0) \\ &\approx \frac{\Delta x}{a^d} P(\mathbf{x}_1 - \mathbf{x}_0, t_1 - t_0) \end{aligned} \quad (10)$$

and thus

$$p(\mathbf{x}_1 - \mathbf{x}_0, t_1 - t_0) = \lim_{a, \tau \rightarrow 0} \int_{-\pi/a}^{\pi/a} \frac{d^d \mathbf{k}}{(2\pi)^d} e^{i\mathbf{k} \cdot (\mathbf{x}_1 - \mathbf{x}_0)} \left(\frac{1}{d} \sum_{\mu} \cos ak_{\mu} \right)^{(t_1 - t_0)/\tau} \quad (11)$$

This limit is nontrivial only when a and τ vanish in such a way that the ratio τ/a^2 is kept fixed, as shown by the expansion of the cosine

$$\left(\frac{1}{d} \sum_{\mu} \cos ak_{\mu} \right)^{(t_1 - t_0)/\tau} = \left(1 - \frac{a^2}{2d} k^2 + \dots \right)^{(t_1 - t_0)/\tau} \rightarrow e^{-(t_1 - t_0)k^2} \quad (12)$$

in which the time scale has been fixed using

$$\tau = \frac{1}{2d} a^2 \quad (13)$$

Hence

$$\begin{aligned} p(\mathbf{x}_1 - \mathbf{x}_0, t_1 - t_0) &= \int_{-\infty}^{+\infty} \frac{d^d \mathbf{k}}{(2\pi)^d} \exp[-(t_1 - t_0)k^2 + i(\mathbf{x}_1 - \mathbf{x}_0) \cdot \mathbf{k}] \\ &= \frac{1}{[4\pi(t_1 - t_0)]^{d/2}} \exp\left(-\frac{(\mathbf{x}_1 - \mathbf{x}_0)^2}{4(t_1 - t_0)}\right) \end{aligned} \quad (14)$$

This is the well-known kernel of the diffusion equation in continuous space. It is positive, symmetric, and satisfies

$$\int d^d \mathbf{x} p(\mathbf{x}, t; \mathbf{x}_0, t_0) = 1 \quad (15a)$$

$$\lim_{t_1 \rightarrow t_0} p(\mathbf{x}_1, t_1; \mathbf{x}_0, t_0) = \delta^d(\mathbf{x}_1 - \mathbf{x}_0) \quad (15b)$$

$$\left(\frac{\partial}{\partial t} - \Delta \right) p(\mathbf{x}, t; \mathbf{x}_0, t_0) = 0 \quad (15c)$$

$$t_2 > t_1 > t_0 \Rightarrow$$

$$\int d^d \mathbf{x}_1 p(\mathbf{x}_2, t_2; \mathbf{x}_1, t_1) p(\mathbf{x}_1, t_1; \mathbf{x}_0, t_0) = p(\mathbf{x}_2, t_2; \mathbf{x}_0, t_0) \quad (15d)$$

Condition (15a) is the conservation law for the probabilities (this is the continuous counterpart of equation (2)), while (15b) describes the initial conditions. Equation (15c) is the diffusion equation. Finally, (15d) expresses the obvious fact that the walker was certainly somewhere at an intermediate time t_1 . This last relation, characteristic of a Markov process, is compatible with the convolution properties of Gaussian integrals and will be used later

to construct a path integral representation which is the continuous counterpart of a similar property on the lattice.

The continuous diffusion law (14) is isotropic and translationally invariant, whereas the discrete version only presents the cubic symmetries. In the limiting procedure, we fixed the physical time intervals $t_1 - t_0$ and distances $x_1 - x_0$ and considered the case where the spatial and time steps a and τ vanished, with the ratio τ/a^2 kept fixed. To be quite rigorous, $t_1 - t_0$ and the components of $x_1 - x_0$ should be multiples of τ and a respectively. Of course, we could have proceeded differently, taking $x_1 - x_0$ and $t_1 - t_0$ large with respect to the spacings. The view adopted above, namely to let a and τ go to zero, is comparable to the ultraviolet limit in field theory, where a cutoff factor (here $1/a$, a natural scale for momenta k) tends to infinity, whereas physical (measurable) quantities are kept fixed. We also used here a cutoff τ^{-1} for frequencies, with $\tau^{-1} \sim a^{-2}$. This will be later interpreted by considering that the typical Brownian curve has a "dimension" 2; this is called the *Hausdorff dimension*. When the curve is followed at a constant speed a/τ , the typical distance behaves as $|x_1 - x_0| \sim (t_1 - t_0)^{1/2}$. In other words, if a Brownian curve is contained almost certainly in a volume Ω and if its length $t_1 - t_0$ is multiplied by λ , the volume subsequently needed to enclose it will be $\lambda^{d/2}\Omega$. Defining an exponent ν characteristic of the end-to-start distance by

$$|x_1 - x_0| \sim (t_1 - t_0)^\nu \quad (16)$$

the value of ν for Brownian motion is therefore

$$\nu = \frac{1}{2} \quad (17)$$

This is a first example of a so-called *critical exponent*. Much more will be said on this subject in the following chapters.

The result (16)-(17) is explained in an elementary way by the central limit theorem of probability theory. Consecutive steps are uncorrelated random variables, and the mean quadratic sum of distances behaves as the number of terms, i.e. the number of steps.

Verify that a random walk on a two-dimensional triangular lattice yields the same continuous limit (14) as on a square lattice.

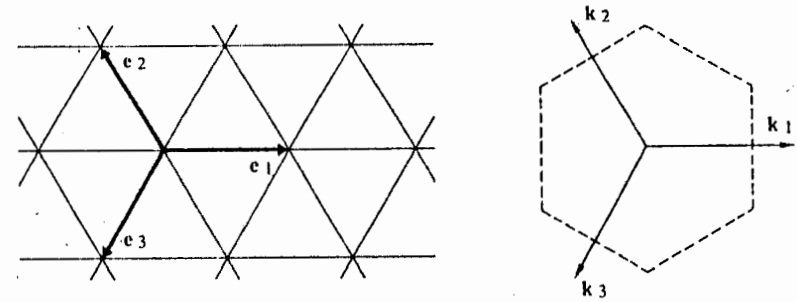


Fig. 3 The triangular lattice and its first Brillouin zone.

On a triangular lattice, each site has $q = 6$ neighbours. It is convenient, as shown in figure 3, to use a redundant basis of 3 vectors $e_{(i)}$ such that $\sum e_{(i)} = 0$, $e_{(i)}^2 = 1$, $e_{(i)} \cdot e_{(j)} = -\frac{1}{2}$ ($i \neq j$). The conditional probability that the walker is on site x at time $t+1$ is expressed as

$$P(x, t+1) = \frac{1}{6} \sum_{i=1}^3 P(x + e_{(i)}, t) + P(x - e_{(i)}, t)$$

Lattice sites are labelled by three integers $\{n_i\}$, with $x = \sum n_i e_{(i)}$. A constant can be added to the n_i 's without changing x . Similarly, the conjugate quantity $k = \sum k_i e_{(i)}$ is characterized by the three numbers k_i , on which we impose the additional constraint $\sum k_i = 0$. Therefore,

$$k \cdot x = \frac{3}{2} \sum k_i n_i, \quad k^2 = \frac{3}{2} (k_1^2 + k_2^2 + k_3^2)$$

The integration domain for k is the hexagon $|k_i| < \pi$, the area of which is $2\pi^2\sqrt{3}$. Consequently,

$$P(x, t) = \int_{\text{hexagon}} \frac{d^2 k}{2\pi^2 \sqrt{3}} e^{ik \cdot x} \left[\frac{1}{3} \left(\cos \frac{3}{2} k_1 + \cos \frac{3}{2} k_2 + \cos \frac{3}{2} k_3 \right) \right]^t$$

Changing the scale according to $x \rightarrow x/a$, $k \rightarrow ka$, $t \rightarrow t/\tau$, $t = a^2/8$, we obtain in the limit $a \rightarrow 0$ the same result as before

$$p(x, t) = \int \frac{d^2 k}{(2\pi)^2} e^{-tk^2 + ik \cdot x} = \frac{1}{4\pi t} e^{-x^2/4t}$$

using the fact that there are $2\Delta x/a^2\sqrt{3}$ lattice sites in an area Δx . This confirms the universality of the isotropic limit.

In arbitrary dimension the previous construction generalizes into the projection of a $(d+1)$ -dimensional hypercubic lattice on the hyperplane $x_1 + \dots + x_{d+1} = 0$. Its dual is the cross section by this hyperplane. When $d = 3$, this gives the body centered (b.c.c.) cubic lattice dual to the face centered cubic (f.c.c.) lattice.

The mean time spent on a given site (which may be either finite or infinite) leads to the Green function

$$G(x_1 - x_0) = \sum_{t_1=t_0}^{\infty} P(x_1, t_1; x_0, t_0) = \int_{-\pi}^{\pi} \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot (x_1 - x_0)}}{1 - \frac{1}{d} \sum_{\mu} \cos k_{\mu}} \quad (18)$$

which obeys

$$G(x_1 - x_0) = \delta_{x_1, x_0} + \frac{1}{2d} \sum_{\mu} G(x_1 + e_{(\mu)} - x_0) + G(x_1 - e_{(\mu)} - x_0) \quad (19)$$

or, equivalently,

$$\Delta_r G(x_1 - x_0) = -\delta_{x_1, x_0} \quad (20)$$

This relation justifies the name of *Green function* for the operator $-\Delta_r$; G is symmetric, positive, and well defined for $d > 2$. When $d = 1$ or 2 , the mean time spent on a given site is infinite, as can be seen from the infrared divergence of the integral (18). The integrand indeed behaves as $d^d k/k^2$ when k is small. However, one can introduce a subtracted quantity

$$\begin{aligned} G^s(x_1 - x_0) &= G(x_1 - x_0) - G(0) \\ &= \sum_{t_1=t_0}^{\infty} P(x_1, t_1; x_0, t_0) - P(x_0, t_1; x_0, t_0) \\ &= \int \frac{d^d k}{(2\pi)^d} \frac{\cos k \cdot (x_1 - x_0) - 1}{1 - \frac{1}{d} \sum_{\mu} \cos k_{\mu}} \end{aligned} \quad (21)$$

which is now well-defined whatever the value of d . It still satisfies equation (20), since we have only formally subtracted a (possibly infinite) constant. However, the positivity property is lost, as shown by the one-dimensional result

$$d = 1 \quad G^s(x_1 - x_0) = -|x_1 - x_0| \quad (22)$$

We define the generating function for the probabilities

$$G(x_1 - x_0, \lambda) = \lambda \sum_{t_1=t_0}^{\infty} \lambda^{t_1-t_0} P(x_1, t_1; x_0, t_0) \quad (23)$$

which coincides for $\lambda = 1$ with the Green function previously considered. When $|\lambda| < 1$, the sum is always convergent and has an integral representation

$$G(x_1 - x_0, \lambda) = \int \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot (x_1 - x_0)}}{\lambda^{-1} - \frac{1}{d} \sum_{\mu} \cos k_{\mu}} \quad (24)$$

It satisfies the relation

$$[-\Delta_r + (\lambda^{-1} - 1)] G(x, \lambda) = \delta_{x,0} \quad (25)$$

This function is the analog of the massive propagator in particle physics. Introducing again the lattice spacing $x \rightarrow x/a$, $k \rightarrow ka$, the continuous limit is obtained for $1/\lambda \equiv 1 + m^2 a^2/2d$, in the limit $a \rightarrow 0$

$$g(x, m^2) = \lim_{a \rightarrow 0} \frac{1}{2da^{d-2}} G\left(\frac{x}{a}, \lambda\right) = \int_{-\infty}^{+\infty} \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot x}}{k^2 + m^2} \quad (26)$$

1.1.2 The sum over paths

The properties of Brownian motion can be understood in terms of sums over paths. This seems quite natural in the present context, and will later be the origin of fruitful generalizations. This interpretation follows directly from relation (3): a path of length $t+1$ is obtained by adding one step to a path of length t . Iterating this relation leads to

$$P(x_1, t_1; x_0, t_0) = \frac{\text{Number of paths joining } x_0 \text{ to } x_1 \text{ with } t_1 - t_0 \text{ steps}}{\text{Number of paths originating from } x_0 \text{ with } t_1 - t_0 \text{ steps}} \quad (27)$$

The denominator is equal to $(2d)^{t_1-t_0} = q^{t_1-t_0}$. Another derivation follows from formula (8), when one replaces cosines by exponentials, expands the integrand and then integrates over k term by term. This equality is interpreted by saying that all paths covered during the time $t_1 - t_0$ are equally probable. Formula

(27) can also be rewritten as

$$P(\mathbf{x}_1, t_1; \mathbf{x}_0, t_0) = \frac{1}{(2d)^{t_1-t_0}} \sum_{\substack{\omega(t) \\ t=1, \dots, t_1-t_0-1}} \prod_{t=0}^{t_1-t_0-1} \delta_{1, \sum_{\mu=1}^d |x^\mu(t+1) - x^\mu(t)|} \quad (28)$$

which states explicitly that elementary steps have unit length, and $\mathbf{x}(0) = \mathbf{x}_0$, $\mathbf{x}(t_1 - t_0) = \mathbf{x}_1$. However, this expression does not yield easily the continuous limit. Let us use a more compact notation. Let ω be a path joining $\mathbf{x}_0$ to $\mathbf{x}_1$. By definition, this is a set of lattice links such that its boundary $\partial\omega$, defined as the set of sites shared by an odd number of links of ω , is $\{\mathbf{x}_0, \mathbf{x}_1\}$. To be precise, the definition of ω implies a prescription on how to describe it as a function of time. Equation (27) is thus rewritten as

$$P(\mathbf{x}_1, t_1; \mathbf{x}_0, t_0) = \sum_{\substack{\omega, \partial\omega = \{\mathbf{x}_0, \mathbf{x}_1\} \\ |\omega| = t_1 - t_0}} \left(\frac{1}{2d}\right)^{|\omega|} \quad (29)$$

where $|\omega|$ is the number of links of ω . These expressions are easily generalized to the Green function, and, from (23) with $\lambda < 1$,

$$G(\mathbf{x}_1 - \mathbf{x}_0, \lambda) = \lambda \sum_{\omega, \partial\omega = \{\mathbf{x}_0, \mathbf{x}_1\}} \left(\frac{\lambda}{2d}\right)^{|\omega|} \quad (30)$$

The mean time spent at $\mathbf{x}_1$ is obtained in the limit $\lambda \rightarrow 1$.

This formula would suggest that its continuous counterpart implies a sum over weighted paths, where the weight should be an exponential of the form $\exp(-A|\omega|)$. With an abusive extension of the language used in quantum mechanics, the argument of the exponential (up to its sign) is called the *action*. In equilibrium statistical mechanics, it would be identified (up to a factor $\beta = 1/kT$) with the energy of a classical configuration (here a path going from $\mathbf{x}_0$ to $\mathbf{x}_1$). In any case, the correct procedure to reach the continuous asymptotic limit is to group together a very large number of infinitesimal terms when the spacing a goes to zero. This results in a very different form of the action. The preceding formulae are not really adapted to this operation. We shall rather proceed as follows.

Let us remark that (with $a^2 = 2d\tau$)

$$\begin{aligned} g(\mathbf{x}, m^2) &= \lim_{a \rightarrow 0} \frac{1}{2da^{d-2}} G\left(\frac{\mathbf{x}}{a}, \lambda \equiv \left(1 + \frac{m^2 a^2}{2d}\right)^{-1}\right) \\ &= \lim_{a \rightarrow 0} \frac{1}{2da^{d-2}} \sum_{t/\tau=0}^{\infty} \left(1 + \frac{m^2 a^2}{2d}\right)^{-t/\tau} P\left(\frac{\mathbf{x}}{a}, \frac{t}{\tau}; 0, 0\right) \\ &= \lim_{a \rightarrow 0} \frac{a^2}{2d} \sum_{t/\tau=0}^{\infty} \exp\left(-\frac{m^2 a^2 t}{2d\tau}\right) p(\mathbf{x}, t) \\ &= \int_0^{\infty} dt e^{-m^2 t} p(\mathbf{x}, t) \end{aligned} \quad (31)$$

One can also obtain this result directly by combining equations (26) and (14). We recall the convolution property (15d)

$$p(\mathbf{x}_1, t_1; \mathbf{x}_0, t_0) = \int d^d \mathbf{x} p(\mathbf{x}_1, t_1; \mathbf{x}, t) p(\mathbf{x}, t; \mathbf{x}_0, t_0) \quad t_0 < t < t_1 \quad (32)$$

Dividing $t_1 - t_0$ into an arbitrary number of intervals and using the expression (14), we write

$$\begin{aligned} p(\mathbf{x}_f, t_f; \mathbf{x}_i, t_i) &= \int \prod_{j=1}^{n-1} \frac{d^d \mathbf{x}_j}{[4\pi(t_{j+1} - t_j)]^{d/2}} \frac{1}{[4\pi(t_1 - t_i)]^{d/2}} \\ &\quad \times \exp\left[-\frac{1}{4} \sum_{j=0}^{n-1} \frac{(\mathbf{x}_{j+1} - \mathbf{x}_j)^2}{t_{j+1} - t_j}\right] \end{aligned} \quad (33)$$

(with $\mathbf{x}_0 = \mathbf{x}_i$, $\mathbf{x}_n = \mathbf{x}_f$). This gives a precise meaning to the symbolic formula

$$p(\mathbf{x}_f, t_f; \mathbf{x}_i, t_i) = \int_{\substack{\mathbf{x}(t_i) = \mathbf{x}_i \\ \mathbf{x}(t_f) = \mathbf{x}_f}} \mathcal{D}\mathbf{x}(t) \exp\left[-\frac{1}{4} \int_{t_i}^{t_f} dt \dot{\mathbf{x}}^2\right] \quad (34)$$

where $\mathcal{D}\mathbf{x}(t)$ implies integrals on the positions at intermediate times, normalized as indicated in (33). Equation (31) leads to a similar relation for $g(\mathbf{x}, m^2)$

$$g(\mathbf{x}, m^2) = \int_0^{\infty} dt \int_{\substack{\mathbf{x}(0)=0 \\ \mathbf{x}(t)=\mathbf{x}}} \mathcal{D}\mathbf{x}(t') \exp\left[-\int_0^t dt' (m^2 + \frac{1}{4} \dot{\mathbf{x}}(t')^2)\right] \quad (35)$$

These path integrals in the continuum – which, despite their symbolic appearance, have a rigorous meaning indicated by formula (33) – seem quite different from their discrete counterparts. The action is now $\int_0^t dt' (m^2 + \frac{1}{4} \dot{x}(t')^2)$, with a distance scale defined by the inverse of m . This form comes from the short-distance behaviour of the continuous limit $p(x, t) \sim \exp(-x^2/4t)$, and not from its discrete partner P . Indeed, a small increment in the continuum is the sum of an infinite number of lattice steps. It should be pointed out that, in the continuous notation, the length of a path would be $\int_0^t dt' \sqrt{\dot{x}(t')^2}$. Finally, the discrete expressions were independent from a given parametrization of the path. This is not the case in the action involved in (35).

1.1.3 The dimension two of Brownian curves

In three-dimensional space, surfaces generically intersect along curves. In four dimensions, they meet only at isolated points. They do not intersect for higher dimensions, apart from exceptional cases. From this point of view, we will see in this section that the characteristics of Brownian motion look more similar to those of a surface than to those of a curve. These properties will be the origin of an intuitive understanding of some fundamental features of critical phenomena.

To be more specific, we consider again a random walk on a cubic lattice, with the probability of being on site x at time t given by equation (8). What is the probability for the walker to have visited the site x during the time interval t , assuming he was at the origin at time $t = 0$? If $x = 0$, this is conventionally the probability of returning to the origin, excluding the starting time. Note that the answer is *not* the sum over t of $P(x, t)$; as we already pointed out, this sum is not a probability, but the mean time the walker spends at site x . However, there are close relations between these two quantities.

Let $P_i(x, t)$ be the probability of being on site x at time t for the i th time. As we deal with exclusive events, it is clear that

$$P(x, t) = \delta_{x,0} \delta_{t,0} + \sum_{i=1}^{\infty} P_i(x, t) \quad (36)$$

Only a finite number of terms contribute to this sum since $P_i(x, t)$ vanishes whenever $i > t$. We deduce a recurrence relation from the fact that the $(i+1)$ th visit occurs just after the i th one. Using the spatial homogeneity of the motion, this leads to

$$P_{i+1}(x, t) = \sum_{t_1+t_2=t} P_i(x, t_1) P_1(0, t_2) \quad (37)$$

Summing over all possible values $i = 1, 2, \dots$, one finds

$$P(x, t) - P_1(x, t) - \delta_{x,0} \delta_{t,0} = \sum_{t_1+t_2=t} P(x, t_1) P_1(0, t_2) - \delta_{x,0} P_1(x, t) \quad (38)$$

The required quantity, i.e. the probability $\Pi(x)$ of visiting the point x at least once, is the sum over all times of the (mutually exclusive) probabilities of visiting this point for the first time. It is convenient to introduce the generating function

$$\Pi(x, \lambda) = \sum_{t=0}^{\infty} \lambda^t P_1(x, t) \quad (|\lambda| \leq 1) \quad (39)$$

so that

$$\Pi(x) = \Pi(x, \lambda = 1) \quad (40)$$

Taking into account the definition of the propagator (33), relation (38) allows us to write

$$\lambda^{-1} G(x, \lambda) = \delta_{x,0} + (1 - \delta_{x,0}) \Pi(x, \lambda) + \lambda^{-1} G(x, \lambda) \Pi(0, \lambda) \quad (41)$$

The additional factor λ in the definition of G is the origin of the combination $\lambda^{-1} G$. Considering the two cases $x = 0$ and $x \neq 0$ in (41) separately, we obtain a new remarkable interpretation of the propagator in terms of the probability Π

$$\Pi(0, \lambda) = 1 - \frac{\lambda}{G(0, \lambda)} \quad (42)$$

$$\Pi(x, \lambda) = \frac{G(x, \lambda)}{G(0, \lambda)} \quad \text{for } x \neq 0 \quad (43)$$

If $d \leq 2$, the mean time spent on a given site is infinite, hence $G(x, \lambda) = \infty$, $\Pi(x, \lambda) = 1$ in the limit $\lambda \rightarrow 1$ for all x . As was previously stated, the walker visits any point with unit probability (Pólya, 1921), according to the well-known adage *all roads lead to Rome*. When $d > 2$, the probability to return to the origin

$\Pi(0) = 1 - G(0, 1)^{-1}$ is positive, less than unity and decreases as $1/d$ for large d . Its expansion in $1/d$ follows from equation (18)

$$\begin{aligned} G(0, 1) = & 1 + \frac{1}{2d} + 3\left(\frac{1}{2d}\right)^2 + 12\left(\frac{1}{2d}\right)^3 + 60\left(\frac{1}{2d}\right)^4 + \\ & + 355\left(\frac{1}{2d}\right)^5 + 2380\left(\frac{1}{2d}\right)^6 + 17430\left(\frac{1}{2d}\right)^7 + \\ & + 134190\left(\frac{1}{2d}\right)^8 + \mathcal{O}\left(\left(\frac{1}{2d}\right)^9\right) \end{aligned} \quad (44)$$

This states that, for d large, the time spent at the origin tends towards the time interval between two jumps. Correspondingly, the probability of returning to the origin is very small

$$\begin{aligned} \Pi(0) = & \left(\frac{1}{2d}\right) + 2\left(\frac{1}{2d}\right)^2 + 7\left(\frac{1}{2d}\right)^3 + 35\left(\frac{1}{2d}\right)^4 + 215\left(\frac{1}{2d}\right)^5 + \\ & + 1501\left(\frac{1}{2d}\right)^6 + 11354\left(\frac{1}{2d}\right)^7 + 88978\left(\frac{1}{2d}\right)^8 + \mathcal{O}\left(\left(\frac{1}{2d}\right)^9\right) \end{aligned} \quad (45)$$

On the other hand, in the vicinity of $d = 2$, we have

$$G(0, 1) = \frac{2}{\pi(d-2)} + \dots, \quad \Pi(0) = 1 - \frac{1}{2}\pi(d-2) + \dots \quad (46)$$

These two expressions are plotted as a function of d in figure 4.

Similarly, for a point distinct from the origin, we have $\Pi(\mathbf{x}) = 1$ whenever $d \leq 2$. If $d > 2$, this probability is less than unity and decreases with the distance to the origin. Asymptotically,

$$\Pi(\mathbf{x}) \xrightarrow{|\mathbf{x}| \rightarrow \infty} (1 - \Pi(0)) \frac{2d}{(d-2)S_d} \times \frac{1}{|\mathbf{x}|^{d-2}} \quad (47)$$

where S_d is the area of the unit sphere in d -dimensional space

$$S_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \quad (48)$$

Notice also the one-dimensional relations

$$\begin{aligned} G(0, \lambda) &= \int_{-\pi}^{\pi} \frac{dq}{2\pi} \frac{1}{1 - \lambda \cos q} = \frac{1}{\sqrt{1 - \lambda^2}} \\ \Pi(0, \lambda) &= 1 - \sqrt{1 - \lambda^2} \end{aligned} \quad (49)$$

on which it is explicitly seen that $G(0, 1) = +\infty$ and $\Pi(0) = 1$.

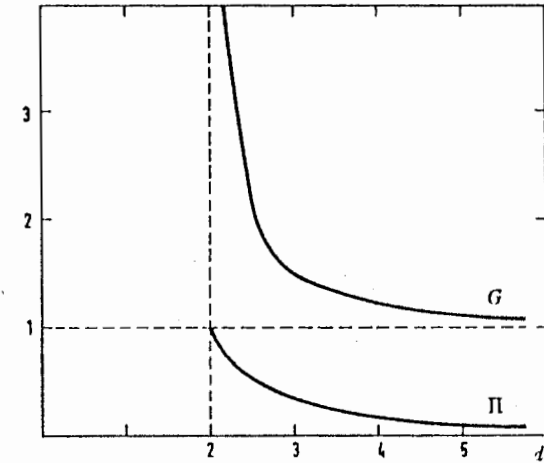


Fig. 4 $G(0, 1)$ and $\Pi(0)$ as functions of d .

The mean time elapsed before the first visit to the site $\mathbf{x}$ (which is, of course, the time elapsed before the first return to the origin for $\mathbf{x} = 0$) can also be computed (see (39)) as

$$t(\mathbf{x}) = \frac{\sum_{t=0}^{\infty} t P_1(\mathbf{x}, t)}{\Pi(\mathbf{x})} = \lambda \frac{\partial}{\partial \lambda} \ln \Pi(\mathbf{x}, \lambda) \Big|_{\lambda=1} \quad (50)$$

This time is infinite for any dimension less than or equal to four. This fact might seem rather surprising at first sight.

It is possible to give a more intuitive proof of formulae (42) and (43), using the fact that $G(\mathbf{x}, 1)$ is the mean time spent at site $\mathbf{x}$ (including the first time interval for $\mathbf{x} = 0$). Indeed, as $\Pi(0)$ is the probability of returning to the origin and $1 - \Pi(0)$ the probability of never returning, the average $G(0, 1)$ can be decomposed into terms corresponding to 0, 1, 2, ... visits with probabilities $1 - \Pi(0)$, $\Pi(0)(1 - \Pi(0))$, $\Pi^2(0)(1 - \Pi(0))$, ...; hence

$$\begin{aligned} G(0, 1) &= 1[1 - \Pi(0)] + 2\Pi(0)[1 - \Pi(0)] + 3\Pi^2(0)[1 - \Pi(0)] + \dots \\ &= [1 - \Pi(0)][1 - \Pi(0)]^{-2} \\ &= [1 - \Pi(0)]^{-1} \end{aligned} \quad (51)$$

Table I. Leading behaviour of the various quantities introduced in the text as $\lambda \rightarrow 1$ or $t \rightarrow \infty$. The d -dependent coefficient K is defined in equation (58).

	$0 < d < 2$	$d = 2$	$2 < d < 4$	$d = 4$
$G(0, \lambda)$	$K(1 - \lambda)^{d/2-1}$	$-\frac{1}{\pi} \ln(1 - \lambda)$	$G(0, 1) + K(1 - \lambda)^{d/2-1}$	$G(0, 1) + \frac{4}{\pi^2}(1 - \lambda) \ln(1 - \lambda)$
$D(\lambda)$	$K^{-1}(1 - \lambda)^{-d/2-1}$	$-\frac{\pi(1 - \lambda)^{-2}}{\ln(1 - \lambda)}$	$G^{-1}(0, 1)(1 - \lambda)^{-2} - \frac{K(1 - \lambda)^{d/2-3}}{G(0, 1)^2}$	$G^{-1}(0, 1)(1 - \lambda)^{-2} - \frac{4(1 - \lambda)^{-1}}{\pi^2 G^2(0, 1)} \ln(1 - \lambda)$
$D(t)$	$\frac{K^{-1} t^{d/2}}{\Gamma(1 + d/2)}$	$\frac{\pi t}{\ln t}$	$G^{-1}(0, 1)t - \frac{K t^{2-d/2}}{G^2(0, 1)\Gamma(3 - d/2)}$	$G^{-1}(0, 1)t + \frac{4 \ln t}{\pi^2 G^2(0, 1)}$
$E(t)$	$t - \frac{K^{-1} t^{d/2}}{\Gamma(1 + d/2)}$	$t(1 - \frac{\pi}{\ln t})$	$\Pi(0)t + \frac{K t^{2-d/2}}{G^2(0, 1)\Gamma(3 - d/2)}$	$\Pi(0)t - \frac{4 \ln t}{\pi^2 G^2(0, 1)}$
$I(t)$	$\frac{K t^{2-d/2}}{\Gamma(3 - d/2)} - t$	$t(\frac{\ln t}{\pi} - 1)$	$t(G(0, 1) - 1) + \frac{K t^{2-d/2}}{\Gamma(3 - d/2)}$	$t(G(0, 1) - 1) - \frac{4 \ln t}{\pi^2}$

while the fraction of sites visited several times behaves as $\Pi(0)t$, where $\Pi(0)$ decreases from 1 for $d = 2$ to 0 as d goes to infinity. The corrections to these behaviours are negligible when d is greater than four. These results are in agreement with the dimensionality two of random curves, but show that some care is required when using this concept. The behaviour $\Pi(0)t$ of $E(t)$ is intuitively understood by remarking that, all along the curve (hence providing the factor t), there is a probability $\Pi(0)$ for the path to make a loop. Since this is an overestimate of $E(t)$, the correction is negative.

As $D(t)$, $E(t)$ is of course bounded by t (recall that $G(0, 1) > 1$). The self-intersections of a Brownian curve can also be discussed by weighting a site visited n times by a factor $n(n-1)/2$ (instead of $n-1$); this quantity counts the mean number of self-intersections of the curve and reads

$$I(t) = \sum_{0 \leq t_1 < t_2 \leq t} P(0, t_2 - t_1) \quad (59)$$

Its generating function is

$$\begin{aligned} I(\lambda) &= \sum_{t=1}^{\infty} \lambda^t I(t) = \sum_{t=1}^{\infty} \lambda^t \sum_{t'=1}^t (t - t' + 1) P(0, t') \\ &= \sum_{t_1=0}^{\infty} (t_1 + 1) \lambda^{t_1} \sum_{t_2=1}^{\infty} \lambda^{t_2} P(0, t_2) \\ &= \frac{G(0, \lambda) - \lambda}{\lambda(1 - \lambda)^2} \end{aligned} \quad (60)$$

The dominant behaviour is also displayed in table I. Notice that $E(t)$ and $I(t)$ are similar in large dimensions (because triple and higher order intersections become very rare).

Finally, let us consider the intersection properties of two different Brownian curves in the continuous limit. $\mathcal{P}(\mathbf{x}_1, \mathbf{x}_2, \Lambda)$ will denote the probability for two curves starting from distinct points $\mathbf{x}_1 \neq \mathbf{x}_2$ to intersect on a nonempty set Λ . The length of both curves is assumed infinite. It is clear that

$$\mathcal{P}(\mathbf{x}_1, \mathbf{x}_2, \Lambda) \leq \sum_{\mathbf{z} \in \Lambda} \Pi(\mathbf{z} - \mathbf{x}_1) \Pi(\mathbf{z} - \mathbf{x}_2) \quad (61)$$

since $\Pi(\mathbf{x})$ is the probability, starting from the origin, to visit $\mathbf{x}$. This is only an inequality because the probabilities are not

independent as $\mathbf{z}$ varies. Let us now perform a dilatation by a factor $\xi \rightarrow \infty$ of the positions and of the domain Λ (with the restrictions imposed by the lattice). This allows the computation of the corresponding continuous quantity $p(\mathbf{x}_1, \mathbf{x}_2, \Lambda)$

$$\begin{aligned} p(\mathbf{x}_1, \mathbf{x}_2, \Lambda) &= \lim_{\xi \rightarrow \infty} \mathcal{P}(\xi \mathbf{x}_1, \xi \mathbf{x}_2, \xi \Lambda) \leq \lim_{\xi \rightarrow \infty} \frac{\xi^d}{G(0, 1)^2} \\ &\times \int_{\Lambda} d^d \mathbf{z} \int_{-\pi}^{\pi} \frac{d^d \mathbf{q}_1 d^d \mathbf{q}_2}{(2\pi)^{2d}} \frac{\exp \{i[\xi \mathbf{q}_1 \cdot (\mathbf{x}_1 - \mathbf{z}) + \xi \mathbf{q}_2 \cdot (\mathbf{x}_2 - \mathbf{z})]\}}{(1 - \frac{1}{d} \sum_{\mu} \cos q_1^{\mu})(1 - \frac{1}{d} \sum_{\mu} \cos q_2^{\mu})} \\ &\leq \lim_{\xi \rightarrow \infty} \left(\frac{2d}{G(0, 1)} \right)^2 \xi^{4-d} \\ &\times \int_{\Lambda} d^d \mathbf{z} \int_{-\infty}^{\infty} \frac{d^d \mathbf{q}_1 d^d \mathbf{q}_2}{(2\pi)^{2d}} \frac{\exp \{i[\mathbf{q}_1 \cdot (\mathbf{x}_1 - \mathbf{z}) + \mathbf{q}_2 \cdot (\mathbf{x}_2 - \mathbf{z})]\}}{\mathbf{q}_1^2 \mathbf{q}_2^2} \end{aligned} \quad (62)$$

The integral is convergent for $d > 2$; the important point is the purely dimensional factor ξ^{4-d} which vanishes for $d > 4$ in the desired limit. This shows directly that two Brownian paths have no chance to cross in an arbitrary region as soon as the dimension is greater than four. Conversely, for d strictly less than four, one can show that they certainly cross in the entire space (this property is evident if $d \leq 2$). Dimension four is the limiting case.

(1) (from G.F. Lawler, *Comm. Math. Phys.* **86**, 539 (1982)).

Consider a four-dimensional lattice.

(i) If two Brownian paths of length t start from two points $\mathbf{x}_1, \mathbf{x}_2$, at a distance of order $\sqrt{t}$, the probability that they intersect behaves as $1/\ln t$ for large t . Hence, in the continuous limit, two paths of infinite length starting from fixed points (at a finite distance) have no chance to intersect.

(ii) If the two paths start from the same point, the corresponding probability tends to unity as their length increases (the correction probably behaves as $(\ln t)^{-1/2}$).

(2) The number of distinct sites visited by both paths starting from the same point, with respective lengths t_1 and t_2 is $D(t_1) + D(t_2) - D(t_1 + t_2)$. If $t_1 = \alpha t$, $t_2 = \beta t$ and $t \rightarrow \infty$, this number diverges for $d \leq 4$ (for $d = 4$, this divergence is logarithmic) and remains finite for $d > 4$.

(3) In one dimension, the probability $P_t(N)$ to visit N distinct sites during the time t ($D(t) = \sum_N N P_t(N)$) is (B. Derrida)

$$P_t(N) = q_t(N) - 2q_t(N-1) + q_t(N-2)$$

with

$$q_t(N) = \frac{2}{N+1} \sum_{\substack{1 \leq q \leq N \\ q \text{ odd}}} \left(\cos \frac{q\pi}{N+1} \right)^t \cotan \frac{q\pi}{2(N+1)}$$

As $t, N \rightarrow \infty$ with $t \gg N^2$, $P_t(N)$ is exponentially small as $(8\pi^2 t^2/N^5) \exp(-\pi^2 t/2N^2)$. It is expected that the d -dimensional analogous property is $P_t(N) \sim \exp(-Kt/N^{2/d})$.

(4) In three dimensions, the mean time spent at the origin $G(0, 1)$ and the escape probability $\Pi(0, 1) = 1 - G(0, 1)^{-1}$ are finite. Construct a probabilistic proof of the following results (Watson, see M.L. Glasser, I.J. Zucker, *Proc. Natl. Acad. Sci. USA*, **74**, 1800 (1977))

$$G(0, 1) = \int_{-\pi}^{\pi} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{1}{1 - c_1 c_2 c_3} = \frac{1}{4\pi^3} \Gamma(\frac{1}{4})^4 \simeq 1.3932039297$$

for the body centered cubic lattice (b.c.c.);

$$\begin{aligned} G(0, 1) &= \int_{-\pi}^{\pi} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{1}{1 - \frac{1}{3}(c_1 c_2 + c_2 c_3 + c_3 c_1)} \\ &= \frac{9\Gamma(\frac{1}{3})^6}{2^{14/3}\pi^4} \simeq 1.3446610732 \end{aligned}$$

for the face centered cubic lattice (f.c.c.); and

$$\begin{aligned} G(0, 1) &= \int_{-\pi}^{\pi} \frac{d^3 \mathbf{k}}{(2\pi)^3} \frac{1}{1 - \frac{1}{3}(c_1 + c_2 + c_3)} \\ &= \frac{4\sqrt{6}}{\pi^2} \Gamma(\frac{1}{24}) \Gamma(\frac{5}{24}) \Gamma(\frac{7}{24}) \Gamma(\frac{11}{24}) \simeq 1.5163860591 \end{aligned}$$

for the simple cubic lattice (s.c.). In these formulae, one uses the short-hand notation $c_i = \cos k_i$.

1.2 Euclidean fields

When one performs an analytic continuation of the time to imaginary values, the Poincaré group of mechanics becomes the Euclidean group of displacements and reflections. From the

superposition principle, quantum probability amplitudes appear now as sums similar to those used in statistical mechanics. Hence an evolution problem in quantum field theory is transformed into questions relative to a static thermodynamical equilibrium, where the fundamental variables are stochastic classical fields. To give a precise and rigorous meaning to the various continuous sums appearing in this formalism, the continuous space $\mathcal{R}^d$ is replaced by a lattice. This procedure leads to the problem of examining the limit of an infinitely dense lattice. Let us recall first the free field formalism. It is based on an extensive use of Gaussian integrals and we assume that the reader is familiar with the formula

$$\int dX \exp(-\frac{1}{2} {}^T X A X + {}^T J X) = \left(\det \frac{A}{2\pi}\right)^{-1/2} \exp\left(\frac{1}{2} {}^T J A^{-1} J\right) \quad (63)$$

where X is a vector and the symmetric matrix A has a positive definite real part. Such integrals have already occurred in the previous section.

1.2.1 Free fields

The function $G(\mathbf{x}, \lambda)$ can be interpreted as the propagator of a Euclidean free field $\varphi(\mathbf{x})$, defined on each lattice site and varying from $-\infty$ to $+\infty$. In this context, the free field notion is equivalent to a statistical Gaussian weight

$$e^{-S(\varphi)} \quad (64a)$$

where the quadratic action $S(\varphi)$ reads

$$S(\varphi) = \frac{1}{2} \sum_{\mathbf{x}, \mathbf{x}'} \varphi(\mathbf{x}) [\lambda^{-1} \delta_{\mathbf{x}, \mathbf{x}'} - J_{\mathbf{x}, \mathbf{x}'}] \varphi(\mathbf{x}') \quad (64b)$$

with

$$J_{\mathbf{x}, \mathbf{x}'} = \begin{cases} 1/2d & \text{if } |\mathbf{x} - \mathbf{x}'| = 1 \\ 0 & \text{otherwise} \end{cases}$$

One introduces the partition function

$$Z = \int \prod_{\mathbf{x}} \frac{d\varphi_{\mathbf{x}}}{\sqrt{2\pi}} e^{-S(\varphi)} \quad (65)$$

which is well defined for a finite – possibly very large – number of integration variables. This is achieved e.g. by taking a finite system in a box of volume $\Omega = L^d$. The size L will be

sent to infinity at the end of the calculations. This operation (*thermodynamic limit*) will always be implicitly assumed. The correlation function between two fields $\varphi(\mathbf{x})$ and $\varphi(0)$ takes the same form as the propagator $G(\mathbf{x}, \lambda)$ for Brownian motion,

$$G(\mathbf{x}, \lambda) = \langle \varphi(\mathbf{x}) \varphi(0) \rangle = \int_{-\pi}^{\pi} \frac{d^d \mathbf{k}}{(2\pi)^d} \frac{e^{i\mathbf{k} \cdot \mathbf{x}}}{\lambda^{-1} - \frac{1}{d} \sum_{\mu} \cos k_{\mu}} \quad (66)$$

$$= Z^{-1} \int \prod_{\mathbf{x}} \frac{d\varphi_{\mathbf{x}}}{\sqrt{2\pi}} e^{-S(\varphi)} \varphi(\mathbf{x}) \varphi(0)$$

The action (64b) is diagonalized by a Fourier transformation

$$S(\varphi) = \frac{1}{2} \int \frac{d^d \mathbf{k}}{(2\pi)^d} \varphi(-\mathbf{k}) \left[\lambda^{-1} - \frac{1}{d} \sum_{\mu} \cos k_{\mu} \right] \varphi(\mathbf{k}) \quad (67)$$

with $\varphi(-\mathbf{k}) \equiv \overline{\varphi(\mathbf{k})}$, so that the denominator of the propagator is the kernel of the quadratic form in the action. Although φ can take both positive and negative values, its nonvanishing correlations are always positive, as G is. This positivity property is associated with the *ferromagnetic* character of the corresponding statistical model, which describes schematically the statistics of ferromagnetic materials. To be more precise, the action is positive definite for $0 < \lambda < 1$, while it is nonnegative for $\lambda = 1$ due to the zero mode $\varphi = Cst$.

The preceding section has shown the relation between Brownian motion and free field Green functions. We also introduced an additional variable t (the length of the path or proper time) which has no interpretation at equilibrium, but which may describe the approach to equilibrium.

To obtain the continuous limit, we rescale the lattice field φ_L to get the continuous field φ_c (a is the lattice spacing).

$$\varphi_L(\mathbf{x}/a) = \sqrt{2d} a^{d/2-1} \varphi_c(\mathbf{x}) \quad (68)$$

with again

$$\lambda^{-1} = 1 + \frac{m^2 a^2}{2d} \quad (69)$$

The dimensionless field φ_L is now replaced by the field φ_c , with a *canonical* dimension (in terms of the wavenumber $1/a$)

$$[\varphi_c] = \frac{1}{2}(d-2) \quad (70)$$

We can rewrite the action as

$$\begin{aligned}
 S(\varphi_L) &= \frac{1}{2} \sum_{\substack{\mathbf{x}/a, \mathbf{x}'/a \\ \text{integers}}} \varphi_L(\mathbf{x}/a) \left[\lambda^{-1} \delta_{\mathbf{x}, \mathbf{x}'} - J_{\mathbf{x}/a, \mathbf{x}'/a} \right] \varphi_L(\mathbf{x}'/a) \\
 &= \frac{1}{2} \int \frac{d^d \mathbf{x}}{a^d} 2da^{d-2} \left[\left(1 + \frac{m^2 a^2}{2d} \right) \varphi_c^2(\mathbf{x}) \right. \\
 &\quad \left. - \varphi_c(\mathbf{x}) \frac{1}{d} \sum_{\mu=1}^d \left(\varphi_c(\mathbf{x}) + \frac{a^2}{2} \partial_\mu^2 \varphi_c(\mathbf{x}) + \dots \right) \right] \quad (71) \\
 &= \frac{1}{2} \int d^d \mathbf{x} \varphi_c(\mathbf{x}) [m^2 - \Delta] \varphi_c(\mathbf{x}) + \mathcal{O}(a^2) \\
 &= \frac{1}{2} \int d^d \mathbf{x} \left[m^2 \varphi_c^2(\mathbf{x}) + (\nabla \varphi_c(\mathbf{x}))^2 \right] + \mathcal{O}(a^2)
 \end{aligned}$$

where, in the last integral, one has performed an integration by parts and neglected the corresponding contributions at infinity.

Up to terms of order a^2 , the action has a limit in terms of φ_c , which is its continuous Euclidean form. We omit now the index of φ and use a compact symbol $\mathcal{D}\varphi$ for the integration measure. The scaling transformation leads of course to a corresponding change for the correlation functions as in equation (31),

$$\begin{aligned}
 \langle \varphi(\mathbf{x}) \varphi(0) \rangle &= g(\mathbf{x}, m^2) = \int_{-\infty}^{\infty} \frac{d^d \mathbf{k}}{(2\pi)^d} \frac{e^{i\mathbf{k} \cdot \mathbf{x}}}{k^2 + m^2} \\
 &= Z^{-1} \int \mathcal{D}\varphi e^{-S(\varphi)} \varphi(\mathbf{x}) \varphi(0) \quad (72)
 \end{aligned}$$

More generally, averages of an odd number of fields vanish, while those with an even number are given by Wick's theorem

$$\begin{aligned}
 \langle \varphi(\mathbf{x}_1) \cdots \varphi(\mathbf{x}_{2n}) \rangle &= \frac{1}{2^n n!} \sum_{\text{perm.}} \langle \varphi(\mathbf{x}_{p_1}) \varphi(\mathbf{x}_{p_2}) \rangle \cdots \langle \varphi(\mathbf{x}_{p_{2n-1}}) \varphi(\mathbf{x}_{p_{2n}}) \rangle \\
 &= \sum_{\text{pairings}} \prod \langle \varphi(\mathbf{x}_s) \varphi(\mathbf{x}_t) \rangle \quad (73)
 \end{aligned}$$

The number of pairings on the r.h.s. is $(2n)!/(2^n n!) = (2n-1)!!$. A compact form which summarizes equations (73) is provided by the generating functional

$$\begin{aligned}
 \frac{Z(J)}{Z(0)} &= Z^{-1} \int \mathcal{D}\varphi \exp \left\{ -S(\varphi) + \int d^d \mathbf{x} J(\mathbf{x}) \varphi(\mathbf{x}) \right\} \\
 &= \exp \left\{ \frac{1}{2} \int d^d \mathbf{x}_1 d^d \mathbf{x}_2 J(\mathbf{x}_1) \langle \varphi(\mathbf{x}_1) \varphi(\mathbf{x}_2) \rangle J(\mathbf{x}_2) \right\} \quad (74)
 \end{aligned}$$

where the two kernels $\langle \varphi(\mathbf{x}_1) \varphi(\mathbf{x}_2) \rangle$ and $\delta^2 S / \delta \varphi(\mathbf{x}_1) \delta \varphi(\mathbf{x}_2)$ are inverse from each other. Correlation functions can be generated by taking functional derivatives with respect to J . This leads to an immediate proof of Wick's theorem (73).

Each term in the right-hand side of (73) appears as a sum over independent Brownian paths. Hence the free field acts, from this point of view, as a generating function for studying symmetrized sets of Brownian particles. Green functions of interacting fields are similarly expected to have an interpretation in terms of particles moving under restrictions (Symanzik, 1969), as we shall now show.

1.2.2 Interacting fields and random walks

Let us consider now a class of models with interacting fields, invariant with respect to the orthogonal group $O(n)$. The n -component vector field Φ is defined on each lattice site. The lattice spacing acts as an ultraviolet cutoff and we modify the action by adding a local interaction term $(\Phi^2)^2$, the simplest possible one. The action is

$$S(\Phi) = -\beta \sum_{(\mathbf{x}, \mathbf{x}')} \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}') + \sum_{\mathbf{x}} V(\Phi^2(\mathbf{x})) \quad (75)$$

with

$$V(\Phi^2) = \frac{1}{2} \mu^2 \Phi^2 + \frac{1}{4} g (\Phi^2)^2 \quad (76)$$

Only two of the three constants β , μ^2 , g are really independent as can be seen using a scaling argument. To deal with a bounded action (from below), we restrict ourselves to a positive coupling constant, $g \geq 0$; however, now we need not exclude negative values of μ^2 . The point $g = 0$, which is the starting point for a perturbation theory around the free field, is singular, because the integrals cannot be defined for arbitrarily small g , including $\text{Re } g < 0$. The notation $(\mathbf{x}, \mathbf{x}')$ refers to all pairs of neighbouring sites, and hence $\sum_{(\mathbf{x}, \mathbf{x}')}$ is equivalent to $d \sum_{\mathbf{x}, \mathbf{x}'} J_{\mathbf{x}, \mathbf{x}'}$. For $n = 1$ the model describes a scalar field in self-interaction. In the limit $g + \mu^2 = 0$, $g \rightarrow \infty$, we obtain the Ising model, with Φ constrained to take only the values ± 1 . When $n > 1$, this limit constrains $|\Phi|$ to be a constant and yields the *classical Heisenberg model*, the continuous counterpart of which is the *nonlinear σ -model*, referring to the model introduced in particle physics by Gell-Mann

and Levy in their description of spontaneous breaking of chiral invariance. These models will be described in greater detail in following chapters.

It is often natural in formal developments to combine the local part of the action $\sum_{\mathbf{x}} V(\Phi^2)$ with the integration measure. The partition function Z - with the infinite volume limit implicitly understood - reads

$$Z = \int \prod_{\mathbf{x}} \left\{ e^{-V(\Phi^2)} \prod_{\alpha=1}^n \frac{d\Phi^\alpha(\mathbf{x})}{\sqrt{2\pi}} \right\} e^{\beta \sum_{(\mathbf{x}, \mathbf{x}')} \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}')} \quad (77)$$

Except for the propagation or kinetic term, the integrand is factorized. In this statistical version, this term therefore acts as an interaction between fields on different sites. Similarly, correlation functions read

$$\begin{aligned} \langle \Phi^{\alpha_1}(\mathbf{x}_1) \dots \Phi^{\alpha_{2k}}(\mathbf{x}_{2k}) \rangle &= Z^{-1} \int \prod_{\mathbf{x}} \left\{ e^{-V(\Phi^2)} \prod_{\alpha=1}^n \frac{d\Phi^\alpha(\mathbf{x})}{\sqrt{2\pi}} \right\} \\ &\times e^{\beta \sum_{(\mathbf{x}, \mathbf{x}')} \Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}')} \Phi^{\alpha_1}(\mathbf{x}_1) \dots \Phi^{\alpha_{2k}}(\mathbf{x}_{2k}) \end{aligned} \quad (78)$$

The relation with a sum over paths appears when one expands the kinetic term in powers of β , exactly as one proceeds in the high temperature expansion of statistical mechanics, with β proportional to the inverse temperature. However, as the weight $\exp(-V(\Phi^2))$ leads to unmanageable integrals, it is convenient to introduce its Fourier transform in order to perform explicitly the subsequent integrals over Φ , owing to their Gaussian form. Therefore one writes the Laplace transform

$$e^{-V(\Phi^2)} = \int d\mu(a) e^{-a\Phi^2} \quad (79)$$

which defines the measure $d\mu(a)$. The integral implicitly contains a prescription for the integration path over the variable a , to be distinguished from the lattice spacing, running along the imaginary axis. For brevity, let us use the operator σ

$$\sigma f(\mathbf{x}) \equiv 2d(\Delta_r + 1)f(\mathbf{x}) \equiv \sum_{\mu=1}^d [f(\mathbf{x} + \mathbf{e}_\mu) + f(\mathbf{x} - \mathbf{e}_\mu)] \quad (80)$$

Hence

$$\begin{aligned} Z &= \int \prod_{\mathbf{x}} \left\{ \prod_{\alpha=1}^n \frac{d\Phi^\alpha(\mathbf{x})}{\sqrt{2\pi}} d\mu(a_{\mathbf{x}}) \right\} e^{-\frac{1}{2} \sum_{\mathbf{x}} \Phi(\mathbf{x}) [2a_{\mathbf{x}} - \beta\sigma] \Phi(\mathbf{x})} \\ &= \int \prod_{\mathbf{x}} \frac{d\mu(a_{\mathbf{x}})}{(2a_{\mathbf{x}})^{n/2}} \det^{-n/2} \left[1 - \frac{\beta\sigma}{2a_{\mathbf{x}}} \right] \end{aligned} \quad (81)$$

In the last expression, we have performed a Gaussian integral providing an inverse square root of a determinant raised to the power n , the number of field components. Similar formulae hold for the Green functions. We only display the two-point function, the generalization being straightforward,

$$\begin{aligned} \langle \Phi^{\alpha_1}(\mathbf{x}_1) \Phi^{\alpha_2}(\mathbf{x}_2) \rangle &= Z^{-1} \int \prod_{\mathbf{x}} \left\{ \prod_{\alpha=1}^n \frac{d\Phi^\alpha(\mathbf{x})}{\sqrt{2\pi}} d\mu(a_{\mathbf{x}}) \right\} \\ &\times e^{-\frac{1}{2} \sum_{\mathbf{x}} \Phi(\mathbf{x}) [2a_{\mathbf{x}} - \beta\sigma] \Phi(\mathbf{x})} \Phi^{\alpha_1}(\mathbf{x}_1) \Phi^{\alpha_2}(\mathbf{x}_2) \\ &= Z^{-1} \int \prod_{\mathbf{x}} \left\{ \prod_{\alpha=1}^n \frac{d\Phi^\alpha(\mathbf{x})}{\sqrt{2\pi}} d\mu(a_{\mathbf{x}}) \right\} \\ &\times e^{-\frac{1}{2} \sum_{\mathbf{x}} \Phi(\mathbf{x}) [2a_{\mathbf{x}} - \beta\sigma] \Phi(\mathbf{x})} \delta^{\alpha_1 \alpha_2} \langle \mathbf{x}_1 | \frac{1}{2a_{\mathbf{x}} - \beta\sigma} | \mathbf{x}_2 \rangle \end{aligned} \quad (82)$$

In the integrand of the last expression, we recognize the Green function for free fields, generalized to the case where the term $a_{\mathbf{x}}$ depends on $\mathbf{x}$. As a result, the expressions implying the Fourier transform cannot be used anymore. However, the expansion in terms of paths going from $\mathbf{x}_1$ to $\mathbf{x}_2$ is always valid, generalizing equation (30)

$$\langle \mathbf{x}_1 | \frac{1}{2a_{\mathbf{x}} - \beta\sigma} | \mathbf{x}_2 \rangle = \sum_{\omega=\{\mathbf{x}_1, \mathbf{x}_2\}} \beta^{|\omega|} \prod_{\mathbf{x}} \left(\frac{1}{2a_{\mathbf{x}}} \right)^{n_{\mathbf{x}}(\omega)} \quad (83)$$

The notation $n_{\mathbf{x}}(\omega) \geq 0$ stands for the number of times the path goes through $\mathbf{x}$ ($\mathbf{x}$ may coincide with one of the end points), that is the "time" spent at point $\mathbf{x}$. Of course, $\sum_{\mathbf{x}} n_{\mathbf{x}}(\omega) = |\omega| - 1$ so that most of the terms in the products on the right-hand side are equal to 1 as $|\omega| < \infty$ (only a finite number of terms contribute).

The auxiliary quantities a_x (varying mass terms) must now be eliminated. This is done by using an integral representation

$$(2a)^{-k} = \int d\nu_k(t) e^{-2at} \quad (84)$$

with

$$d\nu_k(t) = \begin{cases} \theta(t) \frac{t^{k-1}}{(k-1)!} dt & k > 0 \\ \delta(t) dt & k = 0 \end{cases} \quad (85)$$

Combining the expansion in (83) with this representation and substituting in (82), one finally obtains

$$\begin{aligned} \langle \Phi^{\alpha_1}(x_1) \Phi^{\alpha_2}(x_2) \rangle &= \delta^{\alpha_1 \alpha_2} \sum_{\omega = \{x_1, x_2\}} \beta^{|\omega|} Z^{-1} \\ &\times \int \prod_x \left\{ \prod_{\alpha=1}^n \frac{d\Phi^\alpha(x)}{\sqrt{2\pi}} d\nu_{n_x(\omega)}(t_x) \right\} \\ &\times e^{\beta \sum_{(x,x')} \Phi(x) \cdot \Phi(x') - \sum_x V(\Phi^2(x) + 2t_x)} \end{aligned} \quad (86)$$

The factor $\delta^{\alpha_1 \alpha_2}$ expresses the $O(n)$ invariance. The coefficient of $\beta^{|\omega|}$ in the sum over paths is a weight $z(\omega)$ which takes into account the number of times, n_x , the path ω visits each point x . The bigger n_x , the smaller the contribution. In the free field case where $V(\Phi^2) = \frac{1}{2}\mu^2\Phi^2$, the integral over Φ can be done; it compensates the factor Z^{-1} and gives the same result as in the preceding section, up to a change of notation

$$\begin{aligned} \langle \Phi^{\alpha_1}(x_1) \Phi^{\alpha_2}(x_2) \rangle &= \delta^{\alpha_1 \alpha_2} \langle x_1 | \frac{1}{\mu^2 - \beta\sigma} | x_2 \rangle \\ &= \delta^{\alpha_1 \alpha_2} \frac{1}{\mu^2} \sum_{\omega = \{x_1, x_2\}} \left(\frac{\beta}{\mu^2} \right)^{|\omega|} \end{aligned} \quad (87)$$

In the general case however, the positive weight

$$\begin{aligned} z(\omega) &= Z^{-1} \int \mathcal{D}\Phi d\nu \exp \left\{ \beta \sum_{(x,x')} \Phi(x) \cdot \Phi(x') - \sum_x V(\Phi^2(x) + 2t_x) \right\} \end{aligned} \quad (88)$$

cannot be recast in the form $\exp(\text{Cst}|\omega|)$.

For the $2k$ -point function, the result appears as a sum over $(2k-1)!!$ terms corresponding to all pairings. This provides a nontrivial generalization of Wick's theorem. If $\mathcal{P}$ denotes the pairings,

$$\begin{aligned} \langle \Phi^{\alpha_1}(x_1) \dots \Phi^{\alpha_{2k}}(x_{2k}) \rangle &= \sum_{\mathcal{P}} \delta^{\alpha_{p_1} \alpha_{p_2}} \dots \delta^{\alpha_{p_{2k-1}} \alpha_{p_{2k}}} \\ &\times \sum_{\substack{\omega_1, \dots, \omega_k \\ \partial\omega_j = \{x_{p_{2j-1}}, x_{p_{2j}}\}}} \beta^{|\omega_1| + \dots + |\omega_k|} z(\omega_1, \dots, \omega_k) \end{aligned} \quad (89)$$

with

$$\begin{aligned} z(\omega_1, \dots, \omega_k) &= Z^{-1} \int \prod_x \prod_{\alpha} \frac{d\Phi^\alpha(x)}{\sqrt{2\pi}} d\nu_{\mathcal{N}_x} \\ &\times \exp \left\{ \beta \sum_{(x,x')} \Phi(x) \cdot \Phi(x') - \sum_x V(\Phi^2(x) + 2t_x) \right\} \end{aligned} \quad (90)$$

The weight $z(\omega_1, \dots, \omega_k)$ uses the total number $\mathcal{N}_x = \sum_{j=1}^k n_x(\omega_j)$ of visits to site x by all paths.

Fröhlich (1982) has obtained the following remarkable inequalities

$$\begin{aligned} (i) \quad \omega_1 \cap \omega_2 = \emptyset &\Rightarrow z(\omega_1, \omega_2) \geq z(\omega_1) z(\omega_2) \\ (ii) \quad \sum_{\omega_2} z(\omega_1, \omega_2) &\leq z(\omega_1) \sum_{\omega_2} z(\omega_2) \end{aligned} \quad (91)$$

The proof is based on the following idea. The average values of products of fields decrease if the parameters in the action are modified in such a way as to decrease the probability measure for large Φ . Let $\mathcal{D}\nu_\omega(t) = \prod_x d\nu_{n_x(\omega)}(t_x)$. As

$$a^{-n_1 - n_2} = \int d\nu_{n_1 + n_2}(t) e^{-at} = \int \int d\nu_{n_1}(t_1) d\nu_{n_2}(t_2) e^{-a(t_1 + t_2)},$$

the integral over t_x can be replaced by a double integral over t_{1x} and t_{2x} in expression (90), written for $k = 2$,

$$\begin{aligned} z(\omega_1, \omega_2) &= Z^{-1} \int \mathcal{D}\Phi \mathcal{D}\nu_{\omega_1}(t_1) \mathcal{D}\nu_{\omega_2}(t_2) \\ &\quad \times e^{\beta \sum_{(x,x')} \Phi(x) \cdot \Phi(x') - \sum_x V(\Phi^2 + 2t_{1x} + 2t_{2x})} \\ &= \int \mathcal{D}\nu_{\omega_1}(t_1) \mathcal{D}\nu_{\omega_2}(t_2) z(t_1 + t_2) \end{aligned}$$

with

$$z(t_1 + t_2) = e^{-\sum_x \mu^2(t_{1x} + t_{2x}) + g(t_{1x} + t_{2x})^2} \left\langle e^{-\sum_x g \Phi^2(x)(t_{1x} + t_{2x})} \right\rangle$$

the variables t_1 and t_2 being associated with the paths ω_1 and ω_2 respectively. Similar relations hold for $z(\omega_1)$ and $z(\omega_2)$, so that the inequality (i) is satisfied if

$$\omega_1 \cap \omega_2 = \emptyset \Rightarrow z(t_1 + t_2) \geq z(t_1)z(t_2)$$

Defining $\tilde{z}(t) = \langle \exp(-g \sum_x \Phi^2(x)t_x) \rangle$, one finds

$$\ln \tilde{z}(t_1 + t_2) = \ln \tilde{z}(t_2) + \int_0^1 d\lambda \tilde{z}^{-1}(\lambda t_1 + t_2) \frac{\partial}{\partial \lambda} \tilde{z}(\lambda t_1 + t_2)$$

If one uses a notation $\langle \cdot \rangle_t$ to indicate an additional term $g \sum_x t_x \Phi^2(x)$ in the action (with of course $\langle \cdot \rangle_0 = \langle \cdot \rangle$), the integrand reads $-g \sum_x t_{1x} < \Phi^2(x) >_{\lambda t_1 + t_2}$. Now, we can use the remark mentioned previously; as t is positive, $\langle \Phi^2 \rangle_t$ is a decreasing function of t . More precisely, one has the following inequality due to Griffiths

$$\begin{aligned} \langle \Phi^2(x) \Phi^2(y) \rangle_t^{\text{conn}} &= \langle \Phi^2(x) \Phi^2(y) \rangle_t - \langle \Phi^2(x) \rangle_t \langle \Phi^2(y) \rangle_t \geq 0 \\ \frac{\partial}{\partial \lambda'} \langle \Phi^2(x) \rangle_{\lambda t_1 + \lambda' t_2} &= -g \sum_y t_{2y} \langle \Phi^2(x) \Phi^2(y) \rangle_{\lambda t_1 + \lambda' t_2}^{\text{conn}} \leq 0 \end{aligned} \quad (92)$$

The integral is therefore minimized by replacing $\langle \Phi^2 \rangle_{\lambda t_1 + t_2}$ by $\langle \Phi^2 \rangle_{\lambda t_1}$. Hence

$$\ln \tilde{z}(t_1 + t_2) \geq \ln \tilde{z}(t_2) - \int_0^1 d\lambda g \sum_x t_{1x} \langle \Phi^2(x) \rangle_{\lambda t_1} = \ln \tilde{z}(t_1) + \ln \tilde{z}(t_2)$$

and for the quantities $z(t)$,

$$\exp \left(g \sum_x t_{1x} t_{2x} \right) z(t_1 + t_2) \geq z(t_1) z(t_2)$$

Whenever ω_1 and ω_2 do not intersect, $t_{1x} t_{2x} = 0$ for all x , and inequality (91i) is proved. For inequality (ii), taking $\partial \omega_2 =$

$\{x_3, x_4\}$ fixed, one has

$$\begin{aligned} \sum_{\omega_2} z(\omega_1, \omega_2) &= \sum_{\omega_2} \int \mathcal{D}\nu_{\omega_1}(t_1) \mathcal{D}\nu_{\omega_2}(t_2) z(t_1 + t_2) \\ &= \int \mathcal{D}\nu_{\omega_1}(t_1) z(t_1) \langle \Phi^\alpha(x_3) \Phi^\alpha(x_4) \rangle_{t_1} \end{aligned}$$

with the same notation as above. One now uses the inequality

$$\langle \Phi^\alpha(x_3) \Phi^\alpha(x_4) \rangle_t \leq \langle \Phi^\alpha(x_3) \Phi^\alpha(x_4) \rangle = \sum_{\omega_2} z(\omega_2) \quad (93)$$

which completes the proof of the inequalities (91). These formulae are useful for deriving numerous inequalities for $2k$ -point correlation functions.

1.2.3 Self-avoiding walks and the limit $n \rightarrow 0$

We now look at the n -dependence of correlation functions, focusing on the two-point function. When the internal indices coincide, the latter can be written, using equations (82,83), as

$$\begin{aligned} \langle \Phi^1(x_1) \Phi^1(x_2) \rangle &= Z^{-1} \sum_{\omega = \{x_1, x_2\}} \beta^{|\omega|} \int \prod_x (d\mu(a_x) (2a_x)^{-\frac{1}{2}n - n_x(\omega)}) \\ &\quad \times \det^{-n/2} \left(1 - \frac{\beta\sigma}{2a_x} \right) \\ Z &= \int \prod_x (d\mu(a_x) (2a_x)^{-n/2}) \det^{-n/2} \left(1 - \frac{\beta\sigma}{2a_x} \right) \end{aligned} \quad (94)$$

The dependence in n is explicitly written everywhere, except in the integration measure $d\mu(a_x)$. For simple choices of $V(\Phi^2)$, an analytic continuation in the variable n will be possible. We look at the limit $n \rightarrow +0$. A simplification arises, since

$$\lim_{n \rightarrow 0} \det^{-n/2} \left(1 - \frac{\beta\sigma}{2a_x} \right) = 1 \quad (95)$$

so that the weight $z(\omega)$ becomes a sum over factorized contributions

$$\begin{aligned} z(\omega) &= \sum_x \rho_{n_x(\omega)} \\ \rho_k &= \lim_{n \rightarrow 0} \frac{\int d\mu(a) (2a)^{-(n/2)-k}}{\int d\mu(a) (2a)^{-n/2}} \end{aligned} \quad (96)$$

and

$$\lim_{n \rightarrow 0} \langle \Phi^1(x_1) \Phi^1(x_2) \rangle = \sum_{\omega = \{x_1, x_2\}} \beta^{|\omega|} \sum_x \rho_{n_x(\omega)} \quad (97)$$

The normalization is such that $\rho_0 = 1$. With the action (75)-(76), and using $d\mu(a)$ as defined in (79), we can evaluate ρ_k . If $p > 0$,

$$\begin{aligned} \int d\mu(a) (2a)^{-p} &= \int d\mu(a) \int_0^\infty dt \frac{t^{p-1}}{\Gamma(p)} e^{-2at} \\ &= \int_0^\infty dt \frac{t^{p-1}}{\Gamma(p)} e^{-(\mu^2 t + g t^2)} \end{aligned}$$

while

$$\int d\mu(a) = \int d\mu(a) \int_{-\infty}^\infty dt \delta(t) e^{-2at} = 1$$

Inserting these relations in (96), we get

$$\begin{aligned} \rho_0 &= 1 \\ \rho_k &= \int_0^\infty dt \frac{t^{k-1}}{(k-1)!} e^{-(\mu^2 t + g t^2)} \quad (k \geq 1) \end{aligned} \quad (98)$$

We have already observed that, as the number of visits to a site increases, the corresponding weight decreases (i.e. ρ_k decreases as k increases). Let us consider more precisely the limit $n \rightarrow 0$ for the nonlinear σ -model, which constrains Φ^2 to a constant value conveniently fixed to n

$$\Phi^2 = n \quad (99)$$

This is achieved by choosing

$$\mu^2 = -ng \quad (100)$$

and taking the limit $g \rightarrow \infty$ (before $n \rightarrow 0$). Indeed, for a regular test function

$$\lim_{g \rightarrow \infty} \frac{\int \Pi_\alpha \frac{d\Phi^\alpha}{\sqrt{2\pi}} e^{g[n\Phi^2/2 - (\Phi^2)^2/4]} f(\Phi)}{\int \Pi_\alpha \frac{d\Phi^\alpha}{\sqrt{2\pi}} e^{g[n\Phi^2/2 - (\Phi^2)^2/4]}} = \frac{\int \Pi_\alpha d\Phi^\alpha \delta(\Phi^2 - n) f(\Phi)}{\int \Pi_\alpha d\Phi^\alpha \delta(\Phi^2 - n)} \quad (101)$$

Thus we have

$$\rho_k = \lim_{n \rightarrow 0} \frac{\int d\nu_{k+n/2}(t) \delta(t - n/2)}{\int d\nu_{n/2}(t) \delta(t - n/2)} = \lim_{n \rightarrow 0} (n/2)^k \frac{\Gamma(n/2)}{\Gamma(n/2 + k)}$$

and finally

$$\rho_k = \begin{cases} 1 & \text{if } k = 0, 1 \\ 0 & \text{otherwise} \end{cases} \quad (102)$$

The result is quite remarkable (de Gennes, 1972), as it shows that the $n \rightarrow 0$ limit of the nonlinear σ -model leads for the correlation function to a sum over *non-self-intersecting* paths (excluded volume limit), a model which describes dilute solutions of polymers. It is worthwhile to distinguish this sum over self-avoiding paths from a random walk where the walker would choose with equal probabilities the neighbouring sites not previously visited (and which might lead him into traps). The latter model, of interest in itself, belongs to a quite different *universality* class. This term will be defined more precisely as we proceed. In the present context, it means that the two processes have a different large-distance behaviour.

Extend the preceding result to the $2k$ -point function

$$\begin{aligned} \lim_{n \rightarrow 0} \langle \Phi^1(x_1) \cdots \Phi^1(x_{2k}) \rangle &= \sum_{\text{pairings } P} \sum_{\substack{\omega_1, \dots, \omega_k \\ \partial \omega_j = \{x_{P_{2j-1}}, x_{P_{2j}}\}}} \beta^{|\omega_1| + \dots + |\omega_k|} \\ &\quad \times \prod_x \rho_{n_x(\omega_1) + \dots + n_x(\omega_k)} \end{aligned} \quad (103)$$

with ρ_n given by equations (98) or (102). Again, the paths do not intersect.

Looking for a continuous limit of these models, the intersection properties of Brownian curves suggest that, in dimension greater than four, the interactions of φ^4 type will have no effect on a critical limit which will be equivalent to a free field theory, i.e. a Gaussian model. Conversely, for $d = 4$ (as a limiting case) or for $d < 4$, the continuous limit will be nontrivial.

1.2.4 Comparison with the high temperature expansion

The above expansion in terms of paths has suggestive properties and is useful for theoretical purposes, but it does not lead to easy calculations, except in some particular cases. However, it looks very similar to another expansion (*high temperature*, or

strong coupling series) which provides a practical tool for obtaining explicit series and which will be considered in greater detail in chapter 7. It is instructive to perform a comparison in the case of the Ising model ($n = 1$ "σ-model") where the field φ is restricted to the values ± 1 . The integration measure is thus

$$\frac{d\varphi}{\sqrt{2\pi}} e^{-V(\varphi^2)} = 2\delta(\varphi^2 - 1)d\varphi = d\varphi \int_{-\infty}^{\infty} \frac{da}{\pi} e^{ia - ia\varphi^2} \quad (104)$$

Consider the partition function

$$Z = \sum_{\varphi_x = \pm 1} \exp \left(\beta \sum_{(x, x')} \varphi_x \varphi_{x'} \right) = \sum_{\varphi_x = \pm 1} \prod_{(x, x')} \exp(\beta \varphi_x \varphi_{x'}) \quad (105)$$

Each pair of neighbours (associated to a link of the lattice) yields a term $\exp \beta \varphi \varphi'$, which may be expanded as a power series in β

$$e^{\beta \varphi \varphi'} = \sum_{j=0}^{\infty} \frac{\beta^j}{j!} \varphi^j \varphi'^j \quad (106)$$

Substituting in (105) and using

$$\sum_{\varphi = \pm 1} \varphi^n = 2\delta_{n,0}^{(\text{mod } 2)} = \begin{cases} 2 & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases} \quad (107)$$

one obtains a sum of terms corresponding to configurations of non-negative integers $\{q_l\}$ assigned to each link l and satisfying the condition

$$\partial \{q_l\} = \emptyset \quad (108)$$

where $\partial \{q_l\}$ is the set of sites x such that the sum of the q_l 's on incident links is odd. Hence, if Ω is the total volume (the total number of lattice sites),

$$Z = 2^\Omega \sum_{\partial \{q_l\} = \emptyset} \frac{\beta^{\sum q_l}}{\prod_l q_l!} \quad (109)$$

Similarly, one obtains the 2-point function

$$Z \langle \varphi_x \varphi_y \rangle = 2^\Omega \sum_{\partial \{q_l\} = \{x, y\}} \frac{\beta^{\sum q_l}}{\prod_l q_l!} \quad (110)$$

and analogous expressions for higher order functions. Reordering in increasing powers of β , one gets the desired strong coupling

expansion, which looks similar to the sum over paths, as discussed previously. For a precise comparison, we rewrite the latter as (with $d\mu(a) = e^{ia} da / \pi$)

$$Z = \int \prod_x d\varphi_x d\mu(a_x) \exp \left(\beta \sum_{(x, x')} \varphi_x \varphi_{x'} - i \sum_x a_x \varphi_x^2 \right) \\ = \int \prod_x \left(d\mu(a_x) \sqrt{\frac{\pi}{ia_x}} \exp \left(-\frac{1}{2} \text{Tr} \ln \left(1 - \frac{\beta \sigma}{2ia_x} \right) \right) \right) \quad (111)$$

$$-\frac{1}{2} \text{Tr} \ln \left(1 - \frac{\beta \sigma}{2ia_x} \right) = \frac{1}{2} \sum_x \sum_{\omega_x} \frac{\beta^{|\omega_x|}}{|\omega_x|} \prod_y \left(\frac{1}{2ia_y} \right)^{n_y(\omega_x)} \quad (112)$$

The last expression involves a sum over x and over the paths ω_x starting from and returning to x , and visiting $n_y(\omega_x)$ times the site y . Therefore,

$$Z = \int \prod_x \left(da_x \sqrt{\frac{2}{\pi}} e^{ia_x} \right) \sum_{k, r_q} \frac{1}{2^{\sum_{q=1}^k r_q} \prod_{q=1}^k r_q!} \\ \times \sum_{\substack{x_1, \dots, x_k \\ \omega_{x_1}, \dots, \omega_{x_k}}} \frac{\beta^{\sum_{q=1}^k |\omega_{x_q}|}}{\prod_{q=1}^k |\omega_{x_q}|} \prod_y \left(\frac{1}{2ia_y} \right)^{\frac{1}{2} + \sum_{q=1}^k n_y(\omega_{x_q})} \quad (113)$$

where the path ω_x describes the same geometric image r times. The integration over a can be performed to give

$$\sqrt{\frac{2}{\pi}} \int da e^{ia} \left(\frac{1}{2ia} \right)^{k+\frac{1}{2}} = \frac{2}{(2k-1)!!} \quad (114)$$

(using the convention $(-1)!! = 1$). Indeed,

$$2 = \int d\varphi 2\delta(\varphi^2 - 1) \varphi^{2k} = \int \frac{da}{\pi} e^{ia} \int d\varphi e^{-ia\varphi^2} \varphi^{2k} \\ = (2k-1)!! \sqrt{\frac{2}{\pi}} \int da \frac{e^{ia}}{(2ia)^{k+\frac{1}{2}}}$$

Finally

$$Z = 2^\Omega \sum_{k, r_q} \frac{1}{2^{\sum_{q=1}^k r_q} \prod_{q=1}^k r_q!} \sum_{\substack{x_1, \dots, x_k \\ \omega_{x_1}, \dots, \omega_{x_k}}} \frac{\beta^{\sum_{q=1}^k |\omega_{x_q}|}}{\prod_{q=1}^k |\omega_{x_q}|}$$

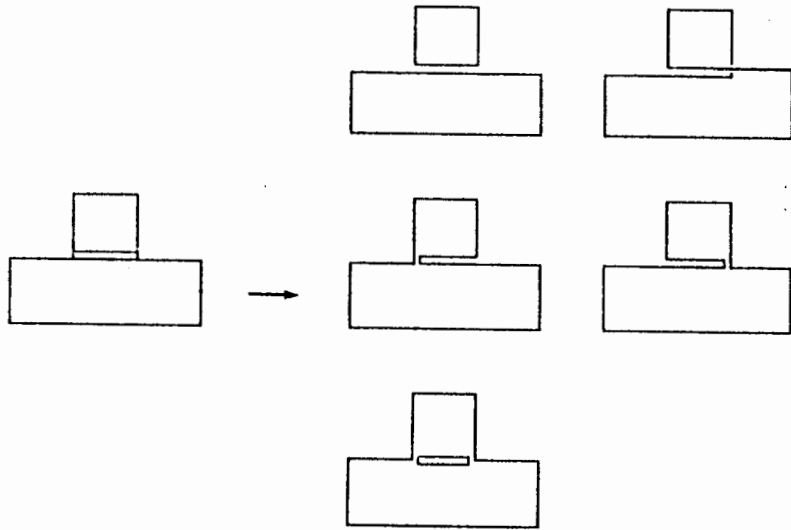


Fig. 5 Example of the relation between high temperature diagrams and closed paths, according to the discussion given in the text.

$$\times \prod_y \frac{1}{(2 \sum n_y(\omega_{x_q}) - 1)!!} \quad (115)$$

This expression can be simplified, observing that the denominator takes partially into account the number of choices of an origin x_q on a closed path, as well as the two possible orientations of these paths (except for some exceptional paths, e.g. those of length 2). A set of closed paths (without a choice of an origin nor an orientation) defines a configuration in the sense of equation (109), which corresponds to the contribution of several terms in the expansion (115).

Let us analyze an example, taking the $\{q_l\}$ depicted on figure 5. The high temperature expansion assigns a contribution $\beta^{12}/2$ to this configuration. In the path decomposition, two sites are visited twice, yielding a factor $[\prod_y (2 \sum n_y(\omega_{x_q}) - 1)!!]^{-1} = 1/3^2$. The power of β is of course 12. The last term in the path decomposition yields an additional factor $\frac{1}{2}$, due to the path of length two which has only one possible orientation (with a fixed origin). Adding all contributions, we recover the expected value $\frac{1}{2} = (4 + \frac{1}{2}) \times (\frac{1}{3^2})$.

The reader may convince himself in the general case that the two methods give identical results.

For the Ising model – as well as for other similar models – the high temperature expansion as presented above is not the most economical one. From the fact that $\varphi\varphi' = \pm 1$, we have the relation

$$e^{\beta\varphi\varphi'} = \cosh \beta + \varphi\varphi' \sinh \beta = \cosh \beta (1 + t\varphi\varphi') \quad (116)$$

with $t = \tanh \beta$

Inserting this identity in the expression (105) for the partition function, and summing over φ_x , one is led to restricted configurations for which $q_l = 0$ or 1. If $L = \Omega d$ stands for the total number of links, we get

$$Z = 2^\Omega (\cosh \beta)^L \sum_{\substack{\theta\{q_l\} = \theta \\ q_l = 0,1}} t^{\sum q_l} \quad (117)$$

This defines again a set of closed curves with, however, an ambiguous separation between connected parts. For correlation functions, we have a similar expansion

$$Z \langle \varphi_{x_1} \cdots \varphi_{x_{2k}} \rangle = 2^\Omega (\cosh \beta)^L \sum_{\substack{\theta\{q_l\} = \{x_1, \dots, x_{2k}\} \\ q_l = 0,1}} t^{\sum q_l} \quad (118)$$

We shall study these formulae in great depth later.

1.2.5 The one-dimensional case

One-dimensional models lead to simple functional integrals. They can be interpreted as solving quantum mechanical problems in terms of stochastic fields, after analytic continuation to imaginary time. As an example, we consider an $O(n)$ symmetric model, using periodic boundary conditions $\Phi_{L+x} = \Phi_x$, with a partition function

$$Z = \int \prod_{x=1}^L \left(\frac{d^n \Phi_x}{(2\pi)^{n/2}} e^{-V(\Phi_x^2)} \right) \exp \beta \sum_x \Phi_x \cdot \Phi_{x+1} \quad (119)$$

Let us introduce the notion of transfer matrix which will be generalized to higher dimensions later. It is an operator $\mathcal{T}$ acting on the states $\psi(\Phi)$ (or wave functions) where Φ denotes a point in n -dimensional space. Its integral kernel satisfies

$$\mathcal{T}\psi(\Phi') = \int \frac{d^n \Phi}{(2\pi)^{n/2}} e^{-V(\Phi^2)} e^{\beta \Phi \cdot \Phi'} \psi(\Phi) \quad (120)$$

so that the partition function may be written as the trace of the L th iterate of $\mathcal{T}$, assuming periodic boundary conditions

$$Z_L = \text{Tr } \mathcal{T}^L = t_0^L + t_1^L + \dots \quad (121)$$

The eigenvalues $\{t_i\}$ of $\mathcal{T}$ have been ordered according to decreasing magnitude. In general, the highest eigenvalue t_0 is not degenerate. This corresponds in quantum mechanics to the uniqueness of the ground state. In the present context, this property can be deduced from a theorem due to Frobenius, which states that a symmetric real matrix with positive coefficients has a nondegenerate highest eigenvalue, and that all coordinates of the corresponding eigenvector have the same sign. The operator $\mathcal{T}$ defined by equation (120) is not symmetric, but we can substitute another equivalent symmetric operator, by replacing $V(\Phi^2)$ by $\frac{1}{2}(V(\Phi^2) + V(\Phi'^2))$ in such a way that equation (121) is still fulfilled. Strictly speaking, Frobenius' theorem only applies to finite matrices, while $\mathcal{T}$ operates here in an infinite dimensional space.

In the thermodynamic limit $L \rightarrow \infty$, Z_L is dominated by the highest eigenvalue, so that the free energy is

$$\mathcal{F} = \lim_{L \rightarrow \infty} \frac{1}{L} \ln Z_L = \ln t_0 \quad (122)$$

In the case of the nonlinear σ -model, the integration measure is

$$\frac{d^n \Phi}{(2\pi)^{n/2}} e^{-V(\Phi^2)} = d^n \Phi \, 2\delta(\Phi^2 - n) \quad (123)$$

The wave functions are defined on a sphere of radius $\sqrt{n}$. Rescaling Φ so that it takes its values on the unit sphere, with a measure $d^{n-1}\hat{\Phi}$ normalized to

$$\int d^{n-1}\hat{\Phi} = S_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

equation (120) takes the form

$$\mathcal{T}\psi(\hat{\Phi}) = n^{(n/2)-1} \int d^{n-1}\hat{\Phi} e^{\beta n \hat{\Phi} \cdot \hat{\Phi}'} \psi(\hat{\Phi}) \quad (124)$$

As $\mathcal{T}$ commutes with rotations in Φ space, its eigenvectors are obtained by an expansion in spherical harmonics. The highest eigenvalue corresponding to $\psi = \text{Cst}$ is thus

$$t_0 = n^{n/2-1} S_{n-1} \int_0^\pi d\theta \sin^{n-2} \theta e^{\beta n \cos \theta} = (2\pi)^{n/2} \frac{I_{n/2-1}(n\beta)}{\beta^{n/2-1}}$$

$$= 2 \cosh \beta \quad \text{for } n = 1 \quad (125)$$

Here $I_\nu(z)$ is the modified Bessel function

$$I_\nu(z) = (z/2)^\nu \sum_{k=0}^{\infty} \frac{(z/2)^{2k}}{k! \Gamma(\nu + k + 1)} \quad (\nu \neq -1, -2, \dots) \quad (126)$$

$$I_{\nu-1} = \frac{2\nu}{z} I_\nu + I_{\nu+1}, \quad I'_\nu = \frac{\nu}{z} I_\nu + I_{\nu+1}$$

The eigenvalue t_0 is therefore an entire even function of β , positive whenever n is an integer. Its expression allows an analytic continuation in n . In particular, $t_0(\beta, n)$ tends to 1 as $n \rightarrow 0$, as expected from section 2.3. For large n , we find

$$t_0(\beta, n) \xrightarrow{n \rightarrow \infty} \left(\frac{2\pi}{\beta} \right)^{\frac{1}{2}(n-1)} \frac{e^{\beta n}}{\sqrt{n}} \quad (127)$$

Let us turn now to correlation functions. Due to rotational invariance,

$$G(1, 2) = \langle \Phi^{\alpha_1}(x_1) \Phi^{\alpha_2}(x_2) \rangle = \delta^{\alpha_1 \alpha_2} \frac{1}{n} \langle \Phi(x_1) \cdot \Phi(x_2) \rangle \quad (128)$$

The eigenvectors (*spherical harmonics*) of the symmetrized $\mathcal{T}$ are orthogonal, since this operator is Hermitian. Denoting them by $|l, m\rangle$ (m being a degeneracy index), one obtains

$$G(1, 2) = \frac{\delta^{\alpha_1 \alpha_2}}{n} \sum_{l, m} \left(\frac{t_l}{t_0} \right)^{|x_1 - x_2|} \langle 0 | \Phi | l, m \rangle \langle l, m | \Phi | 0 \rangle \quad (129)$$

where the ground state $|0\rangle$ corresponds to a constant $\psi_0(\Phi) = S_n^{-\frac{1}{2}}$. This expression is typical of transfer matrix methods. For a large distance $|x_1 - x_2|$, the correlation function is dominated by the leading term in $(t_l/t_0)^{|x_1 - x_2|}$, where t_1 is the second largest eigenvalue, provided the associated matrix elements do not vanish. One observes an exponential behaviour of the correlations, as long as the first two eigenvalues are not degenerate, which is the case here. Applying one component of Φ to the ground state leads to a first order harmonic (let us recall that harmonics of order p are of the form $s_{\alpha_1 \dots \alpha_p} \Phi^{\alpha_1} \dots \Phi^{\alpha_p}$, where $s_{\alpha_1 \dots \alpha_p}$ is a symmetric tensor with vanishing partial traces). Hence, in the present case, the sum (129) reduces to the contributions from the terms with $l = 1$, $m = 1, 2, \dots, n$, and one gets an exponentially decreasing

Identifying the coefficient of the expansion of δA in the basis $A_{(\mu\nu)\alpha\beta} = \delta_{\alpha\mu}\delta_{\beta\nu} - \delta_{\alpha\nu}\delta_{\beta\mu}$, one finds

$$0 = \delta_{\beta\mu} \langle \Phi'^\alpha \Phi^\nu \rangle - \delta_{\beta\nu} \langle \Phi'^\alpha \Phi^\mu \rangle + \beta \langle \Phi'^\alpha \Phi^\beta (\Phi'^\mu \Phi^\nu - \Phi'^\nu \Phi^\mu) \rangle$$

Summing on the indices $\mu = \alpha$, $\nu = \beta$, one obtains the following relation, valid for any n

$$\left(1 - \frac{1}{n}\right) \frac{1}{n} \langle \Phi \cdot \Phi' \rangle = \beta \left(1 - \frac{1}{n^2} \langle (\Phi \cdot \Phi')^2 \rangle\right) \quad (140)$$

For $n = 1$, the relation is trivial, since $(\Phi \cdot \Phi')^2 = 1$. As n increases, one can use the steepest descent method, and the fluctuations of the invariant quantity $\Phi \cdot \Phi'/n$ become negligible. In this limit, mean values of invariant observables factorize. In particular,

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \langle (\Phi \cdot \Phi')^2 \rangle = \left(\frac{1}{n} \langle \Phi \cdot \Phi' \rangle\right)^2 \quad (141)$$

and hence

$$\lim_{n \rightarrow \infty} \left\{ \frac{1}{n} \langle \Phi \cdot \Phi' \rangle + \beta \left[\left(\frac{1}{n} \langle \Phi \cdot \Phi' \rangle \right)^2 - 1 \right] \right\} = 0 \quad (142)$$

which is nothing but the algebraic equation leading to the result (138).

Absence of ordered phase in one dimension.

The very simple expressions obtained for the Ising model can be used to illustrate Landau's proof of the absence of phase transitions in one-dimensional models with short-range interactions. The partition function, normalized to 1 as $\beta \rightarrow \infty$, can be written for a system of size L in the form

$$\bar{Z}_L(\beta) = e^{-L\beta} Z_L(\beta) = (1 + e^{-2\beta})^L = \sum_{k=0}^L \binom{L}{k} e^{-2k\beta}$$

The last expression is interpreted as a partition of the system by k walls separating pure phases (i.e. with all spins up or down). Each wall leads to a variation -2β of the action, and there are $\binom{L}{k}$ ways to choose the positions of these walls. The partition function is a sum over all possible values of k . In the general case of the coexistence of two phases, the term -2β would be replaced by the interface free energy, and the entropy factor $\binom{L}{k}$ would be the same. Now, let us find the average value of k . If this average is

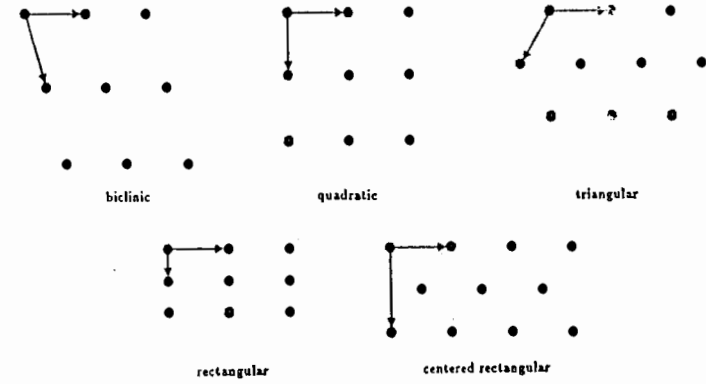


Fig. 6 Two-dimensional lattices.

small, the existence of pure phases over large intervals is favored. However, using Stirling's formula (which is a good approximation to the factorial, even for small values of the argument), we have

$$\bar{F}(\beta) = \frac{1}{L} \ln \bar{Z}_L(\beta) = -(x \ln x + (1-x) \ln(1-x) + 2\beta x) \quad x = \frac{k}{L}$$

Hence, $\partial \bar{F} / \partial x = 0$ yields $k = L / (1 + \exp(2\beta))$. For all finite β , k is proportional to L and disorder will be present in the thermodynamic limit. As $\beta \rightarrow \infty$ (zero temperature limit), k/L vanishes and this explains why the correlation length (L/k) increases indefinitely. One verifies that the substitution $x = k/L = (1 + \exp(2\beta))^{-1}$ gives $\bar{F} = \ln(1 + e^{-2\beta})$ and that $\xi \sim \frac{1}{2} e^{2\beta}$ (as $\beta \rightarrow \infty$) agrees with $L/k \sim e^{2\beta}$.

This reasoning, valid for the coexistence of a finite number of phases, should be modified for a continuous symmetry, since there is a continuum of possible phases. The behaviour of the correlation length is only linear and no longer exponential, the coefficient of β varying from 4 to 2 as n grows from 2 to infinity.

Appendix 1.A Lattices

In this appendix, we remind the reader of some characteristics of two- and three-dimensional regular lattices. Such a lattice is

correlation function

$$G(1, 2) = \delta^{\alpha_1 \alpha_2} \left(\frac{t_1}{t_0} \right)^{|x_1 - x_2|} = \delta^{\alpha_1 \alpha_2} \exp \left(-\frac{|x_1 - x_2|}{\xi} \right) \quad (130)$$

whatever x_1 and x_2 , and ξ has the meaning of a *correlation length*, governing the size of the region over which the fields are correlated. In particular, $t_1/t_0 = n^{-1} \langle \Phi_x \cdot \Phi_{x+1} \rangle = (nL)^{-1} \sum_{x=1}^L \langle \Phi_x \cdot \Phi_{x+1} \rangle$ for two neighbouring sites. This last quantity is nothing but

$$\frac{1}{nL} \frac{\partial}{\partial \beta} \ln Z_L = \frac{1}{n} \frac{\partial}{\partial \beta} \ln t_0$$

Thus,

$$e^{-1/\xi} = \frac{t_1}{t_0} = \frac{1}{n} \frac{\partial}{\partial \beta} \ln t_0 \quad (131)$$

which yields

$$e^{-1/\xi} = \frac{I_{n/2}(n\beta)}{I_{n/2-1}(n\beta)} = \beta - \frac{n}{n+2} \beta^3 + \frac{2n^2}{(n+2)(n+4)} \beta^5 + \mathcal{O}(\beta^7) \quad (132)$$

Note the two particular cases of the Ising model ($n = 1$)

$$\xi^{-1} = -\ln \tanh \beta \quad (133)$$

and of the rigid rotator ($n = 2$)

$$\xi^{-1} = -\ln \frac{I_1(2\beta)}{I_0(2\beta)} \quad (134)$$

In the Ising case, formula (133) can be established directly in a straightforward way. For any integer $n \geq 1$, the correlation length ξ is finite for all β and diverges at zero temperature ($\beta \rightarrow \infty$). For example, for $n = 2$, $\exp(-1/\xi) \sim 1 - 1/4\beta + \mathcal{O}(1/\beta^2)$ and therefore $\xi \sim 4\beta$ as $\beta \rightarrow \infty$. In some sense, we can say that $\beta = \infty$ is the only *critical point*, namely the value of the temperature near which one can substitute a continuous model to the discrete one for the description of the large-distance behaviour. A similar property holds for all one-dimensional models with short-range interactions (Landau).

Two limiting cases are interesting. First, as n vanishes, t_0 tends to unity and one verifies easily that

$$e^{-1/\xi} = \frac{I_{n/2}(n\beta)}{I_{n/2-1}(n\beta)} = \beta \frac{I_{n/2}(n\beta)}{I_{n/2}(n\beta) + \beta I_{n/2+1}(n\beta)} \xrightarrow{n \rightarrow 0} \beta \quad (135)$$

Hence

$$\lim_{n \rightarrow 0} \langle \Phi^1(x_1) \Phi^1(x_2) \rangle = \beta^{|x_1 - x_2|} \quad (136)$$

as expected since there is only one path without self-intersection going from x_1 to x_2 . Positivity properties are lost whenever one performs analytic continuations; this is reflected here by a correlation function which can exceed unity.

Another limiting case is the one corresponding to an infinite number of components, which will be treated in greater detail later. One can use the asymptotic behaviour of Bessel functions

$$I_\nu(\nu z) \underset{\nu \rightarrow \infty}{\sim} \frac{1}{\sqrt{2\pi\nu}} \exp \left\{ \nu \left[\sqrt{1+z^2} + \ln \frac{z}{1+\sqrt{1+z^2}} \right] \right\} \quad (137)$$

or remark that $t_1/t_0 = n^{-1} \langle \Phi_x \cdot \Phi_{x+1} \rangle$ is the value of $\cos \theta$ at the saddle point of the integrand $\exp n(\beta \cos \theta + \ln |\sin \theta|)$ in equation (125), yielding

$$-\beta \sin \theta_{sp} + \cotan \theta_{sp} = 0$$

As a result,

$$\frac{t_1}{t_0} = \cos \theta_{sp} = \beta \left(1 - \frac{t_1^2}{t_0^2} \right) \quad \text{as } n \rightarrow \infty$$

Hence

$$\frac{t_1}{t_0} = e^{-1/\xi} = \frac{\sqrt{1+4\beta^2} - 1}{2\beta} = \beta - \beta^3 + 2\beta^5 + \mathcal{O}(\beta^7) \quad (138)$$

This expansion coincides with the large n limit of equation (132), as it should. The algebraic form of the equation giving t_1/t_0 in the infinite n limit can be derived in another instructive way. From the expression

$$\frac{1}{n} \langle \Phi \cdot \Phi' \rangle = \frac{\int d^n \Phi 2\delta(\Phi^2 - n) e^{\beta \Phi \cdot \Phi'} \Phi \cdot \Phi' / n}{\int d^n \Phi 2\delta(\Phi^2 - n) e^{\beta \Phi \cdot \Phi'}} \quad (139)$$

the $\langle \rangle$ bracket can be interpreted as an average on the vector Φ keeping Φ' fixed of length $\sqrt{n}$. One can also average over the orientations of Φ' . Let us perform an infinitesimal rotation of the integration variables $\Phi \rightarrow (1 + \delta A) \Phi$, where δA is an infinitesimal antisymmetric $n \times n$ real matrix, in the numerator of equation (139). It is clear that the result is unchanged and, owing to the invariance of the measure, one finds

$$\delta \langle \Phi'^\alpha \Phi^\beta \rangle = \delta A_{\beta\gamma} \langle \Phi'^\alpha \Phi^\gamma \rangle + \beta \delta A_{\gamma\delta} \langle \Phi'^\alpha \Phi^\beta \Phi'^\gamma \Phi^\delta \rangle = 0$$

Table II. Fundamental two-dimensional lattices. The lengths of some conventional translation vectors are denoted by a and b , and the angle between them by φ .

lattice	definition
biclinic	$a \neq b, \varphi \neq 90^\circ$
square	$a = b, \varphi = 90^\circ$
triangular	$a = b, \varphi = 120^\circ$
rectangular	$a \neq b, \varphi = 90^\circ$
centered rectangular	$a \neq b, \varphi = 90^\circ$

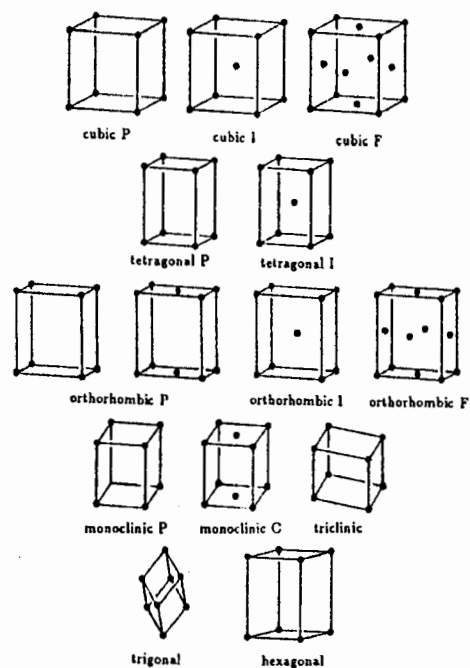


Fig. 7 three-dimensional lattices.

Table III. Fundamental three-dimensional lattices. The lengths of some conventional translation vectors are denoted a, b, c , and their respective angles α, β, γ .

lattice	definition
triclinic	$a \neq b \neq c, \alpha \neq \beta \neq \gamma$
monoclinic $\begin{cases} P \\ C \end{cases}$	$a \neq b \neq c, \alpha = \gamma = 90^\circ \neq \beta$
orthorhombic $\begin{cases} P \\ C \\ I \\ F \end{cases}$	$a \neq b \neq c, \alpha = \beta = \gamma = 90^\circ$
tetragonal $\begin{cases} P \\ I \end{cases}$	$a = b \neq c, \alpha = \beta = \gamma = 90^\circ$
cubic $\begin{cases} P \text{ (simple cubic)} \\ I \text{ (centered cubic)} \\ F \text{ (face centered cubic)} \end{cases}$	$a = b = c, \alpha = \beta = \gamma = 90^\circ$
trigonal	$a = b = c, \alpha = \beta = \gamma \neq 90^\circ$
hexagonal	$a = b \neq c, \alpha = \beta = 90^\circ, \gamma = 120^\circ$

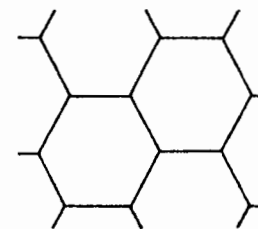


Fig. 8 Honeycomb lattice.

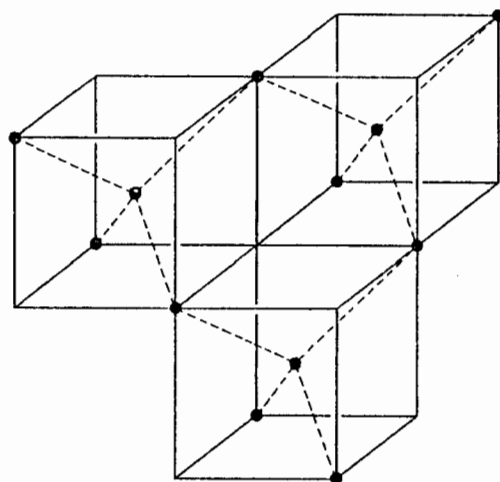


Fig. 9 Diamond lattice.

characterized by a group of translations

$$\mathbf{r} \rightarrow \mathbf{r} + \sum n_i \mathbf{e}_i$$

n_i being integers and $\mathbf{e}_i$ being a set of generators ($i = 1, \dots, d$). Besides translational invariance, there may also exist point symmetries. For instance, lattices may be invariant under cyclic rotation groups of order $n = 2, 3, 4$ or 6 . Other values of n are excluded due to incompatibility with translational symmetry. The existence of a rotational invariance provides relations between the fundamental generators, leading to a classification of lattice types. Figures 6 and 7 display the five two-dimensional and fourteen three-dimensional types, while their characteristics are given in tables II and III.

More complex lattices are obtained by translating a group of sites. Among the various possibilities, let us give the examples of the honeycomb two-dimensional lattice (figure 8) and diamond lattice (figure 9).

We focus on lattices which present a maximal rotational symmetry. This is the case of the triangular and square lattices in two dimensions, and of the simple cubic, body centered cubic and face centered cubic in three dimensions. More generally, in higher

dimensions, one can consider three general types of such lattices. They are the hypercubic lattice, with sites at integral position in an orthonormal frame, the body centered hypercubic lattice composed of two hypercubic lattices displaced from each other by the vector $(\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})$, and the generalized fcc lattice, which is the cross section of a $(d+1)$ -dimensional hypercubic lattice by the hyperplane $x_1 + \dots + x_{d+1} = 0$.

Notes

Our presentation of Brownian motion is very schematic. A more complete treatment can be found in the article by E.W. Montroll and B.J. West in *Studies in Statistical Mechanics VII, Fluctuation Phenomena*, E.W. Montroll and J.L. Lebowitz eds, North Holland, Amsterdam (1979). The complete volume is very instructive reading and provides a wealth of references. Among the numerous mathematical references, we quote the two volume text by W. Feller *An introduction to probability theory and its applications*, second edition, John Wiley and Sons, New York (1971) and *Principles of Random Walk* by F. Spitzer, Springer Verlag, Heidelberg (1976).

Euclidean quantum field theory and its statistical interpretation is presented in a classical course of K. Symanzik in *Rendiconti della Scuola Internazionale di Fisica "Enrico Fermi" XLV Corso*, Teorica quantistica locale, R. Jost ed., Academic Press, New York (1969). The observation that the $O(n)$ -symmetric model describes self-avoiding chains in the $n \rightarrow 0$ limit is due to P.G. de Gennes, *Phys. Lett.* 38A, 339 (1972). For an extensive treatment of polymer physics, see J. des Cloizeaux and G. Jannink, *Les polymères en solution: leur modélisation et leur structure*, Les Editions de Physique (1987). Section 2 on interacting field theory as a weighted sum over paths is patterned after J. Fröhlich, *Nucl. Phys.* B200 [FS4], 281 (1982) who discusses the triviality of $\lambda\phi^4$ theories in dimension larger or equal to four. Similar work is described by M. Aizenman in *Comm. Math. Phys.* 86, 1 (1982). For the ferromagnetic inequalities, we refer to R.B. Griffiths, *J. Math. Phys.* 8, 478 (1967).

GRASSMANNIAN INTEGRALS AND THE TWO-DIMENSIONAL ISING MODEL

Up to now we have considered models arising from classical statistical physics which only involve commuting variables. It is a fascinating aspect of Onsager's solution of the two-dimensional Ising model, that he was able to turn it into a problem involving fermions, indeed into a fermionic free field theory. The natural tools to deal with such systems combine path integral methods with anticommuting variables, described in the first section of this chapter. Such variables originate in the work of 19th century geometers (Grassmann) but their systematic use in a statistical or quantum context is however rather recent (Berezin). We present an overview of Onsager's solution in the second section, leading to the continuum (lattice independent) limit studied in the third section.

2.1 Grassmannian integrals

2.1.1 Anticommuting variables

Consider a finite number of N symbols η_i to be combined linearly using c -number coefficients, which satisfy the following (associative) multiplication rule

$$\eta_i \eta_j + \eta_j \eta_i = 0 \quad (1)$$

i.e. they anticommute, which implies in particular that $\eta_i^2 = 0$. More generally $(\sum f^i \eta_i)^2 = 0$, for c -number coefficients f^i . The Grassmann algebra generated by these symbols contains all expressions of the form

$$f(\eta) = f^0 + \sum_i f^i \eta_i + \sum_{i < j} f^{ij} \eta_i \eta_j + \sum_{i < j < k} f^{ijk} \eta_i \eta_j \eta_k + \dots \quad (2a)$$

$$= \sum_{0 \leq k \leq N} \frac{1}{k!} \sum_{(i)} f^{i_1 \dots i_k} \eta_{i_1} \eta_{i_2} \dots \eta_{i_k} \quad (2b)$$

where we extend the coefficients as antisymmetric tensors with k indices, each ranging from 1 to N . Since there are $\binom{N}{k} = N!/k!(N-k)!$ such linearly independent tensors, summing over k from 0 to N produces a 2^N -dimensional algebra. The word algebra is used here, since equation (1) allows us to define an associative product

$$\begin{aligned} f(\eta)g(\eta) &= f^0 g^0 + \sum_i (f^i g^0 + f^0 g^i) \eta_i \\ &+ \frac{1}{2} \sum_{i,j} (f^{ij} g^0 + f^i g^j - f^j g^i + f^0 g^{ij}) \eta_i \eta_j + \dots \end{aligned} \quad (3)$$

In general note that fg is not equal to $\pm gf$. Nevertheless the subalgebra containing terms with an even number (possibly zero) of η variables commutes with any element f .

We can draw an analogy with the commuting case, where the η_i 's would be replaced by commuting indeterminates X_i , i ranging from 1 to N , and the functions by finite power series (at first) of arbitrary high degree. The coefficients in the latter can be considered as symmetric tensors (of arbitrary high rank). Whereas in the antisymmetric case the number of coefficients of rank k is obtained by expanding $(1+A)^N$ and looking for the term in A^k , i.e. $\binom{N}{k}$, in the symmetric case we have to expand $(1+A)^N$ with a coefficient of A^k given by $\binom{N+k-1}{k} = (N+k-1)!/k!(N-1)!$.

One can present the above expansions in yet another guise which clearly exhibits the relation with Fermi statistics, namely

$$f = \sum_{n_i=0,1} f_{n_1, n_2, \dots, n_N} \eta_1^{n_1} \eta_2^{n_2} \dots \eta_N^{n_N} \quad (2c)$$

The integers $n_i = 0, 1$ can be thought of as occupation numbers of "states" described by η_i . In the commuting case we have a similar expression with X_i replacing η_i and integers n_i unbounded.

A possible choice for a particle-hole-correspondence (in the description given by (2b)) is $f \rightarrow {}^*f$ where

$${}^*f_{\mu_1 \dots \mu_k} = \frac{1}{(N-k)!} \sum_{\mu_{k+1}, \dots, \mu_N} \epsilon_{\mu_1 \dots \mu_k \mu_{k+1} \dots \mu_N} f_{\mu_{k+1} \dots \mu_N} \quad (4a)$$

and ϵ is the totally antisymmetric (Levi Civita) symbol such that $\epsilon_{12 \dots N} = 1$. Show that

(i) in the description (2c) ${}^*f_{n_1 \dots n_N} = \pm f_{n_1+1, \dots, n_N+1}$ where the occupation numbers are understood modulo two. Discuss the corresponding sign.

$$(ii) \quad {}^{**}f_{\mu_1 \dots \mu_k} = (-1)^{k(N-k)} f_{\mu_1 \dots \mu_k} \quad (4b)$$

(iii) the coefficient of $\eta_1 \dots \eta_N$ in ${}^*f(\eta)f(\eta)$ is

$$\sum_k \left(\frac{1}{k!} \sum_{\mu_1 \dots \mu_k} f_{\mu_1 \dots \mu_k}^2 \right) \quad (4c)$$

In the commuting case one introduces a Hilbert space structure, such that two distinct monomials are orthogonal. An orthonormal basis is given by the elements

$$X_{\{n\}} \equiv \frac{X_1^{n_1}}{(n_1!)^{1/2}} \frac{X_2^{n_2}}{(n_2!)^{1/2}} \dots \frac{X_N^{n_N}}{(n_N!)^{1/2}} \rightarrow |\{n\}\rangle \quad (5)$$

Correspondingly the finite series are complemented by infinite ones, defining the Fock-Bargmann space of entire functions. The orthonormality property of the state vectors (5) is expressed using integrals over products of complex planes in the form

$$\delta_{\{n\}, \{m\}} = \langle \{n\} | \{m\} \rangle = \int \prod_{j=1}^N \left(\frac{d^2 X_j}{2\pi} e^{-X_j \bar{X}_j} \right) \bar{X}_{\{n\}} X_{\{m\}} \quad (6)$$

The bar is complex conjugation, and $d^2 X$ stands for $d \operatorname{Re} X d \operatorname{Im} X$. A general vector is represented by an entire series

$$B = \sum_{\{n\}} b_{\{n\}} X_{\{n\}} \quad (7a)$$

with a square norm

$$\langle B | B \rangle = \sum_{\{n\}} |b_{\{n\}}|^2 = \int \prod_{j=1}^N \left(\frac{d^2 X_j}{2\pi} e^{-X_j \bar{X}_j} \right) \overline{B(X)} B(X) \quad (7b)$$

An analogous formula holds for scalar products. This construction is supplemented by the definition of creation and annihilation operators satisfying the Bose commutation rules

$$[a_i, a_j^\dagger] = \delta_{ij} \quad (8a)$$

represented by multiplications and derivations when acting on functions

$$a_i^\dagger \rightarrow X_i \quad a_i \rightarrow \frac{\partial}{\partial X_i} \quad (8b)$$

That these are adjoint operators follows from the exponential factor in the norm.

As Berezin showed, this structure has its exact parallel in the fermionic case. Having already defined sums and products in the Grassmann algebra we now define a left derivative $\partial/\partial \eta_i$. The latter gives zero on a monomial which does not contain the variable η_i . If the monomial contains η_i , the latter is moved to the left, sometimes at the price of a minus sign, then is suppressed. The operation is then extended by linearity to any element of the algebra. One could similarly define a right derivative. It is easily verified that

$$\frac{\partial}{\partial \eta_i} \frac{\partial}{\partial \eta_j} + \frac{\partial}{\partial \eta_j} \frac{\partial}{\partial \eta_i} = 0 \quad (9a)$$

$$\frac{\partial}{\partial \eta_i} \eta_j + \eta_j \frac{\partial}{\partial \eta_i} = \delta_{ij} \quad (9b)$$

The first relation is equivalent to the fact that any linear combination (with c -number coefficients) has a vanishing square. It also shows that derivatives $\partial/\partial \eta_i$ generate in turn a Grassmann algebra, isomorphic with the one expressed in terms of η_i . These quantities may then be viewed as (right-)derivatives on the $\partial/\partial \eta_i$ algebra. Equation (9b) allows one to define fermionic creation and annihilation operators

$$a_i^\dagger \rightarrow \eta_i \quad a_i \rightarrow \frac{\partial}{\partial \eta_i} \quad \{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \quad \{a_i, a_j^\dagger\} = \delta_{ij} \quad (10)$$

The bracket $\{A, B\}$ is the symmetric anticommutator $AB + BA$. The algebra of functions $f(\eta)$ is turned into a Hilbert space

through the definition of a scalar product

$$\langle g, f \rangle = \sum_{k=1}^N \frac{1}{k!} \sum_{i_1, \dots, i_k} \bar{g}^{i_1 \dots i_k} f^{i_1 \dots i_k} \quad (11)$$

where the bar stands again for complex conjugation. If $\bar{\eta}_i$ is a second family of N anticommuting variables (also anticommuting with the η_i 's), one can introduce an antilinear one-to-one map

$$f \rightarrow \bar{f} \quad \sum_k \sum_{\{i\}} f^{i_1 \dots i_k} \eta_{i_1} \dots \eta_{i_k} \rightarrow \sum_k \sum_{\{i\}} \bar{f}^{i_1 \dots i_k} \bar{\eta}_{i_1} \dots \bar{\eta}_{i_k} \quad (12)$$

Note the reverse ordering. This map extends to the 2^{2N} dimensional algebra including products of η 's and $\bar{\eta}$'s.

Returning to the creation and annihilation operators, consider the $2N$ combinations

$$\Gamma_j = a_j + a_j^\dagger \quad \Gamma_{N+j} = i(a_j - a_j^\dagger) \quad 1 \leq j \leq N \quad (13)$$

These can be also viewed as $2^N \times 2^N$ matrices, and they satisfy the relations

$$\{\Gamma_s, \Gamma_{s'}\} = 2\delta_{s,s'} \quad 1 \leq s, s' \leq 2N \quad (14)$$

characteristic of a Clifford algebra. It can be checked that the N pairs $(a_i, a_i^\dagger)$ or the $2N$ combinations Γ_s generate by sums and products all the 2^{2N} linearly independent operators on the original Grassmann algebra. The case $N = 1$ corresponds to quaternions (or Pauli matrices), while if $N = 2$ we have the (four-dimensional) algebra of Dirac matrices.

(i) For $N = 1, 2$ write explicitly the operators $a, a^\dagger, \Gamma$ in the basis of polynomials in η .

(ii) Study the effect of the duality operation, given by equation (4).

(iii) Describe the linear canonical group on creation and annihilation operators in the Bose or Fermi case (symplectic or orthogonal group).

2.1.2 Integrals

The parallel with the Fock-Bargmann construction can be extended by defining anticommuting integrals as linear operations on the functions $f(\eta)$, with the seemingly paradoxical property

that they can be identified with the (left) derivatives! Correspondingly

$$\int d\eta f(\eta) = \frac{\partial f}{\partial \eta}, \quad \int d\eta_1 d\eta_2 f(\eta_1, \eta_2) = \frac{\partial}{\partial \eta_1} \frac{\partial}{\partial \eta_2} f(\eta_1, \eta_2), \quad \dots \quad (15)$$

This yields for the scalar product $\langle g, f \rangle$ defined by equation (11) the following integral representation

$$\langle g, f \rangle = \int d\eta_N d\bar{\eta}_N \dots d\eta_1 d\bar{\eta}_1 \exp\left(-\sum_{i=1}^N \eta_i \bar{\eta}_i\right) \bar{g}(\bar{\eta}) f(\eta) \quad (16)$$

This can be verified by expanding the exponential and using the definition (15)

$$\langle g, f \rangle = \sum \frac{\bar{g}^{j_1 \dots j_q}}{q!} \frac{f^{i_1 \dots i_k}}{k!} \frac{\partial}{\partial \eta_N} \dots \frac{\partial}{\partial \eta_1} (1 + \bar{\eta}_1 \eta_1) \dots (1 + \bar{\eta}_N \eta_N) \bar{\eta}_{j_q} \dots \bar{\eta}_{j_1} \eta_{i_1} \dots \eta_{i_k}$$

Nonvanishing terms arise only when $k = q$ and when the indices j are a permutation of the i 's. Taking into account the antisymmetry of the coefficients, this means that the above expression reduces to

$$\langle g, f \rangle = \sum_k \frac{1}{k!} \sum_{\{i\}} \bar{g}^{i_1 \dots i_k} f^{i_1 \dots i_k} \frac{\partial}{\partial \eta_N} \dots \frac{\partial}{\partial \eta_1} (1 + \bar{\eta}_1 \eta_1) \dots (1 + \bar{\eta}_N \eta_N) \bar{\eta}_{i_k} \dots \bar{\eta}_{i_1} \eta_{i_1} \dots \eta_{i_k}$$

with the indices i all distinct. The last factor can be rewritten $\bar{\eta}_{i_1} \eta_{i_1} \dots \bar{\eta}_{i_k} \eta_{i_k}$. This shows that all other missing pairs $\bar{\eta}\eta$ arise from the other factor $\prod_{i=1}^N (1 + \bar{\eta}_i \eta_i)$ to produce once and only once $\bar{\eta}_1 \eta_1 \dots \bar{\eta}_N \eta_N$, the derivative of which is one. The final result is indeed the combination given in equation (11).

Given this integral expression for the scalar product, we can now proceed to the corresponding representation of operators through integral kernels. The fermionic case (for N finite) is even simpler than its bosonic counterpart, everything being now algebraic as opposed to analytic, which requires us, in principle, to worry about domains of definitions for operators... Only signs have to be watched carefully.

To any linear operator A we associate a kernel $\mathcal{A}$. To comply with a convenient convention we will henceforth denote the members of the Grassmann algebra on which A operates, $f(\xi)$ or $f(\bar{\eta})$... In other words we permute the roles of η and $\bar{\eta}$. Hence

$$(Af)(\bar{\xi}) = \int d\bar{\eta}_N d\eta_N \dots d\bar{\eta}_1 d\eta_1 \exp\left(-\sum_1^N \bar{\eta}_i \eta_i\right) \mathcal{A}(\bar{\xi}, \eta) f(\bar{\eta}) \quad (17)$$

For brevity we shall use the compact notation $d\bar{\eta}d\eta$ for $d\bar{\eta}_N d\eta_N \dots d\bar{\eta}_1 d\eta_1$ and $\bar{\eta}\eta$ for $\bar{\eta}_1 \eta_1 + \dots + \bar{\eta}_N \eta_N$.

A typical example is the integral representation of the identity operator, in analogy with the commutative case

$$f(\bar{\xi}) = \int d\bar{\eta}d\eta e^{-\bar{\eta}\eta} e^{\bar{\xi}\eta} f(\bar{\eta}) \quad (18)$$

Check relation (18) and ascertain the numerical factor (π^{-N}) for the corresponding identity in the Bose case (hint: use the same reasoning as the one leading to equation (16) with $e^{\xi\eta}$ playing the role of $\bar{g}(\eta)$).

Equation (18) enables one to represent any operator by its corresponding kernel $\mathcal{A}(\bar{\xi}, \eta) = Ae^{\bar{\xi}\eta}$, where the right-hand side is interpreted as the action of the operator A on the kernel of the identity considered as a function of $\bar{\xi}$. Thus we have for instance (remember the $\eta \leftrightarrow \bar{\eta}$ interchange)

$$a_i \rightarrow e^{\bar{\xi}\eta} \eta_i \quad a_i^\dagger \rightarrow \bar{\xi}_i e^{\bar{\xi}\eta} \quad (19)$$

and more generally, omitting indices,

$$\begin{aligned} aA &\rightarrow \frac{\partial}{\partial \bar{\xi}} \mathcal{A}(\bar{\xi}, \eta) & Aa &\rightarrow \mathcal{A}(\bar{\xi}, \eta) \eta \\ a^\dagger A &\rightarrow \bar{\xi} \mathcal{A}(\bar{\xi}, \eta) & Aa^\dagger &\rightarrow \mathcal{A}(\bar{\xi}, \eta) \frac{\partial}{\partial \eta} \end{aligned} \quad (20)$$

In the last relation $\overleftarrow{\partial/\partial\eta}$ means that $\partial/\partial\eta$ operates as a right derivative. Let $:$ be the symbol for Wick ordering an expression in a and $a^\dagger$'s. It requires writing the annihilation operators a to the right of the creation ones $a^\dagger$, with due respect for the sign of the corresponding permutation. For instance $:a^\dagger a := a^\dagger a$, $:aa^\dagger := -a^\dagger a$, etc... Observe that we do not use the

commutation rules for otherwise we would be led to the absurd result that $I := I := aa^\dagger + a^\dagger a := -a^\dagger a + a^\dagger a = 0!$ Then, for any monomial

$$:a_{\alpha_1}^\dagger \dots a_{\alpha_s}^\dagger a_{\beta_1} \dots a_{\beta_t}: \rightarrow \bar{\xi}_{\alpha_1} \dots \bar{\xi}_{\alpha_s} e^{\bar{\xi}\eta} \eta_{\beta_1} \dots \eta_{\beta_t} \quad (21)$$

Check formulae (20) and (21).

Therefore to obtain the kernel corresponding to an operator in normal order form, we simply substitute $\bar{\xi}$ for $a^\dagger$, η for a and supply a (commuting) factor $e^{\bar{\xi}\eta}$.

Using the scalar product we endow the Grassmann algebra with a Hilbert space structure. Corresponding to Hermitian conjugation for operators, we have the following operation on kernels

$$A \rightarrow A^\dagger \quad \mathcal{A}(\bar{\eta}, \eta) \rightarrow \overline{\mathcal{A}(\bar{\eta}, \eta)} \quad (22)$$

where the bar over $\mathcal{A}$ implies complex conjugation of coefficients and substitutes for a monomial $\bar{\eta}_{\alpha_1} \dots \bar{\eta}_{\alpha_s} \eta_{\beta_1} \dots \eta_{\beta_t}$ its conjugate $\bar{\eta}_{\beta_t} \dots \bar{\eta}_{\beta_1} \eta_{\alpha_s} \dots \eta_{\alpha_1}$ implying the exchange $\bar{\eta} \leftrightarrow \eta$ together with a reverse ordering. The unit operator being Hermitian, the corresponding kernel is self-conjugate according to (22). The kernels corresponding to a pair $a_i, a_i^\dagger$ are also conjugate.

The kernel for a product is obtained by convolution as

$$\begin{aligned} A &\rightarrow \mathcal{A} & B &\rightarrow \mathcal{B} & AB &\rightarrow \mathcal{C} \\ \mathcal{C}(\bar{\xi}, \xi) &= \int d\bar{\eta}d\eta \mathcal{A}(\bar{\xi}, \eta) e^{-\bar{\eta}\eta} \mathcal{B}(\bar{\eta}, \xi) \end{aligned} \quad (23)$$

Moreover the trace can also be written as

$$\text{Tr } A = \int d\bar{\eta}d\eta e^{-\bar{\eta}\eta} \mathcal{A}(-\bar{\eta}, \eta) \quad (24)$$

Notice the characteristic minus sign in the argument of $\mathcal{A}$. Changing $\eta, \bar{\eta}$ into $-\eta, -\bar{\eta}$, which does not affect the integral, one could instead use $\mathcal{A}(\bar{\eta}, -\eta)$. From equations (18) and (24) it follows that

$$\text{Tr } I = \int d\bar{\eta}d\eta e^{-2\bar{\eta}\eta} = 2^N \quad (25)$$

consistent with the dimension 2^N of the underlying vector space, the Grassmann algebra.

Let us look at some specific properties of anticommuting integrals. Changes of coordinate have to preserve the anticommuting structure of the Grassmann algebra. This allows non-singular linear transformations (with c -number coefficients) of the form

$$\bar{\eta}_i = \sum_{j=1}^N J_i^j \bar{\xi}_j, \quad \det J = \det J_i^j = \frac{D(\bar{\eta})}{D(\bar{\xi})} \quad (26a)$$

One verifies that by setting

$$f(\bar{\eta}) = F(\bar{\xi}) \quad (26b)$$

one obtains the following relation

$$\int d\bar{\xi} F(\bar{\xi}) = \det J \int d\bar{\eta} f(\bar{\eta}) \quad (26c)$$

in contradistinction to the commuting case where the coefficient on the right-hand side would be $|\det J|^{-1}$ involving both an absolute value sign and an inversion. On reflection it is realized that (26c) results in the present case from the identification of derivation and integration. As another consequence we have

$$\int d\bar{\eta} \frac{\partial}{\partial \bar{\eta}} f(\bar{\eta}) = 0 \quad (27)$$

where the commutative analog is a boundary contribution. An important application of these formulae is the one concerning Gaussian integrals, the only nontrivial ones that can be done in practice. First we have to look for the anticommutative analog of quadratic forms. These are obviously antisymmetric forms. Omitting bars on the variables, such a form reads $\sum_{ij} \eta_i Q^{ij} \eta_j$ with $Q^{ij} + Q^{ji} = 0$. The expansion of $\exp \frac{1}{2} \sum_{ij} \eta_i Q^{ij} \eta_j$ only involves even monomials. Consequently an integral over all variables vanishes unless the total number of variables is even: $N = 2n$. In this case

$$\begin{aligned} & \int d\eta_{2n} \dots d\eta_1 \exp \frac{1}{2} \sum_{ij} \eta_i Q^{ij} \eta_j \\ &= \int d\eta_{2n} \dots d\eta_1 \prod_{i < j} (1 + Q^{ij} \eta_i \eta_j) \\ &= \frac{1}{2^n n!} \sum_{\sigma} (-1)^{\sigma} Q^{\sigma_1 \sigma_2} \dots Q^{\sigma_{2n-1} \sigma_{2n}} \equiv \text{Pf}(Q) \end{aligned} \quad (28)$$

Here σ runs over permutations of the $2n$ indices, $(-1)^{\sigma}$ is the parity of the permutation. The prefactor $2^{-n}(n!)^{-1}$ is the number of those permutations which leave any one of the terms invariant. Omitting this factor leads to a sum of distinct pairings (Wick's theorem), each one with the corresponding $\pm$ sign. Finally the whole expression is called a Pfaffian, hence the symbol $\text{Pf}(Q)$. According to (26), a linear change of variables yields

$$\text{Pf}(^T J Q J) = \det(J) \text{Pf}(Q) \quad (29)$$

As a particular case of (28) and (29) suppose that we split the $2n$ variables η_i , $1 \leq i \leq N = 2n$ into two sets $\xi_a, \bar{\xi}_a$ of n variables in such a way that the matrix Q only connects ξ and $\bar{\xi}$ variables. In other words $\{\eta_1 \dots \eta_{2n}\} \equiv \{\xi_1 \dots \xi_n \bar{\xi}_1 \dots \bar{\xi}_n\}$ and $\frac{1}{2} \eta Q \eta = -\bar{\xi} K \xi$, meaning that the matrix Q has the form

$$Q = \begin{pmatrix} 0 & -K \\ K^T & 0 \end{pmatrix} \quad (30)$$

where K is any $n \times n$ matrix. Then

$$(-)^{n(n+1)/2} \text{Pf}(Q) = \int d\bar{\xi}_n d\xi_n \dots d\bar{\xi}_1 d\xi_1 e^{-\bar{\xi} K \xi} = \det K \quad (31)$$

To see this it is sufficient to make the change of variables $\xi' = K\xi$, $\bar{\xi}' = \bar{\xi}$ and to note that $\int d\bar{\eta} d\eta e^{-\bar{\eta} \eta} = 1$. Equation (31) is a widely used formula and has the following consequence. Squaring both sides of equation (28) we get

$$\text{Pf}(Q)^2 = \int d\eta_{2n} \dots d\eta_1 d\xi_{2n} \dots d\xi_1 \exp \frac{1}{2} (\eta Q \eta + \xi Q \xi)$$

Set $\xi_j = (\bar{\rho}_j - \rho_j)/i\sqrt{2}$, $\eta_j = (\bar{\rho}_j + \rho_j)/\sqrt{2}$, for $1 \leq j \leq 2n$. The Jacobian is unity. Applying (31) we get for any antisymmetric even-dimensional matrix Q

$$\text{Pf}(Q)^2 = \int d\bar{\rho}_{2n} \dots d\rho_1 \exp -\bar{\rho} Q \rho = \det Q \quad (32)$$

Of course, if Q is odd-dimensional, $\det Q = 0$.

Mean values with a Gaussian weight satisfy Wick's theorem, again as in the commutative case. This follows from yet another invariance property of the measure $d\eta$ under a translation $\eta \rightarrow \eta + \xi$ with ξ another set of anticommuting variables

$$\int d\eta f(\eta) = \int d\eta f(\eta + \xi) \quad (33)$$

Assuming $\text{Pf}(Q) \neq 0$ (hence $\det Q \neq 0$ and the number of variables is even) define the mean value of any expression $O(\eta)$ as

$$\langle O \rangle = \frac{\int d\eta O(\eta) \exp \frac{1}{2} \eta Q \eta}{\int d\eta \exp \frac{1}{2} \eta Q \eta} \quad (34)$$

Then, using translational invariance, we can derive the analog of the generating functional (1.74)

$$\langle \exp \xi \eta \rangle = \exp -\frac{1}{2} \xi Q^{-1} \xi \quad (35)$$

which is the compact form of Wick's theorem. If we denote by $(Q^{-1})_{ij}$ the ij component of Q^{-1} , it follows from differentiating (35) with respect to the ξ_i 's that

$$\begin{aligned} \langle \eta_1 \eta_2 \rangle &= (Q^{-1})_{12} \\ \langle \eta_1 \eta_2 \eta_3 \eta_4 \rangle &= (Q^{-1})_{12} (Q^{-1})_{34} - (Q^{-1})_{13} (Q^{-1})_{24} + (Q^{-1})_{14} (Q^{-1})_{23} \\ \langle \eta_1 \dots \eta_{2p} \rangle &= \text{Pf}(Q^{-1})_{[1 \dots 2p]} \end{aligned} \quad (36)$$

where the last notation implies a restriction of Q^{-1} to the subspace $[1 \dots 2p]$, for $2p \leq 2n$. This leads to a sum over products of all distinct contractions affected by the sign of the corresponding permutation of indices. A particular case corresponds to a matrix Q of the form indicated in equation (30).

2.2 The two-dimensional Ising model

Let us apply Grassmannian integrals to the solution of the two-dimensional Ising model (Onsager 1944). The model is intended to simulate the Curie transition from a ferro- to a paramagnetic phase. We can also interpret it as a model of the liquid gas transition (the lattice-gas model) where sites are either occupied or empty, or to binary mixtures or alloys. We shall use the magnetic interpretation. The starting point is the partition function in the absence of a magnetic field on a large but finite (regular) lattice with V sites. Pairs of neighbouring sites are denoted (ij) . At each site is attached a "spin variable" σ_i taking the values ± 1 . Boltzmann's constant as well as other energetic units are lumped together in the factor β interpreted as inverse "temperature" in

dimensionless units. The partition function is written

$$Z_V(\beta) = \frac{1}{2^V} \sum_{\sigma_i = \pm 1} \exp \beta \sum_{(ij)} \sigma_i \sigma_j \quad (37)$$

It may seem at first extraordinary that such a simple looking expression is so difficult to compute, as experience has taught us! By going to the infinite volume (or thermodynamic) limit we define the free energy (without a $-\beta$ factor), as

$$F(\beta) = \lim_{V \rightarrow \infty} \frac{1}{V} \ln Z_V(\beta) \quad (38)$$

In the following the infinite volume limit will be implicit most of the time. In the remainder of this chapter, we restrict ourselves to the two-dimensional case, so volume should properly be replaced by area. Also we use a square lattice, each site having four neighbours. Such a lattice admits (geometrically) a notion of duality with the roles of sites and plaquettes (i.e. elementary squares) being interchanged. This leads to an important duality transformation for the Ising model. A generalization is discussed in chapter 6 on lattice gauge fields.

2.2.1 Duality

We have seen in Chapter 1 that $Z_V(\beta)$ can be expanded in powers of β , or even better in powers of

$$t = \tanh \beta \quad (39)$$

after factoring out $(\cosh \beta)^{2V}$. The exact exponent is the number of links belonging to V which may differ from $2V$ by a term negligible compared to V in the thermodynamic limit (it is precisely $2V$ for periodic boundary conditions). The expansion assumes the form

$$Z_V(\beta) = (\cosh \beta)^{2V} \sum_{\partial\{n_i\}=\emptyset} t^{\sum n_i} \quad (40)$$

Each integer $n_i = 0, 1$ refers to a link between neighbours. The boundary $\partial\{n_i\}$ of such a configuration is the set of sites where the sum of n 's attached to incident links is odd. If $\partial\{n_i\} = \emptyset$, each site belongs to an even number (0, 2, 4) of links in the configuration (those for which $n_i = 1$).

Let us now use the dual (plaquette) lattice with corresponding binary variables $\tau_\kappa = \pm 1$. Each link of the new lattice (pair of neighbouring plaquettes) is in one-to-one correspondence with the link of the original lattice on which it is incident. Implicitly we already think in terms of the infinite volume limit and neglect boundary effects. We set up a one-to-two correspondence between configurations $\{n_l\}$ with empty boundaries $\partial\{n_l\} = \emptyset$ and assignments ± 1 to plaquette variables τ_κ , by requiring that adjacent plaquettes carry the same value for τ 's if they are incident on a link l such that $n_l = 0$, or opposite values of τ 's if $n_l = 1$. Obviously reversing all the τ 's leads to a second configuration with the same property, while $\partial\{n_l\} = \emptyset$ insures the consistency of the assignment as one will check by turning around a point of the original lattice. The overall duplication is irrelevant in the calculation of the free energy (38). According to the above procedure

$$\sum n_l = \sum_{(ab)} \frac{1}{2} (1 - \tau_a \tau_b) \quad (41)$$

On the left, the sum runs over the links of the original lattice, on the right, on the dual set of pairs of adjacent plaquettes. Finally, it is realized that the dual lattice is isomorphic to the original one, again in the infinite volume limit. Let us define variables $\tilde{\beta}$ and $\tilde{t} = \tanh \tilde{\beta}$ according to

$$t = \tanh \beta = e^{-2\beta} = \frac{1 - \tilde{t}}{1 + \tilde{t}} \quad \tilde{t} = \tanh \tilde{\beta} \quad (42)$$

Of course $\tilde{\beta}$ runs from zero to infinity and $\tilde{t}$ from zero to one. In the duality transformation, small and large temperatures are interchanged and $\tilde{t} = t$. Using this notation, we rewrite equation (40) as

$$Z'_V(\beta) = (\cosh \beta)^{2V} (2e^{-2\tilde{\beta}})^V \frac{1}{2^{V+1}} \sum_{\tau_a} \exp \left(\tilde{\beta} \sum_{(ab)} \tau_a \tau_b \right) \quad (43)$$

Up to

$$(2 \cosh^2 \beta e^{-2\tilde{\beta}})^V = (\sinh 2\beta)^V = (\sinh 2\tilde{\beta})^{-V}$$

and inessential nonthermodynamic factors, we have reconstructed the original partition function on the dual lattice evaluated at the

dual value $\tilde{\beta}$. Hence we obtain the duality formula (Kramers-Wannier)

$$F(\beta) = \ln \sinh 2\beta + F(\tilde{\beta}) \quad (44)$$

The relation

$$\sinh 2\beta \sinh 2\tilde{\beta} = 1 \quad (45)$$

shows both the consistency of equation (44) and its involutory nature. Duality relates observables and their dual counterparts in the high and low temperature regimes.

The fixed point of the transformation (42), i.e. the one for which $\beta = \tilde{\beta} = \beta_c$, satisfies

$$t_c = \sqrt{2} - 1 \quad \sinh 2\beta_c = 1 \quad \cosh 2\beta_c = \sqrt{2} \\ \beta_c = \frac{1}{2} \ln(\sqrt{2} + 1) = 0.440686794 \dots \quad (46)$$

The forthcoming solution will confirm that β_c is a critical value that separates a disordered high temperature phase ($\beta < \beta_c$) from an ordered low temperature one ($\beta > \beta_c$). The latter is characterized by a spontaneous magnetization M which can be recovered, even in the absence of an external field, from the large distance behaviour of correlations. For instance

$$\lim_{|i-j| \rightarrow \infty} \langle \sigma_i \sigma_j \rangle = M^2 \quad (47)$$

with $M^2 > 0$ for $\beta > \beta_c$.

(i) For a doubly periodic finite square lattice with V points, L links and P plaquettes, supply in the above derivation of duality the exact nonthermodynamic factors. Hint: use Euler's relation $V - L + P = 0$. Discuss the case of free boundary conditions.

(ii) Observe that (a) a triangular and an hexagonal lattice with twice as many points, are (geometrically) dual to each other (b) a decimation procedure may be set up on the latter which maps an Ising model back on a triangular lattice. Therefore find the critical temperature in both cases

$$\begin{array}{ll} \text{triangular} & e^{2\beta_c} = \sqrt{3} \\ \text{hexagonal} & e^{2\beta_c} = 2 + \sqrt{3} \end{array} \quad (48)$$

Since the number of neighbours of a site is 3, 4 and 6 in the hexagonal, square and triangular lattices respectively, one expects for equal densities that the corresponding critical temperatures are increasing in the same order. Check this fact.

2.2.2 Transfer matrix

To compute the partition function we use, as in the one-dimensional case, a transfer matrix T . Properly speaking it is in fact an operator. The latter acts on spin configurations along a row of length L (the x -direction) repeated M times (in the y - or "time" direction), thus enclosing $V = LM$ sites. To simplify future expressions the partition function is written as

$$Z_{LM}(\beta) = \left(\frac{1}{2} \sinh 2\beta\right)^{LM/2} \text{Tr } T^M \quad (49)$$

and therefore assumes periodic boundary conditions in the y -direction as dictated by the trace occurring on the right-hand side. State vectors are described by values $\sigma_x = \pm 1$ with x ranging from 0 to $L-1$. We consider these values as eigenvalues of two by two Pauli matrices $\sigma_x^{(1)}$ with the usual definitions

$$\begin{aligned} \sigma^{(1)} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \sigma^{(2)} &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} & \sigma^{(3)} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ \sigma^{(1)}\sigma^{(2)} &= i\sigma^{(3)}, & \sigma^{(2)}\sigma^{(3)} &= i\sigma^{(1)}, & \sigma^{(3)}\sigma^{(1)} &= i\sigma^{(2)} \end{aligned} \quad (50)$$

The choice of $\sigma^{(1)}$ rather than $\sigma^{(3)}$, which would seem more natural at this stage, is dictated by future convenience. The transfer matrix T is then factorized into a contribution θ collecting the interactions of spins along a row, and $\bar{\theta}$ involving interactions between spins located on two successive rows

$$T = \theta \bar{\theta} \quad (51)$$

This immediately gives

$$\theta(\beta) = \exp \beta \sum_x \sigma_x^{(1)} \sigma_{x+1}^{(1)} \quad (52)$$

where the sum over x runs from 0 to $L-2$ for open boundary conditions or from 0 to $L-1$ for periodic boundary conditions in which case $\sigma_L^{(k)} = \sigma_0^{(k)}$. As for $\bar{\theta}$ it is factorized into commuting contributions from each site along the row. Let V be a matrix which diagonalizes $\sigma^{(1)}$, i.e.

$$V^{-1} \sigma^{(1)} V = -\sigma^{(3)} \quad V^{-1} \sigma^{(2)} V = \sigma^{(2)} \quad V^{-1} \sigma^{(3)} V = \sigma^{(1)} \quad (53)$$

Remembering the factor extracted in (49) as well as the one included in the normalisation of the partition function, the

contribution of site x to $\bar{\theta}_x$ is,

$$\begin{aligned} V^{-1} \bar{\theta}_x V &= \frac{1}{2(\sinh \beta \cosh \beta)^{1/2}} \begin{pmatrix} e^\beta & e^{-\beta} \\ e^{-\beta} & e^\beta \end{pmatrix} \\ &= \frac{e^\beta}{2(\sinh \beta \cosh \beta)^{1/2}} \left[1 + \tanh \beta \sigma_x^{(1)} \right] \end{aligned}$$

Remarking that $e^\beta (4 \sinh \beta \cosh \beta)^{-1/2}$ is $\cosh \bar{\beta}$, and using (53) we deduce that

$$\bar{\theta}(\bar{\beta}) = e^{\bar{\beta}} \sum_x \sigma_x^{(3)} \quad (54)$$

completing the construction of the transfer matrix T . One observes a parallelism between the roles of β and $\bar{\beta}$. There remains however a difference between the operators $\sigma_x^{(1)} \sigma_{x+1}^{(1)}$ and $\sigma_x^{(3)}$, both Hermitian and with a square equal to one (since $\sigma_x^{(k)}$ and $\sigma_{x+1}^{(k)}$ commute) involved respectively in θ and $\bar{\theta}$. This difference will be eliminated by the fermionic formalism to which we turn.

Show that $T^\dagger = \bar{\theta} \theta$ is different from T but nevertheless the partition function is both real (and positive in a finite volume) for real β .

2.2.3 Fermionic representation

On each site the operators $\frac{1}{2} (\sigma_x^{(1)} \pm i \sigma_x^{(2)})$ form a pair of fermionic creation and annihilation operators, but they commute from site to site. There exists a transformation (Jordan-Wigner) which turns them into bona fide fermionic creators and annihilators, after ordering the sites. In the present case, such an ordering is directly afforded by the ordering of the x -axis. Using notations due to Perk, we introduce in addition to the integer labelling sites of the original one-dimensional lattice, a second set of points on the one-dimensional dual lattice with half integral coordinates. We then assign to each pair α consisting of two neighbours, one from each lattice, an operator Γ_α belonging to a Clifford algebra (hence a fermionic one) in such a way that

$$\{\Gamma_\alpha, \Gamma_\beta\} = 2\delta_{\alpha\beta} \quad (55)$$

To be precise, starting from a fixed origin, call it $x = 0$, we write

$$\begin{aligned} \Gamma_{-\frac{1}{2},0} &= \sigma_0^{(1)} & \Gamma_{x-\frac{1}{2},x} &= \left(\prod_0^{x-1} \sigma_{x'}^{(3)} \right) \sigma_x^{(1)} & x \geq 1 \\ \Gamma_{0,\frac{1}{2}} &= \sigma_0^{(2)} & \Gamma_{x,x+\frac{1}{2}} &= \left(\prod_0^{x-1} \sigma_{x'}^{(3)} \right) \sigma_x^{(2)} & x \geq 1 \end{aligned} \quad (56)$$

As a result Γ_α is Hermitian. The starting point $x = 0$ could be made to recede to $-\infty$ at the price of introducing an ill-defined infinite product. The string factor $\prod_0^{x-1} \sigma_{x'}^{(3)} = \exp \frac{1}{2} i \pi \sum_0^{x-1} (\sigma_{x'}^{(3)} - 1)$ reverses the sign of all states with $\sigma^{(3)} = -1$ up to $x-1$. It is readily checked that this factor produces the desired anticommutation rules (55). Should we require periodic boundary conditions $\sigma_0^{(k)} \equiv \sigma_L^{(k)}$, the previous definitions would still make sense up to $x = L-1$ at the price of introducing a slight complication in the transfer matrix (see below). We have now the symmetric expressions

$$\begin{aligned} \sigma_x^{(3)} &= i^{-1} \Gamma_{x-\frac{1}{2},x} \Gamma_{x,x+\frac{1}{2}} \\ \sigma_x^{(1)} \sigma_{x+1}^{(1)} &= i^{-1} \Gamma_{x,x+\frac{1}{2}} \Gamma_{x+\frac{1}{2},x+1} \end{aligned} \quad (57)$$

The second relation would fail for periodic boundary conditions in the subspace $\prod_0^{L-1} \sigma_x^{(3)} = 1$ (it holds in its orthogonal subspace where $\prod_0^{L-1} \sigma_x^{(3)} = -1$). For the particular bond $\sigma_{L-1}^{(1)} \sigma_L^{(1)} \equiv \sigma_{L-1}^{(1)} \sigma_0^{(1)}$ we therefore have to write with our conventions

$$\begin{aligned} \sigma_{L-1}^{(1)} \sigma_0^{(1)} &= i^{-1} \gamma \Gamma_{L-1,L-\frac{1}{2}} \Gamma_{-\frac{1}{2},0} \\ \gamma &= -i^{-L} \prod_0^{L-1} \Gamma_{x-\frac{1}{2},x} \Gamma_{x,x+\frac{1}{2}} \end{aligned} \quad (58)$$

We leave it to the reader to amend the forthcoming expressions appropriately if it is desired to impose periodic boundary conditions. We shall not do so, having ultimately in mind only the discussion of the thermodynamic limit, except for the finite expressions in equations (81) and (82) below.

The two factors of the transfer matrix are then expressed as

$$\begin{aligned} \theta(\beta) &= \exp -i\beta \sum_x \Gamma_{x,x+\frac{1}{2}} \Gamma_{x+\frac{1}{2},x+1} \\ \tilde{\theta}(\tilde{\beta}) &= \exp -i\tilde{\beta} \sum_x \Gamma_{x-\frac{1}{2},x} \Gamma_{x,x+\frac{1}{2}} \end{aligned} \quad (59)$$

The transition from $\theta(\beta)$ to $\tilde{\theta}(\tilde{\beta})$ implies interchanging both β with $\tilde{\beta}$ and the two one-dimensional dual lattices, exhibiting the full symmetry of the problem.

We are now in a position to obtain in a straightforward way a path integral formulation of the partition function. Indeed both θ and $\tilde{\theta}$ appear as exponentials of quadratic forms in fermionic operators. The latter generate a group, the so called spin-group, namely a double covering of the orthogonal group which is familiar in the context of the spin representations of the ordinary three-dimensional rotation group, or of its relativistic four-dimensional extension.

We can carry out the discussion either in operator or path integral language. We use the second option, but we still need to introduce creation and annihilation operators by pairing the Γ 's. This has the unfortunate drawback of destroying the explicit duality. Set for instance

$$\Gamma_{x-\frac{1}{2},x} = a_x + a_x^\dagger \quad \Gamma_{x,x+\frac{1}{2}} = i(a_x - a_x^\dagger) \quad (60)$$

Consequently

$$\tilde{\theta}(\tilde{\beta}) = \exp \tilde{\beta} \sum_x (a_x^\dagger a_x - a_x a_x^\dagger) = \prod_x (e^{-\tilde{\beta}} + 2 \sinh \tilde{\beta} a_x^\dagger a_x) \quad (61)$$

Accordingly, using the same symbol, the corresponding integral kernel is

$$\begin{aligned} \tilde{\theta}(\tilde{\beta}; \bar{\xi}, \xi) &= \prod_x (e^{-\tilde{\beta}} + 2 \sinh \tilde{\beta} \bar{\xi}_x \xi_x) e^{\bar{\xi}_x \xi_x} = \prod_x (e^{-\tilde{\beta}} + e^{\tilde{\beta}} \bar{\xi}_x \xi_x) \\ &= \exp \left\{ -L\tilde{\beta} + \sum_x t^{-1} \bar{\xi}_x \xi_x \right\} \end{aligned} \quad (62)$$

recalling that $t^{-1} = e^{2\tilde{\beta}}$.

To compute the kernel of $\theta(\beta)$ we use the unitary operator U which shifts the coordinate x by a half integral unit

$$\begin{aligned} U^{-1}\Gamma_{x,x+\frac{1}{2}}U &= \Gamma_{x+\frac{1}{2},x+1} & U^{-1}\Gamma_{x-\frac{1}{2},x}U &= \Gamma_{x,x+\frac{1}{2}} \\ U^{-1}a_x^\dagger U &= b_x^\dagger = \frac{1}{2}i(a_{x+1}^\dagger - a_x^\dagger + a_{x+1} + a_x) \\ U^{-1}a_x U &= b_x = \frac{1}{2}i(-a_{x+1}^\dagger - a_x^\dagger - a_{x+1} + a_x) \end{aligned} \quad (63)$$

This assumes implicitly a limit $L \rightarrow \infty$. Defining Fourier transforms as

$$\xi_x = \int_{-\pi}^{+\pi} \frac{dp}{2\pi} e^{ipx} \xi_p \quad \bar{\xi}_x = \int_{-\pi}^{+\pi} \frac{dp}{2\pi} e^{-ipx} \bar{\xi}_p \quad (64)$$

one verifies that the kernels corresponding to U and its adjoint U^{-1} read

$$\begin{aligned} U(\bar{\xi}, \xi) &= 2^{-L/2} \exp \sum_{x \geq x'} 2i\bar{\xi}_x \xi_{x'} - \bar{\xi}_x \bar{\xi}_{x'} + \xi_x \xi_{x'} \\ &= 2^{-L/2} \exp \int_{-\pi}^{+\pi} \frac{dp}{2\pi} \left[\frac{2i}{1 - e^{-ip}} \bar{\xi}_p \xi_p \right. \\ &\quad \left. + \frac{1 + e^{-ip}}{2(1 - e^{-ip})} (\bar{\xi}_{-p} \bar{\xi}_p + \xi_{-p} \xi_p) \right] \end{aligned} \quad (65a)$$

$$\begin{aligned} U^{-1}(\bar{\xi}, \xi) &= 2^{-L/2} \exp \int_{-\pi}^{+\pi} \frac{dp}{2\pi} \left[\frac{-2i}{1 - e^{ip}} \bar{\xi}_p \xi_p \right. \\ &\quad \left. + \frac{1 + e^{ip}}{2(1 - e^{ip})} (\bar{\xi}_p \bar{\xi}_{-p} + \xi_p \xi_{-p}) \right] \end{aligned} \quad (65b)$$

The kernel for $\theta(\beta) = U^{-1}\bar{\theta}(\beta)U$ follows from applying twice the convolution formula (23) to Gaussian kernels with the result

$$\begin{aligned} \theta(\beta; \bar{\xi}, \xi) &= (\cosh \beta)^L \\ &\times \exp \int_{-\pi}^{+\pi} \frac{dp}{2\pi} \frac{(1 - t^2)\bar{\xi}_p \xi_p + it \sin p (\bar{\xi}_{-p} \bar{\xi}_p + \xi_{-p} \xi_p)}{1 + t^2 + 2t \cos p} \end{aligned} \quad (66)$$

In order to obtain the partition function we now need to convolute these expressions M times and to take the trace according to equation (24). Neglecting again a boundary term, and calling y the temporal variable, we obtain

$$Z = \left(\frac{1}{2} \sinh 2\beta\right)^{LM} \int \prod_{x,y} d\bar{\eta}_{xy} d\eta_{xy} d\bar{\rho}_{xy} d\rho_{xy} \exp S \quad (67a)$$

$$S = \sum_y \int_{-\pi}^{+\pi} \frac{dp}{2\pi} \left[-\bar{\eta}_{p,y} \eta_{p,y+1} - \bar{\rho}_{p,y} \rho_{p,y} + t^{-1} \bar{\rho}_{p,y} \eta_{p,y} + \frac{(1 - t^2)\bar{\eta}_{p,y} \rho_{p,y} + it \sin p (\bar{\eta}_{-p,y} \bar{\eta}_{p,y} + \rho_{-p,y} \rho_{p,y})}{1 + t^2 + 2t \cos p} \right] \quad (67b)$$

According to the definitions (64), the variable p is conjugate to x . This seems a very nonlocal form for the action S due to the denominator. It simplifies, however, considerably if we set

$$\begin{aligned} \bar{\eta}_{py} &= \bar{\xi}_{py}^V & \bar{\rho}_{py} &= -\bar{\xi}_{py}^H - \bar{\xi}_{py}^V + \xi_{-py}^H \\ \eta_{py} &= -t \xi_{py}^V & \rho_{py} &= -\bar{\xi}_{py}^V + (1 + te^{ip}) \xi_{py}^H \end{aligned} \quad (68a)$$

It is even clearer to return to configuration space, where the equations (68a) now read

$$\begin{aligned} \bar{\eta}_{xy} &= \bar{\xi}_{xy}^V & \bar{\rho}_{xy} &= -\bar{\xi}_{xy}^H - \bar{\xi}_{xy}^V + \xi_{-xy}^H \\ \eta_{xy} &= -t \xi_{xy}^V & \rho_{xy} &= -\bar{\xi}_{xy}^V + \xi_{xy}^H + t \xi_{x+1,y}^H \end{aligned} \quad (68b)$$

The Jacobian is t per site, hence we have to divide by t^{LM} . Therefore

$$\begin{aligned} Z &= (\cosh \beta)^{2LM} \int \mathcal{D}(\xi, \bar{\xi}) \exp S \\ S &= \sum_{xy} \left\{ t (\bar{\xi}_{x,y}^V \xi_{x,y+1}^V + \bar{\xi}_{x,y}^H \xi_{x+1,y}^H) \right. \\ &\quad \left. + \bar{\xi}_{xy}^H \xi_{xy}^H + \bar{\xi}_{xy}^V \xi_{xy}^V + \bar{\xi}_{xy}^V \xi_{xy}^H + \bar{\xi}_{xy} \bar{\xi}_{xy}^H + \bar{\xi}_{xy}^H \xi_{xy}^V + \xi_{xy}^V \xi_{xy}^H \right\} \end{aligned} \quad (69)$$

an expression first derived by S. Samuel. The action has a suggestive form. Terms proportional to t correspond to propagation from one site to a neighbouring one, while the six remaining terms establish connections between incident links at a site. We have lost explicit duality due to the particular choice made in equation (60). Finally we have two pairs of fermionic variables per site $(\xi^V, \bar{\xi}^V)$ $(\xi^H, \bar{\xi}^H)$ since we kept the variables involved in both kernels for θ and $\bar{\theta}$ separately. We can now turn to the computation of the free energy per site.

2.2.4 Free energy

From equation (69), the partition function is obtained as a Pfaffian, i.e. the square root of the determinant of an antisymmetric matrix. We can take advantage of the translational invariance

by using two-dimensional Fourier transforms, setting $\mathbf{x} \equiv (x, y)$, $\mathbf{p} \equiv (p_x, p_y)$ and letting ξ stand for the vector $(\xi^V \xi^V \xi^H \xi^H)$

$$\xi_{\mathbf{x}} = \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} e^{i\mathbf{p} \cdot \mathbf{x}} \xi_{\mathbf{p}} \quad (70)$$

$$\xi_{\mathbf{p}} \equiv \begin{pmatrix} \xi_{\mathbf{p}}^V \\ \xi_{\mathbf{p}}^H \end{pmatrix}$$

The action is now block diagonal

$$S = \frac{1}{2} \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} \xi_{-\mathbf{p}}^T K_{\mathbf{p}} \xi_{\mathbf{p}} \quad (71a)$$

$$K_{\mathbf{p}} = \begin{pmatrix} 0 & 1 + te^{ip_y} & 1 & 1 \\ -(1 + te^{-ip_y}) & 0 & -1 & 1 \\ -1 & 1 & 0 & 1 + te^{ip_x} \\ -1 & -1 & -(1 + te^{-ip_x}) & 0 \end{pmatrix} \quad (71b)$$

We have

$$\ln Z = 2LM \ln(\cosh \beta) + \frac{1}{2} \text{Tr} \ln K \quad (72a)$$

$$\text{Tr} \ln K = LM \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} \ln \det K_{\mathbf{p}} \quad (72b)$$

$$\det K_{\mathbf{p}} = (1 + t^2) - 2t(1 - t^2)(\cos p_x + \cos p_y) \\ = (\cosh \beta)^{-4} [\cosh^2 2\beta - \sinh 2\beta(\cos p_x + \cos p_y)] \quad (72c)$$

Thus follows Onsager's celebrated result for the free energy per site

$$F(\beta) = \frac{1}{2} \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} \ln [\cosh^2 2\beta - \sinh 2\beta(\cos p_x + \cos p_y)] \quad (73)$$

which indeed satisfies the duality relation (44). There exist many alternative derivations of this expression, some of them much shorter, and perhaps more illuminating than the previous one. The latter however suits our purposes in providing a motivation for the continuous theory in the critical region.

The argument of the logarithm in equation (73) appears in the denominator of the "propagator" $K_{\mathbf{p}}^{-1}$ for the fermionic variables

ξ . In the Brillouin zone $|p_x|, |p_y| \leq \pi$ the combination $(\cos p_x + \cos p_y)$ reaches its maximum value 2 at the origin. Therefore the effective square mass is $\cosh^2 2\beta - 2 \sinh 2\beta$ and is always positive (for real positive β) except for the unique critical value β_c , where $\cosh 2\beta_c = \sqrt{2}$, $\sinh 2\beta_c = 1$ as expected. For β negative corresponding to the antiferromagnetic regime not discussed here, one would have to go to the boundary of the Brillouin zone where $\cos p_x + \cos p_y = -2$ with β_c replaced by $-\beta_c$. Here we limit ourselves to the ferromagnetic system. The free energy F is always regular except possibly at β_c , where its second derivative has a logarithmic singularity. We adopt for brevity the following notation

$$C = \cosh 2\beta \quad \tilde{C} = \cosh 2\tilde{\beta} = C/S \\ S = \sinh 2\beta \quad \tilde{S} = \sinh 2\tilde{\beta} = 1/S \quad (74)$$

Half the internal energy, i.e. the mean value of a neighbouring pair of spins $\langle \sigma \sigma' \rangle$, is continuous at $\beta = \beta_c$ and is given by

$$\frac{1}{2} \frac{d}{d\beta} F(\beta) = \langle \sigma \sigma' \rangle = \frac{\tilde{C}}{2} \left[1 + \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} \frac{S - \tilde{S}}{C\tilde{C} - \cos p_x - \cos p_y} \right] \quad (75)$$

Its behaviour close to β_c is

$$\langle \sigma \sigma' \rangle = \frac{4\sqrt{2}}{\pi} (\beta - \beta_c) \ln \frac{1}{|\beta - \beta_c|} + \text{regular part} \quad (76)$$

Show by performing one of the integrals in (75) that

$$\frac{1}{2} \frac{dF}{d\beta}(\beta) = \langle \sigma \sigma' \rangle = \frac{\tilde{C}}{2} \left[1 + (S - \tilde{S}) \phi(\beta) \right] \quad (77a)$$

$$\phi(\beta) = \int_0^{2\pi} \frac{d\theta}{2\pi} \left[(\cosh 2(\beta + \tilde{\beta}) - \cos \theta) \right. \\ \left. \times (\cosh 2(\beta - \tilde{\beta}) - \cos \theta) \right]^{-1/2} \\ = \frac{k}{2} \cdot \frac{2}{\pi} \int_0^{\pi/2} d\varphi (1 - k^2 \sin^2 \varphi)^{-1/2} \\ = \frac{k}{2} F\left(\frac{1}{2}, \frac{1}{2}; 1; k^2\right) \quad (77b)$$

with $F(a, b; c; x)$ the hypergeometric function, and $k = 2/C\tilde{C} = 2S/C^2$. One recognizes a complete elliptic integral of the first kind.

The specific heat, namely the second derivative of F , has a logarithmic divergence at $\beta = \beta_c$

$$\begin{aligned} \frac{1}{2} \frac{d^2 F(\beta)}{d\beta^2} &= \sum_{(x,x')} \{ \langle \sigma \sigma' \sigma_x \sigma_{x'} \rangle - \langle \sigma \sigma' \rangle \langle \sigma_x \sigma_{x'} \rangle \} \\ &\simeq \frac{4\sqrt{2}}{\pi} \ln \frac{1}{(\beta - \beta_c)} + \text{reg. part} \end{aligned} \quad (78)$$

showing that at criticality energy-energy correlations have an infinite range. Here we call energy the observable $\sigma \sigma'$ product of spins at neighbouring sites. This is our first encounter with second order phase transitions where correlation lengths become infinite and in an appropriate scale the discrete theory can be replaced by a continuous one, ignoring the fine details of the lattice. As it will turn out, close to criticality, the two-dimensional Ising model is nothing but a free Fermi field theory. What is particular to the statistical mechanics aspect is the emphasis on certain observables which derive from the original formulation and interpretation of the model.

(1) Diagonalize $K_p = -K_p^\dagger = -K_p^T$. The eigenvalues are of the form $\pm i\lambda, \pm i\lambda'$, with $\pm\lambda$ and $\pm\lambda'$ solutions of

$$\begin{aligned} \lambda^4 - 2\lambda^2 [2 + (1+t^2) + t(\cos p_x + \cos p_y)] \\ + (1+t^2)^2 - 2t(1-t^2)(\cos p_x + \cos p_y) = 0 \end{aligned} \quad (79a)$$

Let λ be the eigenvalue which vanishes at criticality for $p = 0$, while λ' remains finite,

$$\begin{aligned} \pm i\lambda &= \pm i \left[2 + (1+t^2) + t(\cos p_x + \cos p_y) \right. \\ &\quad \left. - \sqrt{4 + t(\cos p_x + \cos p_y) + 4t^2 - 8} \right]^{1/2} \\ \pm i\lambda' &= \pm i \left[2 + (1+t^2) + t(\cos p_x + \cos p_y) \right. \\ &\quad \left. + \sqrt{4 + t(\cos p_x + \cos p_y) + 4t^2 - 8} \right]^{1/2} \end{aligned} \quad (79b)$$

If the eigenvectors are such that

$$K_p n_p^{(\pm)} = \pm i\lambda n_p^{(\pm)} \quad K_p n_p'^{(\pm)} = \pm i\lambda' n_p'^{(\pm)}$$

one may choose $n_{-p}^{(-)} = n_p^{(+)*}$ and similarly for n' , and adjust the relative phase of n and n' in such a way that the matrix

$$U_p = (n_p^{(+)}, n_p^{(-)}, n_p'^{(+)}, n_p'^{-})$$

be unitary, unimodular, such that

$$U_p^\dagger = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} U_{-p}^T$$

and

$$K_p = U_p \begin{pmatrix} 0 & i\lambda_p & 0 & 0 \\ -i\lambda_p & 0 & 0 & 0 \\ 0 & 0 & 0 & i\lambda_p' \\ 0 & 0 & -i\lambda_p' & 0 \end{pmatrix} U_{-p}^T$$

If we now change variables in (71) by defining

$$\xi_p' = U_{-p}^T \xi_p$$

we obtain the canonical form of the action

$$\xi_{-p} K_p \xi_p = i \xi_{-p}' \begin{pmatrix} 0 & \lambda & 0 & 0 \\ -\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda' \\ 0 & 0 & -\lambda' & 0 \end{pmatrix} \xi_p' \quad (80)$$

The last two components of ξ' are decoupled from the first two and correspond to massive modes for any value of β positive. They can safely be ignored if one is only interested in the critical region.

(2) Find the partition function on a finite $L \times M$ lattice, with periodic boundary conditions on the spins as

$$Z = (\cosh \beta)^{2LM} \frac{1}{2} \left(\int_{-,-} + \int_{-,+} + \int_{+,-} + \int_{+,+} \right) \mathcal{D}(\xi, \bar{\xi}) \exp S \quad (81)$$

Here S is the action (69), with x and y restricted to the range $(0, L-1)$ and $(0, M-1)$ respectively, while the signs refer to the boundary conditions $\xi_{L,y}^H = \pm \xi_{0,y}^H$ and $\xi_{x,M}^V = \pm \xi_{x,0}^V$. Hence obtain the expressions

$$\begin{aligned} Z &= (\cosh \beta)^{2LM} \frac{1}{2} \sum_{\epsilon_1, \epsilon_2=0, \frac{1}{2}} \eta_{\epsilon_1, \epsilon_2} \\ &\times \prod_{q=0}^{M-1} \prod_{p=0}^{L-1} \left\{ (1+t^2)^2 \right. \\ &\quad \left. - 2t(1-t^2) \left[\cos \frac{2\pi(p+\epsilon_1)}{L} + \cos \frac{2\pi(q+\epsilon_2)}{M} \right] \right\}^{1/2} \end{aligned} \quad (82)$$

where $\eta_{\frac{1}{2},0} = \eta_{0,\frac{1}{2}} = \eta_{\frac{1}{2},\frac{1}{2}} = 1$, $\eta_{0,0} = \epsilon(T - T_c)$. Of course, the $(0,0)$ term vanishes at criticality. Generalize this expression when the couplings in the horizontal and vertical directions are distinct.

2.2.5 Spontaneous magnetization

For $\beta > \beta_c$, when we consider far away spins σ_0 and σ_x , the correlation $G(x) = \langle \sigma_0 \sigma_x \rangle$ does not vanish in the limit where x goes to infinity, but rather tends to a positive constant, the square of the spontaneous magnetization, which we denote by $\mathcal{M}$. Thus it is as if the mean value of a spin were equal to $\mathcal{M}$, although in the absence of any symmetry breaking perturbation, such as one induced by a small magnetic field or a specific boundary condition, one should expect the mean value of a spin to vanish. This means in reality that such a symmetric situation beyond β_c is in fact a superposition of two pure states, each one corresponding to a well-defined spontaneous magnetization $\pm \mathcal{M}$. As a convenience in the calculation, we reach $\mathcal{M}$ indirectly by computing the asymptotic behaviour of $G(x)$, the same in any of these two pure states, hence the same in any mixture,

$$\mathcal{M}^2 = \lim_{x \rightarrow \infty} G(x) \quad (83)$$

We say that symmetry is spontaneously broken whenever $\mathcal{M} \neq 0$. It is not *a priori* obvious that $\mathcal{M}$ should be independent of the direction in which we let x go to infinity. We presuppose this property when, as a further matter of convenience, we let x go to infinity along one of the two principal axes, for instance along the first one. Then $x \equiv (x, 0)$ and we note that $\sigma_0 \sigma_x = (\sigma_0 \sigma_1)(\sigma_1 \sigma_2) \cdots (\sigma_{x-1} \sigma_x)$ using the fact that at any site $\sigma_i^2 = 1$. Now we recall that in the high temperature expansion (1.116) each link of the lattice contributed a factor $(1 + t\sigma\sigma')$ to the Boltzmann weight. When computing the numerator in the average value $\langle \sigma_0 \sigma_x \rangle$ we therefore modify each of the links $(01)(12)\dots(x-1, x)$ through $(1 + t\sigma_i \sigma_{i+1}) \rightarrow \sigma_i \sigma_{i+1} (1 + t\sigma_i \sigma_{i+1}) = t(1 + t^{-1} \sigma_i \sigma_{i+1})$. We therefore conclude that, apart from a factor $t^{|x|}$, the required mean value takes the form of a ratio between a "frustrated" partition function Z' and the original one Z

$$G(x) = t^{|x|} \frac{Z'}{Z} \quad (84)$$

where in Z' all the links joining 0 to x one has substituted t^{-1} for t . Thus following the steps which lead to the integral (69), we have a similar representation for Z' with a new kernel $K' = K + Q$ and

$$G(x) = t^{|x|} [\det(1 + K^{-1}Q)]^{1/2} \quad (85a)$$

with the (antisymmetric) matrix Q given by

$$\frac{1}{2} \xi Q \xi = \frac{1 - t^2}{t} \sum_{x'=0}^{x-1} \bar{\xi}_{x',0}^H \xi_{x'+1,0}^H \quad (85b)$$

The matrix $K^{-1}Q$ has lines and columns indexed by the two coordinates x, y , as well as a fourfold symbol $\bar{\xi}^V, \xi^V, \bar{\xi}^H, \xi^H$. Obviously all matrix elements but those in the columns such that $y = 0, x = 0, 1, \dots, x-1$, and the internal index is either $\bar{\xi}^H$ or ξ^H , vanish. As a result, in the computation of the above determinant we can delete all lines but those with the same characterization. This leaves a $(2x) \times (2x)$ determinant which involves the inverse of K . Restricted to the above subspace, the latter, which we denote by $(K^{-1})_{x_1, x_2}^H$, has the form

$$(K^{-1})_{x_1, x_2}^H = \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} \frac{e^{ip_x(x_1 - x_2)} \begin{bmatrix} a_p & b_p \\ c_p & d_p \end{bmatrix}}{(1 + t^2)^2 - 2t(1 - t^2)(\cos p_x + \cos p_y)} \quad (86a)$$

$$\begin{aligned} a_p &= -d_p = -2it \sin p_y \\ b_p &= -c_p^* = 2(1 + t \cos p_y) - (1 + te^{ip_x})(1 + t^2 + 2t \cos p_y) \end{aligned} \quad (86b)$$

Let $\chi(u)$ denote the characteristic function of the interval 1 to x i.e. $\chi(u) = 1$ if $n = 1, 2, \dots, x$ and zero otherwise. Then

$$\begin{aligned} [K^{-1H}Q]_{x_1, x_2} &= \frac{1 - t^2}{t} \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} \frac{e^{ip_x(x_1 - x_2)}}{(1 + t^2)^2 - 2t(1 - t^2)(\cos p_x + \cos p_y)} \\ &\quad \times \begin{bmatrix} -b_p e^{-ip_x} \chi(x_2 + 1) & a_p e^{ip_x} \chi(x_2) \\ -d_p e^{-ip_x} \chi(x_2 + 1) & c_p e^{ip_x} \chi(x_2) \end{bmatrix} \end{aligned} \quad (87)$$

Since a_p and d_p are odd in p_y their contribution vanishes. Using $c_p = -b_p^*$ one discovers that $\det(1 + K^{-1}Q)$ is the square of an

$x \times x$ determinant of the form

$$G(x)^2 = t^{2|x|} \det(1 + K^{-1}Q) = (\det A)^2 \quad (88a)$$

$$A = \begin{vmatrix} a_0 & a_{-1} & a_{-2} & \cdots & a_{-(x-1)} \\ a_1 & a_0 & a_{-1} & \cdots & a_{-(x-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{x-1} & a_{x-2} & a_{x-3} & \cdots & a_0 \end{vmatrix} \quad (88b)$$

$$a_k = \int_{-\pi}^{+\pi} \frac{d^2 p}{(2\pi)^2} e^{-ikp_x} \frac{2t(1+t^2) - t^2(1-t^2)e^{-ip_x} - (1-t^2)e^{ip_x}}{(1+t^2)^2 - 2t(1-t^2)(\cos p_x + \cos p_y)} \quad (88c)$$

We pause to note that one has therefore in "closed form" the value of the spin-spin correlation function at any separation along the principal axes. In particular if $x = 1$, the matrix A reduces to a_0 and (88) to the previous result (75). However, in the general case $x > 1$ equation (88) is still not very manageable. We proceed to evaluate it in the limit $x \rightarrow \infty$. First the integral over p_y can be performed in (88c) and

$$a_k = \int_{-\pi}^{+\pi} \frac{dp}{2\pi} e^{-ikp} \left[\frac{(t - \bar{t}e^{ip})(t\bar{t} - e^{ip})}{(te^{ip} - \bar{t})(\bar{t}te^{ip} - 1)} \right]^{1/2} \quad (88d)$$

The quantity between brackets in the integral is of unit modulus. To extract the square root, its phase is chosen by continuity assuming it is zero for $p = 0$ when $t > t_c$ (i.e. in the low temperature phase). We observe again the very similar roles played by t and $\bar{t}$. A matrix such as A given in (88b) is referred to as a Toeplitz matrix, and the corresponding determinant as a Toeplitz determinant. To obtain its limiting value when $x \rightarrow \infty$, we use a result due to Szegő, which goes as follows. Let a function $f(z)$ analytic in the vicinity of the unit circle be expanded in a Laurent (or Fourier) series as

$$f(z) = \sum_{q=-\infty}^{+\infty} a_q z^q \quad |z| = 1 \quad (89a)$$

We require furthermore that as one rotates around the unit circle the phase of f returns to its original value, making its logarithm well-defined. Thus the image of the unit circle through f does not

encircle the origin in the (complex) f plane. Expand the logarithm $h(z) = \ln f(z)$ in Laurent series

$$h(z) = \ln f(z) = \sum_{q=-\infty}^{+\infty} h_q z^q \quad |z| = 1 \quad (89b)$$

Then Szegő's lemma states that

$$\lim_{x \rightarrow \infty} e^{-x h_0} \det \begin{vmatrix} a_0 & a_{-1} & \cdots & a_{-(x-1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{x-1} & a_{x-2} & \cdots & a_0 \end{vmatrix} = \exp \sum_1^{\infty} q h_q h_{-q} \quad (89c)$$

We present below an interpretation of this formula which makes it sound plausible. For the time being, we apply it to the function appearing in (88d), namely with $z = e^{ip}$

$$f(z) = \left(\frac{t - \bar{t}z}{tz - \bar{t}} \frac{t\bar{t} - z}{t\bar{t}z - 1} \right)^{1/2} \quad (90)$$

Now $f(z)$ has two branch points inside the unit circle, one located at $z_1 = t\bar{t}$, the second one z_2 is either at $t/\bar{t}$ or its inverse, according to whichever one has a modulus smaller than one. Two other branch points lie outside the unit circle. Let us look at the low temperature regime, corresponding to $t > t_c > \bar{t}$. It is readily seen that under such circumstances the total variation of the argument of f as one describes the unit circle vanishes, and as a result $h = \ln f$ is well-defined. Moreover

$$h_0 = 0, \quad h_q = -h_{-q} = \frac{1}{2q} \left((t\bar{t})^{|q|} - \left(\frac{\bar{t}}{t} \right)^{|q|} \right) \quad q \neq 0 \quad (91)$$

Applying (89c) we conclude that in the low temperature phase

$$\mathcal{M}^2 = \lim_{x \rightarrow \infty} \langle \sigma_0 \sigma_x \rangle = \left[1 - \frac{\bar{t}^2 (1 - t^2)^2}{t^2 (1 - \bar{t}^2)^2} \right]^{1/4} = [1 - (\sinh 2\beta)^{-4}]^{1/4} \quad (92a)$$

Therefore the spontaneous magnetization $\mathcal{M}$ is (Onsager, Yang)

$$\mathcal{M} = [1 - (\sinh 2\beta)^{-4}]^{1/8} \quad (92b)$$

At zero temperature, or $\beta \rightarrow \infty$, $\mathcal{M}$ takes the value 1, then decreases down to zero as β recedes to β_c where $\sinh 2\beta_c = 1$.

More significant is the nontrivial power law behaviour as β comes close to β_c

$$\mathcal{M} \sim (\beta - \beta_c)^{1/8} \quad (92c)$$

What is somehow surprising is the simple appearance of the expression (92b), considering the complexity of the intermediate stages in the calculation. A simpler derivation is still awaiting. Furthermore, the power law behaviour in $(\beta - \beta_c)^{1/8}$ or $(1 - T/T_c)^{1/8}$ (with $T = \beta^{-1}$) is not at all expected. We shall see later that a mean field approximation, ignoring fluctuations, would have led to a $(1 - T/T_c)^{1/2}$ behaviour. In the present case the departure from zero at the critical point is much steeper. Finally it is customary to characterize the behaviour of the spontaneous magnetization close to T_c as $(1 - T/T_c)^\beta$ and hence

$$\beta = \frac{1}{8} \quad (92d)$$

The reader will notice a conflict in notation between the inverse temperature β and the critical exponent for the magnetization. We can only offer the apology that both are in common use and should not lead to confusion. That the index β comes out as a rational fraction is part of a much more involved story which will be described in chapter 9. Below T_c , two pure phases with magnetizations $\pm \mathcal{M}$ may coexist, and one can compute the associated surface tension (it would be more appropriate to speak of perimeter tension in a two-dimensional world). This is discussed below.

Szcgő's lemma and averages over unitary matrices Instead of reproducing the original proof of Szcgő and Kac, let us present an heuristic proof, the interest of which is to rely on averages over the unitary group. Let A denote the Toeplitz matrix built from the Fourier coefficients of the function f of size $x \times x$. Expressing these coefficients as integrals we find

$$\det A = \det [a_{k_1 - k_2}] = \int \prod_{k=0}^{x-1} \frac{d\varphi_k}{2\pi} f(e^{i\varphi_k}) \\ \times \det \begin{vmatrix} 1 & e^{i\varphi_0} & \dots & e^{i(x-1)\varphi_0} \\ e^{-i\varphi_1} & 1 & \dots & e^{i(x-2)\varphi_1} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i(x-1)\varphi_{x-1}} & e^{-i(x-2)\varphi_{x-1}} & \dots & 1 \end{vmatrix}$$

The determinant under the integral sign is recognized as a Vandermonde determinant up to a factor

$$\exp \left(-i \sum_{k=0}^{x-1} k\varphi_k \right) \prod_{0 \leq k_1 < k_2 \leq x-1} (e^{i\varphi_{k_1}} - e^{i\varphi_{k_2}})$$

Since the measure is invariant under any permutation of the k -indices, the latter expression can be replaced by its average over permutations. Under such permutation, the above product picks a sign, the signature of the permutation, and the combination of prefactors builds up the complex conjugate Vandermonde determinant. Hence the equality

$$\det A = \frac{1}{x!} \int \prod_{k=0}^{x-1} \frac{d\varphi_k f(e^{i\varphi_k})}{2\pi} \prod_{0 \leq k_1 < k_2 \leq x-1} |e^{i\varphi_{k_1}} - e^{i\varphi_{k_2}}|^2$$

If f is replaced by unity, the determinant is obviously one, hence the measure is normalized. Now let U stand for any unitary $x \times x$ matrix with eigenvalues $e^{i\varphi_k}$ and set

$$F(U) = \sum a_k U^k$$

Then $\det F(U) = \prod_{k=0}^{x-1} f(e^{i\varphi_k})$ and

$$\frac{1}{x!} \prod_{k=0}^{x-1} \frac{d\varphi_k}{2\pi} \prod_{0 \leq k_1 < k_2 \leq x-1} |e^{i\varphi_{k_1}} - e^{i\varphi_{k_2}}|^2$$

is the unique normalized measure on the conjugacy classes of the unitary group. Denoting by brackets averages over the unitary group, our expression reads

$$\det A = \langle \det F(U) \rangle = \langle \exp \text{Tr } H(U) \rangle$$

where in analogy with $f = \exp h$ we have $F = \exp H$ and the expansion of H in powers of U involves the same coefficients as those of h , as follows readily once we bring U to a diagonal form. We then evaluate the last average using the standard cumulant expansion

$$\det A = \exp \left\{ \langle \text{Tr } H \rangle + \frac{1}{2} \left[\langle (\text{Tr } H)^2 \rangle - \langle \text{Tr } H \rangle^2 \right] \right. \\ \left. + \frac{1}{6} \left[\langle (\text{Tr } H)^3 \rangle - 3 \langle (\text{Tr } H)^2 \rangle \langle \text{Tr } H \rangle + 2 \langle \text{Tr } H \rangle^3 \right] + \dots \right\}$$

The mean values of the powers of $\text{Tr } H$ are readily computed, leading to

$$\det A = \exp x h_0 + \sum_{p=1}^{\infty} \text{Min}(p, x) h_p h_{-p} + \\ + \frac{1}{2} \sum_{\substack{q, r=1 \\ q+r \geq x}}^{\infty} \text{Min}(q, r, q+r-x, x) (h_q h_r h_{-q-r} + h_{-q} h_{-r} h_{q+r}) + \dots$$

If h_k tends to zero fast enough as $x \rightarrow \infty$, the generic term tends to zero and we are left with the desired result

$$\lim_{x \rightarrow \infty} e^{-x h_0} \det A = \exp \sum_{p=1}^{\infty} p h_p h_{-p}$$

2.2.6 Correlation function in the high temperature phase

We follow here the derivation given by McCoy and Wu. We return to the expressions (88) and assume $t < t_c < \bar{t}$, in which case the phase of f varies by 2π as we describe the unit circle. By factoring out z , we write $f(z) = z g(z)$, with $\ln g(z)$ well-defined for $|z| = 1$ and

$$a_k = \int_{-\pi}^{\pi} \frac{d\varphi}{2\pi} e^{-i(k-1)\varphi} g(e^{i\varphi}) = g_{k-1} \\ g^2(z) = \frac{\bar{t} - t z^{-1}}{\bar{t} - t z} \times \frac{1 - t \bar{t} z^{-1}}{1 - t \bar{t} z} \quad (93)$$

If B stands for the $(x+1) \times (x+1)$ matrix

$$B = \begin{vmatrix} g_0 & g_{-1} & \dots & g_{-x} \\ g_1 & g_0 & \dots & g_{-(x-1)} \\ \vdots & \vdots & \ddots & \vdots \\ g_x & g_{x-1} & \dots & g_0 \end{vmatrix} \quad (94)$$

the determinant we wish to evaluate is the minor of the element in the last row and first column, i.e.

$$\det A = \det B (B^{-1})_{0x} \quad (95)$$

From the above, as $x \rightarrow \infty$

$$\det B \xrightarrow{x \rightarrow \infty} \left\{ (1-t^2)^2 (1-t^2 \bar{t}^2) (1-t^2 \bar{t}^{-2}) \right\}^{1/4} \quad (96)$$

To compute the inverse matrix B^{-1} , we have to invert the equation

$$uB = v \quad (97)$$

Given the vector $v \equiv \{v_0, \dots, v_x\}$ we want to find the vector $u \equiv \{u_0, \dots, u_x\}$. If v reduces to $v_k = \delta_{k,0}$, then $u_x = (B^{-1})_{x,0}$. We can rephrase equation (97) by introducing the two polynomials

$$u(z) = \sum_0^x u_k z^k \quad v(z) = \sum_0^x v_k z^k \quad (98)$$

and compare it with the expression $u(z)g(z^{-1}) - v(z)$. The remainder can be written for $|z| = 1$

$$u(z)g(z^{-1}) - v(z) = z^x \varphi(z) + \psi(z^{-1}) \quad (99a)$$

where

$$\varphi(z) = \sum_{q=0}^x \sum_{k=1}^{\infty} u_q g_{q-k-x} z^k \\ \psi(z) = \sum_{q=0}^x \sum_{k=1}^{\infty} u_q g_{q+k} z^k \quad (99b)$$

The two terms on the r.h.s. of (99a) are analytic respectively inside and outside the unit circle, assuming the coefficients to decrease fast enough, while u and v being polynomials are analytic in the entire complex plane. Let us now factor out $g(z^{-1})$ into a function analytic inside the unit circle times a function analytic outside, in the form

$$g(z^{-1}) = \frac{1}{P(z)Q(z^{-1})} \\ P(z) = \frac{1}{\sqrt{\left(1 - \frac{t}{\bar{t}}z\right)(1 - t\bar{t}z)}} \quad Q(z) = \sqrt{\left(1 - \frac{t}{\bar{t}}z\right)(1 - t\bar{t}z)} \quad (100)$$

Of course, $P(z)Q(z) = 1$. As a consequence, $P(z)^{-1}$ is analytic inside the unit circle. It then follows that, for $|z| = 1$,

$$\frac{u(z)}{P(z)} = Q(z^{-1})v(z) + Q(z^{-1})[z^x \varphi(z) + \psi(z^{-1})] \quad (101)$$

In the last term, let us split the series into two parts. One involves only non-negative powers of z , the other negative powers, denoted

respectively by brackets with a subscript + or -. As a result

$$\frac{u(z)}{P(z)} - [Q(z^{-1})v(z) + Q(z^{-1})z^x\varphi(z)]_+ = \quad (102)$$

$$Q(z^{-1})\psi(z^{-1}) + [Q(z^{-1})v(z) + Q(z^{-1})z^x\varphi(z)]_-$$

The left-hand side admits an analytic continuation inside the unit circle, the right-hand side a continuation outside and vanishes at infinity. Consequently, both sides represent the same analytic function which, according to Liouville's theorem, vanishes. The conclusion is therefore that

$$u(z) = P(z) [Q(z^{-1})v(z) + Q(z^{-1})z^x\psi(z)]_+ \quad (103)$$

$$\psi(z^{-1}) = -\frac{1}{Q(z^{-1})} [Q(z^{-1})v(z) + Q(z^{-1})z^x\varphi(x)]_-$$

The argument can be repeated for $z^xu(z^{-1})$ analytic inside the unit disk with the result that

$$z^xu(z^{-1}) = Q(z) [P(z^{-1})z^xv(z^{-1}) + P(z^{-1})z^x\psi(z)]_+ \quad (104)$$

$$\varphi(z^{-1}) = -\frac{1}{P(z^{-1})} [P(z^{-1})z^xv(z^{-1}) + P(z^{-1})z^x\psi(x)]_-$$

In the case of interest, $v(z) = 1$. As $x \rightarrow \infty$, we can neglect in the second equation (103) the term $[Q(z^{-1})z^x\varphi(z)]_-$. Furthermore $[Q(z^{-1})v(z)]_- = [Q(z^{-1})]_- = Q(z^{-1}) - 1$. Thus $\psi(z) \simeq 1/Q(z) - 1$. This value is then inserted in (104) with the result that, for large x ,

$$z^xu(z^{-1}) \simeq Q(z) \left[\frac{P(z^{-1})z^x}{Q(z)} \right]_+$$

Finally, the quantity of interest is u_x , i.e. the value of the above function at $z = 0$, and referring to (95) and (96)

$$\langle \sigma_0 \sigma_x \rangle \xrightarrow{x \rightarrow \infty} \left[(1-t^2)^2 (1-t^2\bar{t}^2) (1-t^2\bar{t}^{-2}) \right]^{1/4} \times \quad (105)$$

$$\oint \frac{dz}{2i\pi} z^{x-1} \left[\left(1 - \frac{t}{\bar{t}} z \right) \left(1 - \frac{t}{\bar{t}} z^{-1} \right) (1 - t\bar{t}z) (1 - t\bar{t}z^{-1}) \right]^{-1/2}$$

with exponentially small corrections for $x \rightarrow \infty$. The evaluation of the integral is obtained by contour deformation and leads to the final result that in the high temperature phase the correlation

function decreases asymptotically with the distance as

$$\langle \sigma_0 \sigma_x \rangle \xrightarrow{x \rightarrow \infty} \left(\frac{1-t^2\bar{t}^2}{1-t^2\bar{t}^{-2}} \right)^{1/4} \frac{1}{(1-t^2)^{1/2}} \frac{1}{(\pi|x|)^{1/2}} \left(\frac{t}{\bar{t}} \right)^x [1 + O(1/x)] \quad (106)$$

We recall that here 0 and x are points on the same crystallographic axis. It is not expected that far from criticality we find an isotropic law. However, if we are close to T_c , such should be the case, and (106) can be compared to the behaviour of a massive boson propagator

$$\int \frac{d^2k}{(2\pi)^2} \frac{e^{ik \cdot x}}{k^2 + m^2} \simeq \left(\frac{2}{\pi m |x|} \right)^{1/2} e^{-m|x|} \left(1 + O\left(\frac{1}{m|x|} \right) \right) \quad (107)$$

In both cases we find an exponential decay law with characteristic length $\xi = m^{-1}$, corrected by a prefactor in $|x|^{-1/2}$.

At any rate, we conclude that the $(x-x)$ or $(y-y)$ correlation length is given by

$$\exp -1/\xi = t/\bar{t} = \exp[-2(\bar{\beta} - \beta)] \quad (108)$$

$$\xi = 1/2(\bar{\beta} - \beta)$$

an expression only valid of course for $\beta < \beta_c$ or $t < \sqrt{2} - 1$. As a consequence the correlation length diverges linearly as $T \rightarrow T_c$ from above

$$\xi \sim \left(\frac{T}{T_c} - 1 \right)^{-1} \quad (109)$$

The critical exponent ν is defined through $\xi \sim [(T/T_c) - 1]^{-\nu}$, meaning in the present case that

$$\nu = 1 \quad (110)$$

The expression (106) describes bosonic excitations with a mass $m_B = 2(\bar{\beta} - \beta)$. Let us compare this scale with the fermionic propagator, the denominator of which is proportional to

$$[2 \sinh(\bar{\beta} - \beta)]^2 + 2(1 - \cos p_x) + 2(1 - \cos p_y)$$

the normalization being chosen in such a way that, for long wavelengths, the behaviour is $m_F^2 = p_x^2 + p_y^2$. Therefore

$$m_F^2 = [2 \sinh(\bar{\beta} - \beta)]^2$$

Thus as $\beta \rightarrow \tilde{\beta}$, we find $m_F \sim m_B$. Finally the relation $\exp(-1/\xi) = t/\tilde{t} = t(1+t)/(1-t)$ is to be compared with the analogous one valid in one dimension, $\exp(-1/\xi_1) = t = \exp -2\tilde{\beta}$, $\xi_1 = 1/2\tilde{\beta}$, which diverges as $\frac{1}{2} \exp 2\beta$ as $\beta \rightarrow \infty$ or $T \rightarrow 0$.

The relation $\exp(-1/\xi) = t(1+t)/(1-t) = t + 2t^2 + 2t^3 + \dots$ suggests the following approximation for the correlation function. Recall the high temperature expansion

$$\langle \sigma_0 \sigma_x \rangle = \frac{\sum_{\partial\{n\}=(0,x)} t^{\sum n_i}}{\sum_{\partial\{n\}=\emptyset} t^{\sum n_i}}$$

with $n_i = 0$ or 1 . Let us consider an anisotropic model with t_x assigned to horizontal links (parallel to the direction $0x$) and t_y to the vertical ones. Assume a limit (the so called "solid-on-solid" limit) where $t_x/t_y \rightarrow 0$. In the denominator no term different from unity survives in this limit. In the numerator the dominant terms have t_x^x as a factor, the corresponding contributions being depicted as a broken curve with no overhangs nor disconnected parts of the type shown on figure 1. It is easy to show that in this approximation

$$\langle \sigma_0 \sigma_x \rangle_{\text{sos}} = \left(t_x \frac{1+t_y}{1-t_y} \right)^x$$

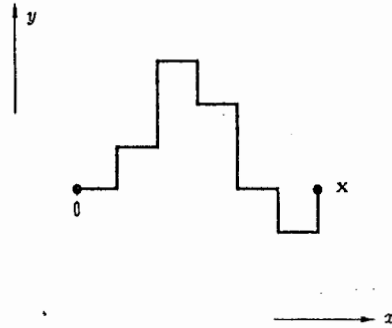


Fig. 1 A typical "solid-on-solid" curve joining 0 and x

It is intriguing to find that, if one is only interested in computing $\exp(-1/\xi)$, i.e. in the large x limit, it remains valid up to the isotropic case $t_x = t_y$.

Having found that the mass, or inverse correlation length, vanishes linearly with $T - T_c$, and assuming that, close to criticality, the correlation function becomes isotropic, we can represent its asymptotic behaviour up to a factor by the integral (107). This prefactor is in fact proportional to $m^{1/4}$. Thus

$$\langle \sigma_0 \sigma_x \rangle \simeq \text{cst } m^{1/4} \int \frac{d^2 k}{(2\pi)^2} \frac{e^{ik \cdot x}}{k^2 + m^2} \quad m \equiv m_B \sim T - T_c \quad (111)$$

The susceptibility χ is the integral of $\langle \sigma_0 \sigma_x \rangle$ over x . Thus

$$\chi = \sum_x \langle \sigma_0 \sigma_x \rangle \sim m^{-7/4} \sim (T - T_c)^{-7/4} \quad (112)$$

It diverges at criticality as a power law $(T - T_c)^{-\gamma}$ and one finds in the present case

$$\gamma = \frac{7}{4} \quad (113)$$

In concluding this section, let us consider the correlation function in the limit $T \rightarrow T_c + 0$. We shall discuss more general expressions for multipoint correlation functions in section 3. Let ξ and x be much larger than the lattice spacing, but x much smaller than ξ , in such a way that $\exp(-|x|/\xi)$ can safely be replaced by unity. We then proceed to let $T \rightarrow T_c$. This means that x scales as $1/(1 - t/\tilde{t})$ and, returning to equation (106), we find

$$\langle \sigma_0 \sigma_x \rangle \sim \frac{1}{|x|^{1/4}} \quad (114)$$

Assuming again isotropy at criticality, the scaling behaviour for such a correlation is parametrized as $1/|x|^{d-2+\eta}$. Thus the critical exponent η is

$$\eta = \frac{1}{4} \quad (115)$$

This result can be recovered directly from the determinantal expression for the correlation function. The function $g(z)$ defined in (93) reduces at criticality to

$$g^2(z) = \frac{1-z^{-1}}{1-z} \times \frac{1-t_c^2 z^{-1}}{1-t_c^2 z}$$

The Fourier coefficients of g are asymptotically dominated by those of the first factor with a phase discontinuity of π as we travel around the unit circle. Let us altogether neglect the second factor. It can be

shown that this amounts to replacing the computation of $\langle \sigma_0 \sigma_x \rangle$ by the one corresponding to a diagonal correlation $\langle \sigma_{00} \sigma_{xx} \rangle$. Within this approximation, we have $g(e^{ip}) = -ie^{ip/2}$. The corresponding Fourier coefficients are

$$a_k = -i \int_0^{2\pi} \frac{dp}{2\pi} \exp[-i(k-1)p - ip/2] = \frac{2}{\pi(1-2k)}$$

To obtain the corresponding determinant, we use a result of Cauchy which states that

$$\det \frac{1}{(a_i + b_j)} = \prod_{i < j} \frac{(a_i - a_j)(b_i - b_j)}{\prod_{i,j} (a_i + b_j)}$$

Thus

$$\langle \sigma_{00} \sigma_{xx} \rangle = \left(\frac{2}{\pi} \right)^x \prod_{0 < k < q \leq x-1} [4(k-q)(q-k)] \frac{1}{\prod_{k,q=0}^{x-1} (1-2(k-q))}$$

Grouping terms by pairs in the denominator, this simplifies to

$$\langle \sigma_{00} \sigma_{xx} \rangle = \left(\frac{2}{\pi} \right)^x \prod_{k=1}^{x-1} \left(1 - \frac{1}{4k^2} \right)^{k-x}$$

Recall that

$$\frac{\sin \pi a}{\pi a} = \prod_{k=1}^{\infty} \left(1 - \frac{a^2}{k^2} \right)$$

and set $a = 1/2$ to find

$$\frac{2}{\pi} = \prod_{k=1}^{\infty} \left(1 - \frac{1}{4k^2} \right).$$

Thus

$$\begin{aligned} \langle \sigma_{00} \sigma_{xx} \rangle &= \prod_{k=1}^{\infty} \left(1 - \frac{1}{4k^2} \right)^x \prod_{k=1}^{x-1} \left(\frac{1}{1-4k^2} \right)^{k-x} \\ &= \exp \left\{ \sum_{k=1}^{x-1} k \ln \left(1 - \frac{1}{4k^2} \right) + x \sum_{k=x}^{\infty} \ln \left(1 - \frac{1}{4k^2} \right) \right\} \\ &\simeq \exp - \sum_{k=1}^{x-1} \frac{1}{4k} + \text{cst} \approx \frac{\text{cst}}{x^{1/4}} \end{aligned}$$

which is the expected behaviour.

As suggested by the previous examples, numerous quantities can be analytically evaluated for the two-dimensional Ising model, at the price of an ever increasing amount of mathematical sophistication. An important point for our future development is that one can check in detail the restoration of continuous (geometrical) symmetries as one enters the critical domain. For instance, correlation functions become isotropic on the scale of the correlation length, when the latter becomes larger and larger compared to the lattice spacing, as studied in detail by B. McCoy, T.T. Wu and their collaborators. Further elaboration is to be found in the work of Sato, Jimbo and Miwa, who used powerful mathematical tools relating the computation of correlation functions to the Riemann-Hilbert problem. The crucial result is the existence of a critical continuous domain, where the underlying field theory is a free fermionic one, with scaling laws characterized by critical exponents. We return to the continuous theory in the last section of this chapter.

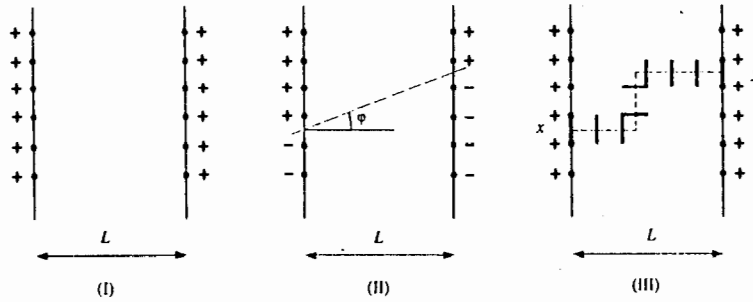
2.2.7 Surface tension

At low temperature, when two macroscopic pure phases with opposite magnetization are brought into thermal equilibrium, one expects an additional contribution to the free energy proportional to the area (in three dimensions) or to the length (in two dimensions) of the interface. To compute this quantity in the present case, we can proceed as follows. Figure 2 shows a strip of width L under two sets of boundary conditions. In case I, these are uniform, say $\sigma = +1$ on the boundary. In case II, we impose mixed boundary conditions, $\sigma = +1$ above a certain line at angle ϕ , $\sigma = -1$ below. Case I will correspond in the thermodynamic limit ($L \rightarrow \infty$) to an homogeneous pure phase, while we expect an inhomogeneous mixture in case II.

We define the surface tension through the thermodynamic limit of the ratio Z_{II}/Z_I . More precisely, and restoring the conventional definition,

$$-\beta \Sigma(\varphi) = \lim_{L \rightarrow \infty} \frac{\cos \varphi}{L} \ln \frac{Z_{II}}{Z_I} \quad (116a)$$

The tension depends on the angle φ and we note that $L/\cos \varphi$ is the length of a presumed flat interface. We can replace

Fig. 2 Comparison of boundary conditions on a strip of width L .

the calculation of Z_{II} with mixed boundary conditions by an equivalent frustrated partition function with uniform boundary conditions. This Z_{III} is obtained by drawing an arbitrary line between two sites x and y of the dual lattice, chosen at the left (respectively right) of the link with opposite boundary conditions. One reverses the sign of the coupling, $\beta \rightarrow -\beta$, on each link of the original lattice crossing the frustration line. These are the links shown in case III on figure 2. Alternatively, this would follow starting from case II if we change all variables $\sigma \rightarrow -\sigma$ below the frustration line in the computation of Z_{II} . Except from being fixed at its end points, the line joining x to y is arbitrary. Thus we also have

$$-\beta\Sigma(\varphi) = \lim_{L \rightarrow \infty} \frac{\cos \varphi}{L} \ln \frac{Z_{III}}{Z_I} \quad (116b)$$

This suggests a more general definition of the surface tension Σ in the two-dimensional case. Consider in the infinite volume limit two points x and y on the dual lattice and assume that $y - x$ is at an angle φ with the x -axis. Join these two points by a line drawn on the dual lattice and change $\beta \rightarrow -\beta$ on each link crossing this line. This is described as having set two disorder operators at x and y . Denote these operators as $\tilde{\sigma}_x$ and $\tilde{\sigma}_y$, observing that except for its fixed boundary the location of the frustration line is again arbitrary. Then, in the thermodynamic limit, if Z_{xy} denotes the frustrated partition function

$$\langle \tilde{\sigma}_x \tilde{\sigma}_y \rangle = \frac{Z_{xy}}{Z} \quad (117)$$

and

$$-\beta\Sigma(\varphi) = \lim_{|x-y| \rightarrow \infty} \frac{\ln \langle \tilde{\sigma}_x \tilde{\sigma}_y \rangle}{|x-y|} \quad (118a)$$

The above expression implies that $\beta\Sigma(\varphi)$ is identified with the inverse correlation length for disorder operators (in the ordered phase $T < T_c$). Asymptotically

$$\langle \tilde{\sigma}_x \tilde{\sigma}_y \rangle \sim \exp -\beta\Sigma(\varphi) |x-y| \quad (118b)$$

The terminology of disorder operators is due to Kadanoff and Ceva. Off the critical region $\Sigma(\varphi)$ is expected to depend on φ (and on the lattice, here a square one) exhibiting the general anisotropy of the underlying setting. Of course, $\Sigma(\varphi)$ reflects the symmetries of this lattice, i.e. in our case

$$\Sigma(\varphi) = \Sigma(-\varphi) = \Sigma(\pi - \varphi) = \Sigma(\frac{1}{2}\pi - \varphi) \quad (119)$$

The last equality might not at first be obvious in the strip geometry where it was first defined.

Let us insist on the fact that the frustration line is arbitrary except for its boundary. It can therefore be chosen in such a way as to minimize the number of frustrated links. One can indeed change a spin σ into $-\sigma$ in the computation of any partition function, thus frustrating four lines originating from the relevant site, or equivalently introducing a closed frustration curve around a plaquette of the dual lattice. Since two successive frustrations amount to no frustration at all, this means that any closed contour of frustration (on the dual lattice) can be introduced at no cost and therefore two distinct frustration lines having as common boundary the two points x, y are equivalent. As readily verified, this extends to the introduction of an even number of frustration operators (and corresponding frustration lines).

As $T \rightarrow 0$, or $\beta \rightarrow \infty$, the energy "cost" of a frustrated link is -2β . Hence in the limit $|x-y| \rightarrow \infty$, $\Sigma(\varphi)$ tends to twice the ratio of the minimal number of frustrated links to the distance $|x-y|$. Per unit distance, the number of vertical and horizontal frustrated links in the minimal case are respectively $|\sin \varphi|$ and $|\cos \varphi|$. Thus

$$\lim_{T \rightarrow 0} \Sigma(\varphi) = 2(|\cos \varphi| + |\sin \varphi|) \quad (120)$$

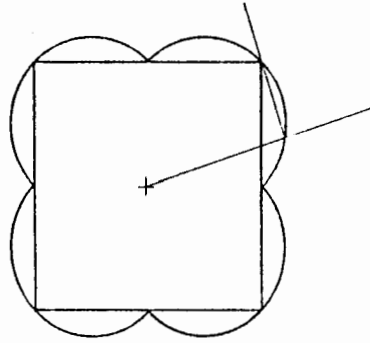


Fig. 3 Surface tension $\Sigma(\varphi)$ at zero temperature. The surface tension is the external envelope of these circles with four cusps along the horizontal and vertical axis.

In polar coordinates, this is the equation of four circles through the origin and centered at the points $(\pm 1, \pm 1)$ (Figure 3).

The duality property of the two-dimensional Ising model on a square lattice described in section 2.1 allows one to relate correlation functions of disorder operators $\tilde{\sigma}$ to those of the spins σ (the order operators of course, according to the same terminology) at the dual temperature. Indeed one has only to compare their low and high temperature expansions respectively to find

$$\langle \tilde{\sigma}_{\mathbf{x}} \tilde{\sigma}_{\mathbf{y}} \rangle_{\beta} = \langle \sigma_{\mathbf{x}} \sigma_{\mathbf{y}} \rangle_{\tilde{\beta}} \quad (121)$$

both at equal relative coordinates $\mathbf{x} - \mathbf{y} = \mathbf{x}' - \mathbf{y}'$. Comparing the asymptotic behaviour at large separation, we get

$$\beta \Sigma(\varphi) = \frac{1}{\xi_{\tilde{\beta}}(\varphi)} \quad (122)$$

where ξ stands for the correlation length for spins. In particular it follows that the surface tension vanishes at β_c with the same exponent ν with which ξ diverges. Up to now we had only computed ξ along the main crystallographic axis. But the same method extends to arbitrary directions (Cheng and Wu, 1967). From its symmetry relations, it is sufficient to compute $\Sigma(\varphi)$ for φ between 0 and $\pi/4$

$$\beta \Sigma(\varphi) = \rho(\varphi) \cos \varphi + \rho(\pi/2 - \varphi) \sin \varphi$$

$$\cos 2\varphi \cosh \rho(\varphi) = 2 \cosh^2(\beta - \tilde{\beta}) \cos^2 \varphi - \left(\cosh^4(\beta - \tilde{\beta}) \sin^2 2\varphi + \cos^2 2\varphi \right)^{1/2} \quad (123)$$

When φ is equal to 0 or $\pi/2$ we recover $\xi_{\tilde{\beta}}^{-1} = 2(\beta - \tilde{\beta})$, while for $\varphi = \pi/4$,

$$\cosh \frac{\beta \Sigma}{\sqrt{2}} = \cosh \frac{1}{\sqrt{2} \xi_{\tilde{\beta}}} = \cosh^2(\beta - \tilde{\beta})$$

As $\beta \rightarrow \infty$, $\rho(\varphi) \sim \rho(\pi/2 - \varphi) \sim 2\beta$ and equation (123) reduces to the previous result (120). It is remarkable that, for any finite temperature, $\rho(\pi/2)$ vanishes and the cusps in $\Sigma(\varphi)$ disappear. Finally, in the vicinity of β_c , $\rho(\varphi) \sim 2(\beta - \tilde{\beta}) \cos \varphi$, and the surface tension (in the low temperature phase) as well as the correlation length (in the high temperature phase) become isotropic

$$\beta \rightarrow \beta_c \quad \beta \Sigma(\varphi) = 1/\xi_{\tilde{\beta}} \sim 2(\beta - \tilde{\beta}) \quad (124)$$

The same is true of the correlation functions in the scaling regime, showing the emergence of an isotropic (i.e. rotationally invariant) continuous theory. This is illustrated by the equilibrium shape of crystals, here two-dimensional. A side remark is that equilibrium shapes of crystals are seldom observed in nature. But under favorable conditions, they can be obtained in the laboratory for specific materials, generally at rather low temperature.

The equilibrium shape results from a competition between the bulk free energy and the surface tension. In two dimensions and at a fixed temperature, one wants therefore to minimize $\int \Sigma(\varphi) ds$ given the total area. Choosing an arbitrary origin, the element of area is $\frac{1}{2} |\mathbf{x} \wedge d\mathbf{s}| = \frac{1}{2} \mathbf{x} \cdot \mathbf{n} ds$, $\mathbf{n}$ standing for the normal to the bounding curve (Figure 4).

Should $\Sigma(\varphi)$ be independent of the direction φ , the equilibrium shape would be a circle. The dependence on φ or equivalently on the normal $\mathbf{n}$, can then be accounted for by minimizing

$$\int \Sigma(\mathbf{n}) ds - \lambda \int \mathbf{x} \cdot \mathbf{n} ds \quad (125)$$

using a Lagrange multiplier λ for the constraint of fixed area. The quantity to be minimized is thus the integral of $\Sigma - \lambda \mathbf{x} \cdot \mathbf{n}$. Let us describe a geometrical construction (Wulff 1901) yielding the desired result. Let us draw a representation of Σ in polar

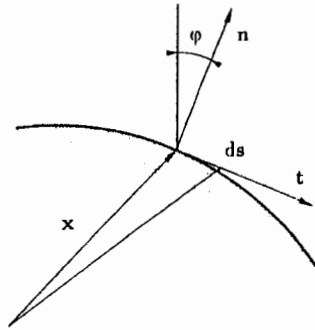


Fig. 4 Geometry for the minimization of the surface tension.

coordinates by following the point $\mathbf{n}\Sigma(\mathbf{n})$. From a point on this curve, the straight line $\Sigma - \mathbf{x} \cdot \mathbf{n}$ represents a perpendicular to the radius vector, absorbing the Lagrange multiplier in a rescaling of coordinates (or of Σ). The equilibrium shape is the envelope of these successive perpendicular lines (figure 5) up to a scale. At a point where the curve Σ is regular, let φ be the angle of $\mathbf{n}$ with the y -axis $\mathbf{n} = (-\sin \varphi, \cos \varphi)$. Let $\mathbf{t}$ denote the orthogonal unit vector tangent to the envelope, $\mathbf{t} = (\cos \varphi, \sin \varphi)$.

By minimizing (125), we find that the equation of the latter is

$$\mathbf{x} = -\mathbf{t} \frac{d\Sigma}{d\varphi} + \mathbf{n}\Sigma \quad (126)$$

When Σ is stationary, $\mathbf{x}$ is along $\mathbf{n}$. If Σ is a constant, the shape is a circle. Finally, if the slope of the Σ -curve is discontinuous (due to the presence of cusps), $d\Sigma/d\varphi$ has a discontinuity and $\Delta(\mathbf{x} \cdot \mathbf{t}) = \Delta|d\Sigma/d\varphi|$. The location of the point $\mathbf{x}$ on the line perpendicular to $\mathbf{n}$ has a jump proportional to $\Delta|d\Sigma/d\varphi|$ corresponding to a planar face on the crystal. Thus cusps in $\Sigma(\varphi)$ lead to planar faces, an example of which was given at zero temperature. The envelope was then given by the four points $(\pm 2, \pm 2)$ and four faces joining them, yielding a square equilibrium shape. The disappearance of cusps is then a signal of the disappearance of the corresponding faces. The conventional name is that the face becomes rough. At any finite temperature smaller than T_c in the case of the two-dimensional Ising model, interfaces are rough (at a

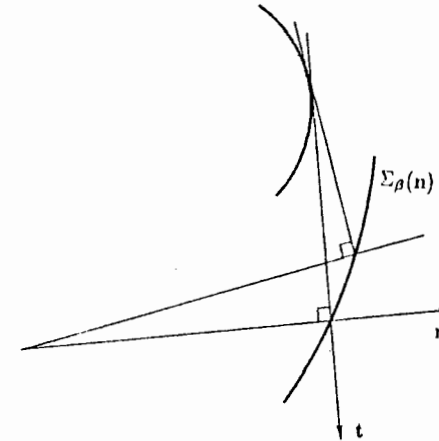


Fig. 5 The Wulff construction.

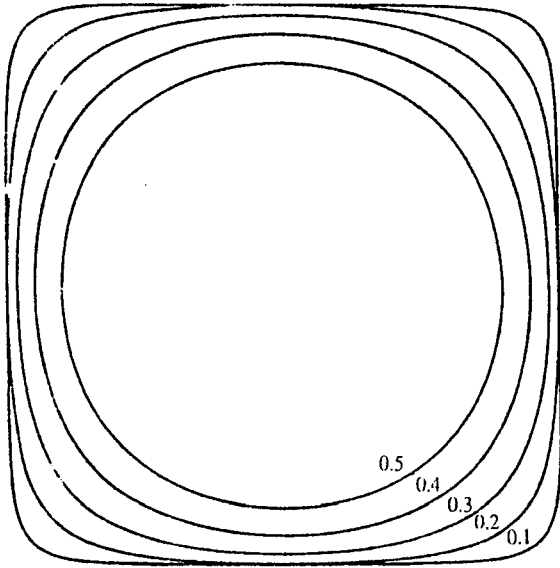
molecular scale) i.e. on a macroscopic scale the equilibrium shape is a continuous curve which does not (except in its overall shape) reveal the direction of the crystallographic axes. The equilibrium curves shown on figure 6 provide a visual demonstration for the progressive rounding off of the crystal as it tends to a circle when $T \rightarrow T_c$.

Study the curvature of the equilibrium shapes close to the would be corners as $T \rightarrow 0$.

2.3 Critical continuous theory

2.3.1 Effective action

Let us return to the Grassmannian path integral with action (69) and ask how to simplify it in the vicinity of the critical temperature. We saw that out of the four sets of modes, two remained massive at T_c . We want to decouple them from the genuine critical ones. This is accomplished if we perform a linear

Fig. 6 Successive equilibrium shapes as a function of T/T_c .

transformation of the variables ξ of the following type

$$\begin{aligned}\bar{\xi}^H &= \frac{1}{2} \sum_{s=\pm 1, \pm 3} e^{is\pi/4} \psi_s & \bar{\xi}^V &= -\frac{1}{2} \sum_{s=\pm 1, \pm 3} e^{2is\pi/4} \psi_s \\ \xi^H &= -\frac{1}{2} \sum_{s=\pm 1, \pm 3} e^{3is\pi/4} \psi_s & \xi^V &= \frac{1}{2} \sum_{s=\pm 1, \pm 3} e^{4is\pi/4} \psi_s\end{aligned}\quad (127)$$

As $t \rightarrow t_c = \sqrt{2} - 1$, the ψ_s modes with $s = \pm 3$ remain massive and do not contribute to the critical singularities. We can therefore define an effective action in two steps. First neglect $\psi_{\pm 3}$, and arrange the remaining components in a two component spinor to obtain

$$S \rightarrow S' = \sum_{\mathbf{x}} t_c \frac{1}{2} \psi(\mathbf{x}) \sigma_2 \psi(\mathbf{x}) - \frac{1}{4} t \psi(\mathbf{x} + \mathbf{u}_1) (1 + \sigma_2) \psi(\mathbf{x}) - \frac{1}{4} t \psi(\mathbf{x} + \mathbf{u}_2) (i\sigma_3 + \sigma_2) \psi(\mathbf{x}) \quad (128)$$

with $\mathbf{u}_1$ and $\mathbf{u}_2$ unit vectors (in lattice spacing units) along the x and y axes. Then we proceed to a formal continuous limit where sums are replaced by integrals, ψ by a continuous differentiable spinor and $\psi(\mathbf{x} + \mathbf{u})$ is expanded up to first order as $\psi(\mathbf{x}) + \mathbf{u} \cdot \nabla \psi(\mathbf{x}) + \dots$. This second step amounts to the replacement $S' \rightarrow S_{\text{eff}}$. Recall that ψ 's are anticommuting variables, and use instead of $\mathbf{x} = (x, y)$ a complex notation $\mathbf{x} = (z, \bar{z})$ with

$$z = x + iy \quad \text{and} \quad \partial \equiv \frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right)$$

the Laplacian being $4\partial\bar{\partial}$. We find

$$S' \rightarrow S_{\text{eff}} \quad (129)$$

$$= \frac{1}{2} t \int d^2 z \left\{ \psi_+ \bar{\partial} \psi_+ + \psi_- \partial \psi_- + \left(\frac{t - t_c}{t} \right) (i\psi_+ \psi_- - i\psi_- \psi_+) \right\}$$

The field (ψ_+, ψ_-) is referred to as a real (or Majorana) two-dimensional Fermi field. At criticality the two components decouple, while in the vicinity of this point they are coupled by a "mass" term, the mass being proportional to $t - t_c$. Finally, we rescale the field by absorbing a $\sqrt{t}$ factor in ψ . It is convenient to set

$$\sqrt{\pi t} \psi_+ = \psi \quad \sqrt{\pi t} \psi_- = \bar{\psi} \quad \frac{t - t_c}{t} = -\frac{m}{2} \quad (130)$$

with ψ and $\bar{\psi}$ standing now for one component fields, and m agreeing with our previous definition of m_F as $t \rightarrow t_c$. Then

$$S_{\text{eff}} = \int \frac{d^2 z}{2\pi} \{ (\psi \bar{\partial} \psi + \bar{\psi} \partial \bar{\psi}) + im \bar{\psi} \psi \} \quad (131)$$

The corresponding equation of motion, i.e. the two-dimensional analog of the Dirac equation, is the pair

$$\bar{\partial} \psi - \frac{1}{2} im \bar{\psi} = 0 \quad \partial \bar{\psi} + \frac{1}{2} im \psi = 0 \quad (132a)$$

and as a consequence each component satisfies

$$(-4\partial\bar{\partial} + m^2) \psi \equiv (-\Delta + m^2) \psi = 0 \quad (132b)$$

The corresponding $\psi, \bar{\psi}$ propagator is the inverse of the matrix valued operator

$$\frac{1}{\pi} \begin{pmatrix} \bar{\partial} & -\frac{im}{2} \\ \frac{im}{2} & \partial \end{pmatrix},$$

Measured in units of inverse length, this means that the dimension of ε is 1. More generally we have

$$\langle \varepsilon_1 \cdots \varepsilon_{2p} \rangle = (-1)^p \langle \bar{\psi}_1 \psi_1 \cdots \bar{\psi}_p \psi_p \rangle = \text{Pf} \frac{1}{z_i - z_j} \text{Pf} \frac{1}{\bar{z}_i - \bar{z}_j} \quad (142)$$

Using the identity (137) we can obtain an interesting interpretation of this expression by squaring it, namely

$$\langle \varepsilon_1 \cdots \varepsilon_{2n} \rangle^2 = \text{Hf} \frac{1}{(z_i - z_j)^2} \text{Hf} \frac{1}{(\bar{z}_i - \bar{z}_j)^2} \quad (143)$$

This invites us to compare the model obtained by taking two distinct copies of the Ising model with a massless bosonic theory. In analogy with the effective fermionic action (131), we could consider a bosonic one, written omitting a mass term, as

$$S_{\text{bosonic}} = \int \frac{d^2 z}{8\pi} \varphi (-\Delta \varphi) = \int \frac{d^2 z}{2\pi} \bar{\partial} \varphi \partial \varphi \quad (144)$$

Using an arbitrary unit of length, the corresponding propagator

$$\langle \varphi_1 \varphi_2 \rangle = \ln \left[\frac{1}{(z_1 - z_2)(\bar{z}_1 - \bar{z}_2)} \right] \quad (145)$$

is of course unbounded at large distance and cannot therefore describe a decent observable. But φ should merely be understood as an intermediate quantity enabling one, through derivation or exponentiation, to construct bona fide observables. From (145), it follows that

$$\begin{aligned} \langle \partial \varphi_1 \partial \varphi_2 \rangle &= -\frac{1}{(z_1 - z_2)^2} & \langle \bar{\partial} \varphi_1 \bar{\partial} \varphi_2 \rangle &= -\frac{1}{(\bar{z}_1 - \bar{z}_2)^2} \\ \langle \partial \varphi_1 \bar{\partial} \varphi_2 \rangle &= 0 \end{aligned} \quad (146)$$

so that

$$\langle \varepsilon_1 \varepsilon_2 \rangle^2 = \langle (\partial \varphi \bar{\partial} \varphi)_1 (\partial \varphi \bar{\partial} \varphi)_2 \rangle = \frac{1}{(z_1 - z_2)^2 (\bar{z}_1 - \bar{z}_2)^2} \quad (147a)$$

and more generally

$$\begin{aligned} \langle \varepsilon_1 \cdots \varepsilon_{2p} \rangle^2 &= \langle (\partial \varphi \bar{\partial} \varphi)_1 \cdots (\partial \varphi \bar{\partial} \varphi)_{2p} \rangle \\ &= \text{Hf} \frac{1}{(z_i - z_j)^2} \text{Hf} \frac{1}{(\bar{z}_i - \bar{z}_j)^2} \end{aligned} \quad (147b)$$

The interpretation is then that by doubling the Fermi theory, the operator $\varepsilon \varepsilon'$ (prime denotes the second copy) can be thought of as $\partial \varphi \bar{\partial} \varphi$, or $\psi \psi'$ as $\partial \varphi$, both functions of z alone, and $\bar{\psi}' \bar{\psi}$ as $\bar{\partial} \varphi$. We achieve therefore a *bosonization* of the duplicated Majorana theory which may be thought of as describing a complex Fermi field. The quantities

$$\begin{aligned} J &= \psi \psi' = \partial \varphi & \bar{\partial} J &= 0 \\ \bar{J} &= \bar{\psi}' \bar{\psi} = \bar{\partial} \varphi & \partial \bar{J} &= 0 \end{aligned} \quad (148)$$

are then conserved currents (since $\bar{\partial} \partial \varphi = 0$) associated to an invariance under an Abelian rotation group in "isotopic" space (ψ, ψ') through Noether's theorem.

We record here also the analogous expressions for the spin operators, referring to the literature for a complete discussion. It is convenient to use the following normalization of the two-point function

$$\langle \sigma_1 \sigma_2 \rangle = \frac{\sqrt{2}}{|z_{12}|^{1/4}} \quad (149)$$

Then the squared correlation functions may be written as

$$\langle \sigma_1 \cdots \sigma_{2n} \rangle^2 = \sum_{\epsilon_i = \pm 1, \epsilon_1 + \cdots + \epsilon_{2n} = 0} \prod_{1 \leq i < j \leq 2n} |z_{ij}|^{\epsilon_i \epsilon_j / 2} \quad (150)$$

In terms of the boson field φ we have

$$\langle : e^{i\alpha \varphi_1} :: e^{-i\alpha \varphi_2} : \rangle = \frac{1}{(z_1 - z_2)^{\alpha^2} (\bar{z}_1 - \bar{z}_2)^{\alpha^2}} \quad (151)$$

where the $::$ symbol implies the omission of self-contractions, and mean values of such exponentials are understood to vanish whenever the total "charge" i.e. the sum of coefficients of φ 's, is nonzero. Choosing $\alpha = \frac{1}{2}$, we find the identification

$$\langle \sigma_1 \cdots \sigma_{2n} \rangle^2 = \langle : e^{i\varphi_1/2} + e^{-i\varphi_1/2} : \cdots : e^{i\varphi_{2n}/2} + e^{-i\varphi_{2n}/2} : \rangle \quad (152)$$

or in operator form

$$\sigma \sigma' = : e^{i\varphi/2} + e^{-i\varphi/2} : \quad (153)$$

Starting from the original formulation of the Ising model and the Jordan-Wigner transformation, derive the representation (150)–(153) for the spin correlation function.

The above continuum formulation will be expanded in chapter 9 when we discuss the general two-dimensional critical theories.

Appendix 2.A Quadratic differences and Painlevé equations

The crowning piece of the work on the two-dimensional Ising model is a set of quadratic difference equations obtained by McCoy, Perk and Wu, which determine the lattice correlation functions at any separation and any temperature. These relations reduce in the vicinity of the critical point to Painlevé equations for the two-point functions and generalizations thereof for higher correlations. A similar discussion has been given by Sato, Jimbo and Miwa, in a more mathematical setting. It would be unfair to the subject not to briefly mention these aspects, even though their developments (in the framework of integrable models) will not be pursued here. We shall limit ourselves to the case of the spin-spin correlation function, although the method extends to higher point functions. As we shall see, duality plays a crucial role in the derivation.

Let us return to the transfer matrix operating on a row of size L (tending to infinity). Consider the correlation functions of two spins located at (x_0, y_0) and (x_1, y_1) . Assuming for instance $y_1 \geq y_0$

$$\langle \sigma_{x_1 y_1} \sigma_{x_0 y_0} \rangle = \lim_{M, L \rightarrow \infty} \frac{\text{Tr} \left(T^{M-y_1} \sigma_{x_1}^{(1)} T^{y_1-y_0} \sigma_{x_0}^{(1)} T^{y_0} \right)}{\text{Tr} T^M} \quad (154)$$

Recall from the definitions (56) that

$$\begin{aligned} \sigma_x^{(1)} &= \sigma_{x-1/2}^{(1)} \Gamma_{x-1/2, x} \\ \sigma_{x-1/2}^{(1)} &= \prod_{i=0}^{x-1} \frac{1}{i} \Gamma_{k-1/2, k} \Gamma_{k, k+1/2} = \prod_{i=0}^{x-1} \sigma_k^{(3)} \end{aligned} \quad (155)$$

where $\sigma_{x-1/2}^{(1)}$ commutes with $\Gamma_{x-1/2, x}$ and can also be expressed as an exponential of a quadratic form in the Γ 's since $\exp \frac{1}{2} \pi \Gamma_1 \Gamma_2 =$

$\Gamma_1 \Gamma_2$. Set

$$\begin{aligned} A &= T^{y_1-y_0} & B &= T^{M-(y_1-y_0)} \\ \tilde{A} &= \sigma_{x_1-1/2}^{(1)} T^{y_1-y_0} \sigma_{x_0-1/2}^{(1)} & AB &= T^M \end{aligned} \quad (156)$$

Then

$$\langle \sigma_{x_1 y_1} \sigma_{x_0 y_0} \rangle = \lim_{L, M \rightarrow \infty} \frac{\text{Tr} \left(\Gamma_{x_1-1/2, x_1} \tilde{A} \Gamma_{x_0-1/2, x_0} B \right)}{\text{Tr} AB} \quad (157)$$

where $A, \tilde{A}, B$ are all exponentials of quadratic forms in the Clifford (or Fermi) algebra. We now use a variant of Wick's theorem in operator form which says the following. Let U and V be exponentials of quadratic forms in the Γ 's and O_1 and O_2 monomials in the Γ 's. Then

$$\frac{\text{Tr} O_1 U O_2 V}{\text{Tr} UV} = \text{Sum over products of contractions}$$

where by contractions we understand either $\text{Tr} \Gamma_1 \Gamma_2 UV / \text{Tr} UV$, $\text{Tr} \Gamma_1 \Gamma_2 VU / \text{Tr} UV$, or $\text{Tr} \Gamma_1 U \Gamma_2 V / \text{Tr} UV$ if Γ_1 and Γ_2 both belong to O_1 or to O_2 , or one to O_1 and the other to O_2 respectively. A sign depending of the permutation of the Γ 's is understood on the r.h.s.. Clearly this property follows from a path integral representation. Apply this with $U = \tilde{A}$ and $V = B$ to obtain

$$\begin{aligned} \text{Tr} \tilde{A} B &\times \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \Gamma_{x_1, x_1+1/2} \tilde{A} \Gamma_{x_0-1/2, x_0} \Gamma_{x_0, x_0+1/2} B \right) = \\ &= \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \Gamma_{x_1, x_1+1/2} \tilde{A} B \right) \text{Tr} \left(\tilde{A} \Gamma_{x_0-1/2, x_0} \Gamma_{x_0, x_0+1/2} B \right) \\ &\quad - \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \tilde{A} \Gamma_{x_0-1/2, x_0} B \right) \text{Tr} \left(\Gamma_{x_1, x_1+1/2} \tilde{A} \Gamma_{x_0, x_0+1/2} B \right) \\ &\quad + \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \tilde{A} \Gamma_{x_0, x_0+1/2} B \right) \text{Tr} \left(\Gamma_{x_1, x_1+1/2} \tilde{A} \Gamma_{x_0-1/2, x_0} B \right) \end{aligned} \quad (158)$$

Let us divide both sides by $(\text{Tr} AB)^2$ to turn all quantities into mean values, and interpret the six terms, using the Γ -algebra, the correlation functions of spins $\sigma_x^{(1)}$ called $\mathcal{C}$ and the correlation functions of disorder operators $\sigma_{x+1/2}^{(1)}$, or equivalently of spins at the dual temperature, called $\tilde{\mathcal{C}}$.

Thus, in thermodynamic limit ($L, M \rightarrow \infty$)

$$\frac{1}{\text{Tr} AB} \text{Tr} \tilde{A} B = \frac{1}{\text{Tr} AB} \text{Tr} \sigma_{x_1-1/2}^{(1)} A \sigma_{x_0-1/2}^{(1)} B = \tilde{\mathcal{C}}(x_1 - x_0, y_1 - y_0).$$

$$\begin{aligned}
\frac{1}{\text{Tr } AB} \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \Gamma_{x_1, x_1+1/2} \bar{A} B \right) &= \frac{i}{\text{Tr } AB} \text{Tr} \left(\sigma_{x_1+1/2}^{(1)} A \sigma_{x_0-1/2}^{(1)} B \right) \\
&= i \tilde{C}(x_1 - x_0 + 1, y_1 - y_0) \\
\frac{1}{\text{Tr } AB} \text{Tr} \left(\bar{A} \Gamma_{x_0-1/2, x_0} \Gamma_{x_0, x_0+1/2} B \right) &= i \tilde{C}(x_1 - x_0 - 1, y_1 - y_0) \\
\frac{1}{\text{Tr } AB} \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \bar{A} \Gamma_{x_0-1/2, x_0} B \right) &= C(x_1 - x_0, y_1 - y_0)
\end{aligned}$$

With the notations (74), we compute using the equation of motion, i.e. the commutation with the transfer matrix T ,

$$\begin{aligned}
C(x_1 - x_0, y_1 - y_0 + 1) &= \frac{1}{\text{Tr } AB} \text{Tr} \left(T^{-1} \sigma_{x_1}^{(1)} T A \sigma_{x_0}^{(1)} B \right) \\
&= \tilde{C} C(x_1 - x_0, y_1 - y_0) - i \tilde{S} \frac{1}{\text{Tr } AB} \text{Tr} \left(\Gamma_{x_1, x_1+1/2} \bar{A} \Gamma_{x_0-1/2, x_0} B \right)
\end{aligned}$$

and

$$\begin{aligned}
C(x_1 - x_0, y_1 - y_0) &= \left\langle \sigma_{x_1, y_1+1} \sigma_{x_0, y_0+1} \right\rangle \\
&= \tilde{C}^2 \frac{1}{\text{Tr } AB} \text{Tr} \left(T_{x_1-1/2, x_1} \bar{A} \Gamma_{x_0-1/2, x_0} B \right) \\
&\quad - i \tilde{C} \tilde{S} \frac{1}{\text{Tr } AB} \text{Tr} \left(\Gamma_{x_1, x_1+1/2} \bar{A} \Gamma_{x_0-1/2, x_0} B \right) \\
&\quad - i \tilde{C} \tilde{S} \frac{1}{\text{Tr } AB} \text{Tr} \left(T_{x_1-1/2, x_1} \bar{A} \Gamma_{x_0, x_0+1/2} B \right)
\end{aligned}$$

Therefore we get

$$\begin{aligned}
C(x_1 - x_0, y_1 - y_0)^2 - C(x_1 - x_0, y_1 - y_0 - 1) C(x_1 - x_0, y_1 - y_0 + 1) \\
= -\tilde{S}^2 \frac{1}{(\text{Tr } AB)^2} \left[\text{Tr} \left(\Gamma_{x_1-1/2, x_1} \bar{A} \Gamma_{x_0-1/2, x_0} B \right) \text{Tr} \left(\Gamma_{x_1, x_1+1/2} \bar{A} \Gamma_{x_0, x_0+1/2} B \right) \right. \\
\left. - \text{Tr} \left(\Gamma_{x_1-1/2, x_1} \bar{A} \Gamma_{x_0, x_0+1/2} B \right) \text{Tr} \left(\Gamma_{x_1, x_1+1/2} \bar{A} \Gamma_{x_0-1/2, x_0} B \right) \right]
\end{aligned}$$

i.e. we have an interpretation of the last terms in (158). Putting everything together, we obtain the first quadratic difference equation

$$\begin{aligned}
S [C^2(x, y) - C(x, y+1)C(x, y-1)] \\
+ \tilde{S} [\tilde{C}^2(x, y) - \tilde{C}(x+1, y)\tilde{C}(x-1, y)] = 0
\end{aligned} \quad (159a)$$

The second relation is obtained by interchanging the roles of the coordinates x and y , a discrete symmetry of the lattice (or by

exchanging β and $\bar{\beta}$)

$$\begin{aligned}
S [C^2(x, y) - C(x+1, y)C(x-1, y)] \\
+ \tilde{S} [\tilde{C}^2(x, y) - \tilde{C}(x, y+1)\tilde{C}(x, y-1)] = 0
\end{aligned} \quad (159b)$$

Applying again Wick's theorem to another combination of operators one derives in the same vein a third relation (symmetric in x and y)

$$\begin{aligned}
S [C(x+1, y+1)C(x, y) - C(x+1, y)C(x, y+1)] = \\
\tilde{S} [\tilde{C}(x+1, y+1)\tilde{C}(x, y) - \tilde{C}(x+1, y)\tilde{C}(x, y+1)]
\end{aligned} \quad (159c)$$

These relations exhibit an interesting symmetry (or antisymmetry) under duality. Supplemented by appropriate boundary conditions, they determine fully the two-point function. They also generalize to higher correlations. We now assume that a continuum limit is meaningful in the vicinity of T_c , and we study their continuum approximation. To do so, we first absorb a factor $S^{1/2} \equiv \sinh 2\beta^{1/2}$ into the definition of C

$$S^{1/2} C \rightarrow C \quad (160)$$

expand all terms in equations (159) to second order in the lattice spacing, and use complex coordinates $z = x + iy$, $\bar{z} = x - iy$, to recast the system into the form

$$(C\partial\bar{\partial}C - \partial C\bar{\partial}C) + (\bar{C}\partial\bar{\partial}\bar{C} - \partial\bar{C}\bar{\partial}\bar{C}) = 0 \quad (161a)$$

$$C\partial^2 C - (\partial C)^2 = \bar{C}\partial^2 \bar{C} - (\partial\bar{C})^2 \quad (161b)$$

$$C\bar{\partial}^2 C - (\bar{\partial}C)^2 = \bar{C}\bar{\partial}^2 \bar{C} - (\bar{\partial}\bar{C})^2 \quad (161c)$$

Let us assume C (and $\bar{C}$) isotropic i.e. function only of the distance $|z|$. Then the two equations (161b) and (161c) collapse into a single one. For a while, let us set

$$z\bar{z} = e^{2t} \quad C = e^f \quad \bar{C} = e^{\bar{f}} \quad (162)$$

Then

$$e^{2f} f'' + e^{2\bar{f}} \bar{f}'' = 0 \quad (163a)$$

$$e^{2f} (f'' - 2f') = e^{2\bar{f}} (\bar{f}'' - 2\bar{f}') \quad (163b)$$

where primes denote derivatives with respect to t . These equations have to satisfy a compatibility condition. To obtain this, we

rewrite the system in the form

$$e^{2f} f' = e^{2\tilde{f}} (\tilde{f}' - \tilde{f}'') \quad (164a)$$

$$e^{2\tilde{f}} \tilde{f}' = e^{2f} (f' - f'') \quad (164b)$$

and take a derivative of the second equation with respect to t

$$e^{2\tilde{f}} (\tilde{f}'' + 2\tilde{f}'^2) = e^{2f} (f'' - f''' + 2f'^2 - 2f'f'')$$

we use (163a) to replace $e^{2\tilde{f}} \tilde{f}''$ by $-e^{2f} f''$ and (164b) to replace $2e^{2\tilde{f}} \tilde{f}'^2$ by $2\tilde{f}' e^{2f} (f' - f'')$, thus getting

$$\begin{aligned} f''' - 2f'^2 - 2f'' + 2f'f'' + 2\tilde{f}'(f' - f'') &= 0 \\ \tilde{f}''' - 2\tilde{f}'^2 - 2\tilde{f}'' + 2\tilde{f}'\tilde{f}'' + 2f'(\tilde{f}' - \tilde{f}'') &= 0 \end{aligned}$$

where the second equality follows from duality. Adding the two equations yields

$$\frac{d}{dt} [f'' + f'^2 + \tilde{f}'' + \tilde{f}'^2 - 2f'\tilde{f}'] - 2[f'' + f'^2 + \tilde{f}'' + \tilde{f}'^2 - 2f'\tilde{f}'] = 0$$

Hence by integration

$$f'' + f'^2 + \tilde{f}'' + \tilde{f}'^2 - 2f'\tilde{f}' = \text{cst } e^{2t} = \text{cst } z\bar{z} \quad (165)$$

Returning to the original function, this is seen to be equivalent to

$$\frac{1}{C\tilde{C}} [\tilde{C}\partial\bar{\partial}C - C\partial\bar{\partial}\tilde{C} - \partial C\bar{\partial}\tilde{C} - \bar{\partial}C\partial\tilde{C}] = \text{cst} = \frac{1}{4}m^2 \quad (166)$$

The constant on the right hand side is identified as $m^2/4$ assuming the following behaviour at large separation for $T > T_c$

$$\tilde{C} \rightarrow \text{cst} \quad C \rightarrow \frac{\text{cst}}{|z|^a} \exp -m|z| \quad (167)$$

We have now what might be called the Painlevé system

$$C\partial\bar{\partial}C + \tilde{C}\partial\bar{\partial}\tilde{C} - \partial C\bar{\partial}\tilde{C} - \bar{\partial}C\partial\tilde{C} = 0 \quad (168a)$$

$$\tilde{C}\partial\bar{\partial}C + C\partial\bar{\partial}\tilde{C} - \partial C\bar{\partial}\tilde{C} - \bar{\partial}C\partial\tilde{C} = \frac{1}{4}m^2 C\tilde{C} \quad (168b)$$

The second equation provides a scale in which to measure length. We change variables by defining

$$u = \frac{1}{2}mz \quad (169)$$

and now understand derivatives as carried with respect to u and $\bar{u}$, which amounts to replace $m^2/4$ by unity in (168b). By adding

and subtracting the two equations (168), we obtain the following system

$$C + \tilde{C} = e^L \quad C - \tilde{C} = e^M \quad u = \frac{r}{\alpha} z \quad (170a)$$

$$4\partial\bar{\partial}(L - M) = e^{2(L-M)} - e^{-2(L-M)} \quad (170b)$$

$$4\partial\bar{\partial}(L + M) = 2 - e^{2(L-M)} - e^{-2(L-M)} \quad (170c)$$

Recalling that C and $\tilde{C}$, and therefore L and M , only depend on $|z|$, equation (170b) determines $(L - M)$, while (170c) yields $L + M$ provided we include some information on boundary conditions. The system (170) therefore determines C and $\tilde{C}$. Painlevé's equation proper is (170b), when we insist that $L - M$ depends on $|z|$ only. When this is taken into account, it becomes a second order nonlinear differential equation with the specific property that its solutions have only poles as moveable singularities (i.e. dependent on the boundary conditions). Specifically, if

$$\eta = e^{-(L-M)} = \frac{\tilde{C} - C}{\tilde{C} + C} \quad r = |u| = \frac{1}{2}m|z| \quad (171a)$$

$$\frac{1}{\eta} \frac{d^2\eta}{dr^2} - \left(\frac{1}{\eta} \frac{d\eta}{dr} \right)^2 + \frac{1}{r} \frac{d\eta}{dr} = \eta^2 - \frac{1}{\eta^2} \quad (171b)$$

Formally this is invariant under $\eta \rightarrow 1/\eta$. We can satisfy ourselves with the solution for T slightly above T_c , since determining C and $\tilde{C}$ yields then also C below T_c . Conditions (167) imply that η goes to 1 for large r , up to exponentially small terms. If we neglect $(\eta'/\eta)^2$ as compared to η''/η , this yields

$$\eta = 1 + \text{cst} \frac{e^{-2r}}{r^{1/2}}$$

thus the exponent in (167) is $a = 1/2$, as expected from equation (107). We let the reader convince himself that the leading asymptotic behaviour is

$$|z| \rightarrow \infty \begin{cases} \eta = 1 - \frac{2}{\pi} K_0(2r) + \dots = 1 - \sqrt{\frac{2}{\pi m |z|}} e^{-m|z|} + \dots \\ \tilde{C} \rightarrow \text{cst} \\ C/\tilde{C} \sim \frac{1}{\pi} K_0(2r) \sim \sqrt{\frac{1}{2m |z|}} e^{-m|z|} \end{cases} \quad (172)$$

Close to the origin, $\mathcal{C}$ and $\tilde{\mathcal{C}}$ have the expected $|z|^{-1/4}$ scaling behaviour

$$|z| \rightarrow 0 \begin{cases} \mathcal{C} \sim \frac{1}{(m|z|)^{1/4}} [1 + \frac{1}{2}m|z| \ln \frac{1}{2}m|z| + \dots] \\ \tilde{\mathcal{C}} \sim \frac{1}{(m|z|)^{1/4}} [1 - \frac{1}{2}m|z| \ln \frac{1}{2}m|z| + \dots] \end{cases} \quad (173)$$

Notes

Grassmannian variables and their properties are discussed by F.A. Berezin in *The Method of Second Quantization*, Academic Press, New York (1966). They are now used in many contexts, including the treatment of disordered systems. Examples will be given in chapter 10.

It is not possible to give here a serious and complete bibliography on the Ising model introduced in fact by Lenz, *Z. Physik* **21**, 613 (1920) and first studied by E. Ising, *Z. Physik* **31**, 253 (1925). We list here some of the fundamental works pertaining to the two-dimensional solution, as well as those related to our presentation. A proof of the existence of a low temperature ordered phase, not discussed in the text is the one of R. Peierls, *Proc. Cambridge Phil. Soc.* **32**, 477 (1936). Transfer matrix methods and duality appear in H.A. Kramers and G.H. Wannier, *Phys. Rev.* **60**, 252, 263 (1941). The free energy was computed, using Clifford algebra methods, by L. Onsager, *Phys. Rev.* **65**, 117 (1944). The proof was simplified by B. Kaufman, *Phys. Rev.* **76**, 1232 (1949). Onsager also announced the expression for the spontaneous magnetization in *Nuovo Cimento* **6**, Supplement, 261 (1949). A complete derivation was given by C.N. Yang, *Phys. Rev.* **85**, 808 (1952). M. Kac and J.C. Ward developed a combinatorial proof in *Phys. Rev.* **88**, 1332 (1952). Other reviews, and alternative methods can be traced in G.F. Newell and E.W. Montroll, *Rev. Mod. Phys.* **25**, 353 (1953), E.W. Montroll, R.B. Potts, J.C. Ward, *J. Math. Phys.* **4**, 308 (1963), T.D. Sluutz, D.C. Mattis, E.H. Lieb, *Rev. Mod. Phys.* **36**, 856 (1964), H.S. Green and C.A. Hurst *Order disorder phenomena*, Interscience, London (1964). A discussion on finite size

effects is found in A.E. Ferdinand and M.E. Fisher, *Phys. Rev.* **185**, 832 (1969).

The lemma on Töplitz determinants is due to G. Szegő in *Communications du séminaire mathématique de l'université de Lund*, tome supplémentaire dédié à M. Riesz, 228 (1952) and to M. Kac, *Duke Math. J.* **21**, 501 (1954).

A general summary, including many references and a presentation of their early contributions, is to be found in B.M. McCoy and T.T. Wu, *The Two-Dimensional Ising Model*, Harvard University Press, Cambridge, Mass. (1973).

The concept of order and disorder variables is developed in L.P. Kadanoff, H. Ceva, *Phys. Rev.* **B3**, 3918 (1971).

The construction of the equilibrium shape of crystals from the interfacial tension is due to G. Wulff, *Z. Kristallogr. Mineral.* **34**, 449 (1901). See also C. Herring, *Phys. Rev.* **82**, 87 (1951). The computation for the two-dimensional Ising case is based on the work of H. Cheng and T.T. Wu, *Phys. Rev.* **164**, 719 (1967), P.G. Watson, volume 3 of the Domb and Green series, D.B. Abraham and P. Reed, *J. Phys.* **A10**, L121 (1977) and C. Rottman and M. Wortis, *Phys. Rev.* **B24**, 6274 (1981). The curves in the text are from J.E. Avron, H. van Beijeren, L.S. Schulman and R.K.P. Zia, *J. Phys.* **A15**, L81 (1982). See also R.K.P. Zia and J. Avron, *Phys. Rev.* **B25**, 2041 (1982).

The representation of the Ising model in terms of Grassmannian integrals is patterned after the work of S. Samuel, *J. Math. Phys.* **21**, 2806, 2815, 2820 (1980), and of one of the authors in *Nucl. Phys.* **B210 (FS6)**, 448 (1982).

The calculation of correlation functions in the scaling limit is found in the work of T.T. Wu, B.M. McCoy, C.A. Tracy, E. Barouch, *Phys. Rev.* **B13**, 316 (1976). For a relation with continuous field theory, see A. Luther, I. Peschel, *Phys. Rev.* **B12**, 3908 (1975), the work of one of the authors in collaboration with M. Bander, *Phys. Rev.* **D15**, 463 (1977), and with J.-B. Zuber, *Phys. Rev.* **D15**, 2875 (1977), as well as M. Bander and J.L. Richardson, *Phys. Rev.* **B17**, 1464 (1978).

Quadratic difference equations are discussed in B.M. McCoy and T.T. Wu, *Phys. Rev. Lett.* **45**, 675 (1980), J.H.H. Perk, *Phys. Lett.* **79A**, 1,3 (1980), B.M. McCoy, J.H.H. Perk, and T.T. Wu, *Phys. Rev. Lett.* **46**, 757 (1981).

Finally the work of M. Sato, T. Miwa and M. Jimbo is to be found in a series of publications, *Proc. Jap. Acad.* **53A**, 147, 153, 183 (1977) and *Publications R.I.M.S.*, Kyoto University, **14**, 223 (1978), **15**, 201, 577, 871 (1979).

For the Painlevé property, see E. Ince *Ordinary differential equations*, Dover, New York (1945).

3

SPONTANEOUS SYMMETRY
BREAKING, MEAN FIELD

In many statistical models (but not all) one can introduce a local order parameter associated with a finite, discrete or continuous group of symmetries. The higher the temperature, the more important the fluctuations. One expects therefore in general to find pure phases at low temperature, with a reduced symmetry group, as a consequence of a nonvanishing expectation value of some order parameter. This situation is referred to as spontaneous symmetry breaking. A typical example is the classical Heisenberg model describing short-range interactions of an n -vector field φ , with an orthogonal $O(n)$ symmetry group. For $n = 1, 2, 3$, this can account for a Curie transition from a ferromagnetic to a paramagnetic phase. In particle physics the σ -model of Gell-Mann and Levy involves a spontaneous symmetry breaking of chiral invariance, typical of massless spinor fields, accompanied by soft excitation modes, the so-called Goldstone modes, associated to a π -meson triplet and leading to a nonvanishing dynamical fermion mass. In a first and rather crude approximation, one can analyse the action itself or an effective one incorporating some fluctuation effects, and look for extrema as a function of field configurations, generally translationally invariant. The remaining fluctuations are then treated perturbatively. This mean field method is common to a great variety of domains, ranging from the Clausius-Mossotti formula for a polarizable medium, the Weiss molecular field in the theory of magnetism, Landau's effective action in various statistical contexts, the effective medium approximation in disordered systems, the Hartree-Fock method in atomic or many-body physics, to the semiclassical approximation in the study of quantum systems. A common mathematical tool is, at one stage or another, the steepest descent or saddle point evaluation of a functional integral, implying a stability analysis and a perturbative expansion of the fluctuation effects.

Finally the work of M. Sato, T. Miwa and M. Jimbo is to be found in a series of publications, *Proc. Jap. Acad.* **53A**, 147, 153, 183 (1977) and *Publications R.I.M.S.*, Kyoto University, **14**, 223 (1978), **15**, 201, 577, 871 (1979).

For the Painlevé property, see E. Ince *Ordinary differential equations*, Dover, New York (1945).

3

SPONTANEOUS SYMMETRY
BREAKING, MEAN FIELD

In many statistical models (but not all) one can introduce a local order parameter associated with a finite, discrete or continuous group of symmetries. The higher the temperature, the more important the fluctuations. One expects therefore in general to find pure phases at low temperature, with a reduced symmetry group, as a consequence of a nonvanishing expectation value of some order parameter. This situation is referred to as spontaneous symmetry breaking. A typical example is the classical Heisenberg model describing short-range interactions of an n -vector field φ , with an orthogonal $O(n)$ symmetry group. For $n = 1, 2, 3$, this can account for a Curie transition from a ferromagnetic to a paramagnetic phase. In particle physics the σ -model of Gell-Mann and Levy involves a spontaneous symmetry breaking of chiral invariance, typical of massless spinor fields, accompanied by soft excitation modes, the so-called Goldstone modes, associated to a π -meson triplet and leading to a nonvanishing dynamical fermion mass. In a first and rather crude approximation, one can analyse the action itself or an effective one incorporating some fluctuation effects, and look for extrema as a function of field configurations, generally translationally invariant. The remaining fluctuations are then treated perturbatively. This mean field method is common to a great variety of domains, ranging from the Clausius-Mossotti formula for a polarizable medium, the Weiss molecular field in the theory of magnetism, Landau's effective action in various statistical contexts, the effective medium approximation in disordered systems, the Hartree-Fock method in atomic or many-body physics, to the semiclassical approximation in the study of quantum systems. A common mathematical tool is, at one stage or another, the steepest descent or saddle point evaluation of a functional integral, implying a stability analysis and a perturbative expansion of the fluctuation effects.

The physical motivation is to substitute for the dynamics or statistics of a very large number of coupled degrees of freedom, an effective one, pertaining to a small number of carefully chosen relevant observables, interacting via a self-consistent mean field, representing the overall effect of the remaining ones. The examples will show how these ideas are implemented in practice.

3.1 Mean field approximation

3.1.1 Dielectric constant of a polarizable medium

A classical instance of mean field theory is the derivation of an effective dielectric constant of a polarizable medium. The introduction of a test charge in such an environment induces a charge density $\rho_{\text{ind}}(\mathbf{x})$, with a vanishing integral (since the medium is charge-neutral). The latter can therefore be written as the divergence of a dipole density

$$\rho_{\text{ind}}(\mathbf{x}) = -\nabla \cdot \mathbf{P}(\mathbf{x}) \quad (1)$$

Multiply both sides by $\mathbf{x}$ and integrate over space. If $\rho_{\text{ind}}(\mathbf{x})$ and $\mathbf{P}(\mathbf{x})$ are sufficiently localized so that an integration by parts is justified, we find for the total induced dipole moment

$$\mathbf{P}_{\text{tot}} = \int d^3\mathbf{x} \, \mathbf{x} \rho_{\text{ind}}(\mathbf{x}) = \int d^3\mathbf{x} \, \mathbf{P}(\mathbf{x}) \quad (2)$$

A shift $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{x}_0$ does not affect this result since the total induced charge vanishes. As a consequence, Gauss's law $\text{div} \mathbf{E} = \rho/\epsilon_0$, giving the electric field produced by the total charge density $\rho = \rho_{\text{ext}} + \rho_{\text{ind}}$, can be reshuffled to read

$$\text{div} \left(\mathbf{E} + \frac{\mathbf{P}}{\epsilon_0} \right) \equiv \text{div} \mathbf{D} = \frac{\rho_{\text{ext}}}{\epsilon_0} \quad (3)$$

For sufficiently small fields and an isotropic medium, the induced moment is proportional to the electric field. The electric induction can thus be written as

$$\mathbf{D} = \mathbf{E} + \frac{\mathbf{P}}{\epsilon_0} = \epsilon \mathbf{E} \quad (4)$$

where ϵ is the effective dielectric constant. Let us now look at this expression from a microscopic point of view. The dipole moment $\mathbf{P}$ can be written as $n\mathbf{p}$ where $\mathbf{p}$ denotes the individual molecular

dipole moment and n is the density of carriers. In turn each molecular dipole $\mathbf{p}$ is proportional to a local electric field $\mathbf{E}_{\text{loc}}$, with $\mathbf{p} = \alpha \mathbf{E}_{\text{loc}}$, and α the polarizability of the isolated molecule. The point is now to find the effective local field. Assuming weak fields as in (3), the standard analysis decomposes $\mathbf{E}_{\text{loc}}$ into three parts. The first one is the electric field $\mathbf{E}$ due to external sources, the second $\mathbf{E}_1$ being the result of far away dipoles outside a volume, large on the microscopic scale but small on a macroscopic one (a classical, poorly defined, but convenient concept). This volume is chosen spherical to respect isotropy. Finally $\mathbf{E}_2$ will be the field produced by the nearby molecules in this spherical cavity. The latter vanishes in a medium with a large symmetry (a disordered one, as a liquid, or even in a crystalline cubic environment for instance). The problem of evaluating $\mathbf{E}_1$ at the center of the cavity is of course a well-known elementary exercise. It is obtained as the one deriving from a surface density $\sigma = -\mathbf{n} \cdot \mathbf{P}$, $\mathbf{n}$ being the unit vector along the normal to the sphere pointing outwards. Choosing the z -axis along $\mathbf{P}$, R being the radius of the sphere, the corresponding potential along the z axis is

$$\phi(z) = -P \frac{R^2 2\pi}{4\pi\epsilon_0} \int_0^\pi d\cos\theta \frac{\cos\theta}{\sqrt{R^2 + z^2 - 2Rz\cos\theta}} = -z \frac{P}{3\epsilon_0}$$

The field $\mathbf{E}_1 = -\nabla \phi$ at the center is

$$\mathbf{E}_1 = \frac{\mathbf{P}}{3\epsilon_0}$$

The local field is the sum $\mathbf{E} + \mathbf{E}_1 = \mathbf{E} + \mathbf{P}/3\epsilon_0$, and the individual polarization is

$$\mathbf{p} = \alpha \mathbf{E}_{\text{loc}} = \alpha \left(\mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \right)$$

Multiplying both sides by the density n

$$\mathbf{P} = n\alpha \left(\mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \right)$$

From this and the definition (4) it follows that

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{1}{3} \frac{n\alpha}{\epsilon_0} \quad (5)$$

which is the Clausius-Mossotti relation (or the Lorentz-Lorentz relation in an optical context, with ϵ replaced by the square of

the refraction index) expressing the phenomenological dielectric constant in terms of the microscopic polarizability α . Note the self-consistent character of the computation of the local field.

The mean field approximation (5) does not take into account fluctuations and correlations. It is interesting to see how these come about using the standard Kirkwood-Yvon theory. Consider the case of a uniform nonpolar medium of mean density n , from a microscopic point of view. Let $\mathbf{x}_i$ denote the positions of the molecules or atoms, and $\mathbf{E}$ stand as before for the external field. The induced dipole $\mathbf{p}_i$ on molecule i is given by

$$\mathbf{p}_i = \alpha \left[\mathbf{E} - \sum_{j, j \neq i} K_{ij} \mathbf{p}_j \right]$$

with K_{ij} the matrix

$$K_{ij} = \frac{1}{4\pi\epsilon_0} \nabla_i \otimes \nabla_j \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|}$$

Set $\langle \mathbf{p} \rangle = \langle \mathbf{p}_i \rangle$ independent of i , where the bracket is the statistical average. We have

$$\mathbf{p}_i = \alpha \left[\mathbf{E} - \sum_{j, j \neq i} \langle K_{ij} \rangle \langle \mathbf{p}_j \rangle - \sum_{j, j \neq i} (K_{ij} \mathbf{p}_j - \langle K_{ij} \rangle \langle \mathbf{p}_j \rangle) \right]$$

A calculation similar to the one above yields

$$\mathbf{p}_0 = \alpha \left[\mathbf{E} - \sum_{j, j \neq i} \langle K_{ij} \rangle \langle \mathbf{p} \rangle \right] = \alpha \left[\mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0} \right] = \frac{\alpha}{3} (\epsilon + 2) \mathbf{E}$$

If one neglects the fluctuating contribution to $\mathbf{p}_i$, thus identifying $n\mathbf{p}_0$ with $\mathbf{P}$, one recovers the mean field result (5). To obtain the correction, we have only to iterate

$$\mathbf{p}_i = \mathbf{p}_0 - \alpha \sum_{j, j \neq i} [K_{ij} \mathbf{p}_j - \langle K_{ij} \rangle \langle \mathbf{p}_j \rangle]$$

by considering α formally as a small parameter. Let us content ourselves with a first iteration and then take the average. This yields

$$\langle \mathbf{p} \rangle = \mathbf{p}_0 + \alpha^2 \sum_{\substack{j, k \\ j \neq i, k \neq j}} [\langle K_{ij} K_{jk} \rangle - \langle K_{ij} \rangle \langle K_{jk} \rangle] \mathbf{p}_0 + \dots$$

Since $\mathbf{p}_0 = \frac{1}{3}\alpha(\epsilon + 2)\mathbf{E}$, it is clear that $\mathbf{p}$ is along $\hat{\mathbf{E}}$, the unit vector in the direction of the external field. Multiplying both sides by n/ϵ_0 , we readily find

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{n\alpha}{3\epsilon_0} [1 + S] \quad (6a)$$

$$S = \alpha^2 \sum_{\substack{j, k \\ j \neq i, k \neq j}} \hat{\mathbf{E}} [\langle K_{ij} K_{jk} \rangle - \langle K_{ij} \rangle \langle K_{jk} \rangle] \hat{\mathbf{E}} + \mathcal{O}(\alpha^3)$$

We split the above sum into a contribution with $j \neq i \neq k$ and another one with $k = i$. If N and V stand respectively for the total number of molecules and the total volume, both tending to infinity, we define the molecular distribution functions as

$$n_2(\mathbf{x}_1 - \mathbf{x}_2) = \sum'_{i, j} \langle \delta(\mathbf{x}_1 - \mathbf{x}_i) \delta(\mathbf{x}_2 - \mathbf{x}_j) \rangle = n^2 g_2(|\mathbf{x}_1 - \mathbf{x}_2|)$$

where the prime means the omission of the contribution $i = j$, so that for instance

$$\begin{aligned} \sum'_{i, j \neq i} \langle f(\mathbf{x}_i - \mathbf{x}_j) \rangle &= \frac{V}{N} \frac{1}{V} \int d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 f(\mathbf{x}_1 - \mathbf{x}_2) \\ &\quad \times \sum'_{i, j} \langle \delta(\mathbf{x}_1 - \mathbf{x}_i) \delta(\mathbf{x}_2 - \mathbf{x}_j) \rangle \\ &= \frac{1}{n} \int d^3 \mathbf{x} n_2(\mathbf{x}) f(\mathbf{x}) = n \int d^3 \mathbf{x} g_2(\mathbf{x}) f(\mathbf{x}) \end{aligned}$$

Inserting such expressions in the correction term S and performing the sums and angular integrals gives

$$\begin{aligned} S &= \frac{\alpha^2}{(4\pi\epsilon_0)^2} \left\{ 8\pi n \int_0^\infty dx \frac{g_2(x)}{x^4} \right. \\ &\quad \left. + n^2 \int d^3 \mathbf{x}_1 d^3 \mathbf{x}_2 \frac{g_3(\mathbf{x}_1, \mathbf{x}_2) - g_2(x_1)g_2(x_2)}{x_1^3 x_2^3} [3\hat{\mathbf{x}}_1 \cdot \hat{\mathbf{x}}_2 - 1] \right\} \\ &\quad + \mathcal{O}(\alpha^3) \end{aligned} \quad (6b)$$

predicting to leading order a parabolic shape for the correction as a function of density or pressure. Long-range correlations close to a continuous transition invalidate the preceding perturbative approach and lead to critical opalescence.

3.1.2 Classical spin model with a finite symmetry group

Consider an Ising model on a hypercubical d -dimensional lattice where the coordination number $q = 2d$ grows linearly with d . The Boltzmann weight of a configuration in an external field h is

$$p(\sigma) = Z^{-1} \exp \left(\beta \sum_{(ij)} \sigma_i \sigma_j + h \sum_i \sigma_i \right) \quad (7)$$

Strictly, the true external field is proportional to h/β ; we keep however the above terminology. For $h = 0$, the weights of the configurations $\{\sigma_i\}$ and $\{-\sigma_i\}$ are identical, exhibiting $Z/2Z$, or for short Z_2 , the group with two elements ± 1 , as the symmetry group of the model. We saw in chapter 2 that at least for $d = 2$, the Ising model has a low temperature ordered phase, with pure states characterized by a spontaneous magnetization breaking the Z_2 symmetry. Let us carry over the discussion to higher dimension in the mean field context. The probability weight (7) suggests that we consider the random variable

$$H_i = \beta \sum_{j(i)} \sigma_j \quad (8)$$

as a field acting on the spin σ_i , very much as we did with the other dipoles which contributed to the local electric field in the previous example. The difference is here that interactions are of short range. As d increases, so does the number of neighbours contributing in (8). It is therefore tempting for large d to neglect the fluctuations of H_i , and to replace it by its mean value, borrowing the idea from the central limit theorem. If a nonvanishing mean value of H_i persists in the limit $h \rightarrow 0$, there is a spontaneous symmetry breaking. This mean field has to be computed self-consistently. Let us present several derivations to illustrate the point.

When computing the partition function corresponding to (7), which normalizes the weights, it is possible to reverse any set of (dummy) variables σ , since they are summed over. For instance, Z is invariant if we change the sign of a single spin σ_i , before summation. Taking into account the definition of H_i , this leads to the identity

$$1 = \langle \exp(-2\sigma_i(H_i + h)) \rangle \quad (9a)$$

or equivalently

$$\langle \cosh 2(H_i + h) \rangle = 1 + \langle \sigma_i \sinh 2(H_i + h) \rangle \quad (9b)$$

The approximation is to replace the random variable H_i by its translational invariant mean H , while $\langle \sigma_i \rangle$ can be identified with the magnetization M . From (9), we obtain

$$M = \frac{\cosh 2(H + h) - 1}{\sinh 2(H + h)} = \tanh(H + h) \quad (10a)$$

and for consistency with (8), we must have

$$H = 2\beta dM \quad (10b)$$

What is remarkable is that, even for zero external field h , these equations can have a nonvanishing solution. When $h = 0$, the Z_2 symmetry implies that for every solution (H, M) we have the opposite one $(-H, -M)$ and H and M have obviously the same sign. Therefore we can assume $H > 0$, which has then to satisfy

$$\frac{H}{2d\beta} = \tanh H \quad (10c)$$

In general, $2d$ should be replaced by the coordination number of the lattice, assumed to be large. If $2d\beta > 1$, i.e. at sufficiently small temperature, equation (10c) admits a nonzero solution, such that $M = H/2d\beta$ decreases from 1 at $\beta \rightarrow \infty$, down to zero at $\beta_c = (2d)^{-1}$. The graphical solution is shown in figure 1. At high temperature, $\beta < \beta_c$, the spontaneous magnetization M vanishes with the local field H . One expects therefore a disordered symmetric phase. Observe that as the disorder increases, so does the symmetry. In a similar sense, the crystalline patterns observed after solidification express a loss of (translational or rotational) symmetry as compared to the liquid or gaseous phase.

Close to the transition point, one finds a square root behaviour typical of the present approximation

$$M \sim H \sim \sqrt{3}(\beta/\beta_c - 1)^{1/2} \quad (11)$$

This critical exponent is denoted β , following the convention introduced in the previous chapter, where we saw that its value for $d = 2$ is $1/8$, rather than $1/2$ as predicted by the mean field approximation.

For $h \neq 0$, equations (10a) and (10b) admit either a unique solution or several ones, which have to be distinguished on the

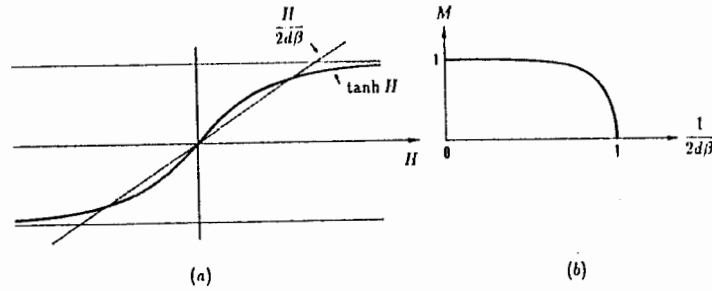


Fig. 1 (a) Graphical solution of the mean field equation. (b) Spontaneous magnetization.

basis of the maximization of $(\ln Z)/\Omega$, where Ω is the total volume. The interesting situation is the one near the critical point. At $\beta = \beta_c$, by expanding the hyperbolic tangent, we find $h \sim (M + h)^3/3$, thus

$$\beta = \beta_c \quad M \sim (3h)^{1/3} \quad (12a)$$

or

$$\beta = \beta_c \quad h \sim M^3, \quad \delta_{\text{mean field}} = 3 \quad (12b)$$

Approaching criticality in the high temperature phase, M is linear in h , but as suggested by (12a) the corresponding coefficient, the susceptibility must diverge as $\beta \rightarrow \beta_c - 0$. Indeed we find

$$\beta \rightarrow \beta_c - 0 \quad M \sim \chi h \quad \chi \sim \frac{1}{1 - \beta/\beta_c} \quad (13a)$$

or

$$\chi \sim (1 - \beta/\beta_c)^{-\gamma} \quad \gamma_{\text{mean field}} = 1 \quad (13b)$$

Recall that $\gamma = 7/4$ in dimension two.

We return to the derivation of the mean field equations from another point of view, namely the minimization of the free energy $-\ln Z/\beta\Omega$. The latter is related to the convexity property of the exponential function, i.e.

$$\langle \exp A \rangle \geq \exp \langle A \rangle \quad (14)$$

Let us apply this inequality by taking as a trial Boltzmann weight a factorized expression, the only one for which we can do an explicit computation. Alas, one has frequently no other choice but to search for one's keys under the street lamp! For N sites (the volume Ω being Na^d , with a the lattice spacing), write

$$Z = \frac{1}{2^N} \sum_{\sigma_i = \pm 1} \exp \left(\beta \sum_{(ij)} \sigma_i \sigma_j + h \sum_i \sigma_i \right) = Z_H \left\{ Z_H^{-1} \frac{1}{2^N} \sum_{\sigma_i = \pm 1} \exp \left((H + h) \sum_i \sigma_i + A \right) \right\} \quad (15a)$$

with

$$Z_H = \frac{1}{2^N} \sum_{\sigma_i = \pm 1} \exp \left[(H + h) \sum_i \sigma_i \right] = [\cosh(H + h)]^N \quad (15b)$$

and

$$A = \beta \sum_{(ij)} \sigma_i \sigma_j - H \sum_i \sigma_i \quad (15c)$$

In other words, we have added and subtracted $H \sum_i \sigma_i$ in the action (or energy), and presented Z as the product of Z_H times a mean value taken on a factorized measure. The latter cannot be evaluated exactly, but we can apply the inequality (14), with the result that

$$F = \frac{1}{N} \ln Z \geq \sup_H \left\{ \ln \cosh(h + H) + \frac{1}{N} \langle A \rangle_H \right\} \quad (16)$$

We have taken advantage of the arbitrariness on H to majorize the right-hand side over H , and of course $\langle \dots \rangle_H$ means the average over the factorized measure. In the thermodynamic limit

$$M = \langle \sigma_i \rangle_H = \tanh(H + h) \quad (17)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \langle A \rangle_H = \beta d \tanh^2(H + h) - H \tanh(H + h)$$

Thus

$$F \geq \sup_H \mathcal{F}(H, h; \beta)$$

$$\mathcal{F}(H, h; \beta) = \ln \cosh(h + H) + \beta d \tanh^2(H + h) - H \tanh(H + h) \quad (18)$$

Identifying $\tanh(H + h)$ with the magnetization M , the interpretation of (18) is straightforward. The piece

$$\beta d \tanh^2(H + h) + h \tanh(H + h) = \beta d M^2 + h M$$

is the energy contribution, while the remaining part

$$\begin{aligned} \ln \cosh(H + h) - (H + h) \tanh(H + h) = \\ = -\ln 2 - \left[\frac{1}{2}(1 + M) \ln \frac{1}{2}(1 + M) + \frac{1}{2}(1 - M) \ln \frac{1}{2}(1 - M) \right] \end{aligned}$$

represents entropy. The $-\ln 2$ contribution arises from our normalization of Z . Thus

$$\begin{aligned} \mathcal{F}(H, h; \beta) &= -V(M, \beta) + h M \\ V(M, \beta) &= -\beta d M^2 + \frac{1}{2}(1 + M) \ln \frac{1}{2}(1 + M) \\ &\quad + \frac{1}{2}(1 - M) \ln \frac{1}{2}(1 - M) + \ln 2 \end{aligned} \quad (19)$$

and we have to minimize $V(M, \beta) - h M$. The relation between the expressions (18) and (19) is known as a Legendre transformation. In either form, the condition for an extremum is the previous mean field equation

$$H = 2\beta d \tanh(H + h) \quad (20)$$

The advantage of the present derivation is that it is based on a variational principle, thus allowing us to distinguish the various solutions. Figure 2 shows that for $h = 0$, we have a minimum for $V(M, \beta)$ at $M = 0$ for $\beta < \beta_c = 1/2d$, resulting from a dominance of entropy over energy, while in the low temperature regime ($\beta > \beta_c$), two new equivalent minima emerge. An external field h , no matter how small, distorts the picture and leads to a unique absolute minimum with M of the same sign as h . For $\beta = \beta_c$, the curve $V(M)$ has a very flat minimum at the origin with a vanishing coefficient of the quadratic term. In the vicinity of the critical point, the expansion of $V(M, \beta)$ for small M reads

$$V(M, \beta) = \frac{1}{2} \left(1 - \frac{\beta}{\beta_c} \right) M^2 + \frac{1}{12} M^4 + \dots \quad (21)$$

The quartic term always has a positive coefficient, while the coefficient of the quadratic term vanishes linearly (and therefore changes sign). For $h = 0$, the extremum value of V is therefore 0 for $\beta < \beta_c$, and nonvanishing for $\beta > \beta_c$. In the immediate

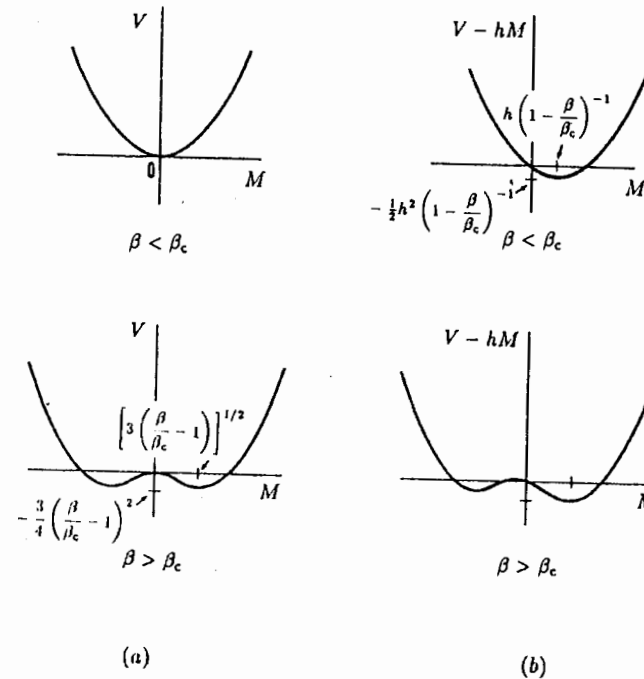


Fig. 2 Minimization of the trial free energy (a) in the absence of an external field and (b) in the presence of such a field.

vicinity, we have

$$\beta > \beta_c \quad V_{\min} = -\frac{3}{4} \left(\frac{\beta}{\beta_c} - 1 \right)^2; \quad \beta < \beta_c \quad V_{\min} = 0 \quad (22)$$

Thus the specific heat c , the second derivative of $V_{\min}$ with respect to β , is discontinuous at β_c . If one insists on characterizing this behaviour as a power law in $|1 - \beta/\beta_c|^{-\alpha}$, with an exponent α , the only choice is to say that $\alpha = 0$. This same α was also zero for $d = 2$, but then the behaviour was logarithmic.

We have applied the mean field ideas in the case of the Ising model with a Z_2 symmetry group. The same reasoning generalizes when the symmetry group is finite. In general, such a group possesses nontrivial subgroups. According to the range

of parameters involved in the effective action, the nature of the various ordered phases will depend on the residual symmetry group, as well as the transformation properties of the order parameter (there may be several ones). We shall not pursue this direction, except to say that the corresponding discussion generally requires some physical insight.

3.1.3 Continuous symmetry group

The spontaneous breaking of a continuous symmetry group leads to an interesting phenomenon in the low temperature phase, namely the existence of soft modes (Goldstone 1961). We shall investigate the simple case of an $O(n)$ -invariant model, using as a dynamical variable an n -dimensional "vector" field φ with fixed length – say unity: $\varphi^2 = 1$. There exist many other possibilities – including variables transforming according to unitary groups, to be discussed at later stages – as well as tensor fields of various kinds.

We have already encountered the $O(n)$ vector model in chapter 1, where we introduced in (1.125) the integral ($H \equiv |\mathbf{H}|$)

$$\begin{aligned} \exp n u_n(H) &= \int \frac{d^n \varphi}{S_n} 2\delta(\varphi^2 - 1) \exp n \mathbf{H} \cdot \varphi \\ &= \Gamma\left(\frac{n}{2}\right) \left(\frac{2}{nH}\right)^{n/2-1} I_{n/2-1}(nH) \\ &= \sum_{k=0}^{\infty} \left(\frac{1}{4} n^2 H^2\right)^k \frac{1}{k!} \frac{\Gamma(\frac{1}{2}n)}{\Gamma(\frac{1}{2}n + k)} \end{aligned} \quad (23)$$

This is a generalization of $N^{-1} \ln Z_H$ appearing in equations (15). For $n = 1$, $u_1(H) = \ln \cosh H$. In general, for H small,

$$u_n(H) = \frac{1}{2} H^2 - \frac{n}{4(n+2)} H^4 + \dots \quad (24)$$

We write the partition function as

$$Z = \int \prod_i \left\{ \frac{d^n \varphi_i}{S_n} 2\delta(\varphi_i^2 - 1) \right\} \exp n \left\{ \beta \sum_{(i,j)} \varphi_i \cdot \varphi_j + \sum_i \mathbf{h} \cdot \varphi_i \right\} \quad (25)$$

a direct generalization of the Ising case (corresponding to $n = 1$). The mean field approximation in variational form reads

$$F \geq \sup_{\mathbf{H}} n \mathcal{F}(\mathbf{H}, \mathbf{h}; \beta) \quad (26)$$

$$\mathcal{F}(\mathbf{H}, \mathbf{h}; \beta) = u(\mathbf{H} + \mathbf{h}) + \beta d [\nabla u(\mathbf{H} + \mathbf{h})]^2 - \mathbf{H} \cdot \nabla u(\mathbf{H} + \mathbf{h})$$

Here the gradient operator ∇ refers to the argument of the function u and not of course to a spatial dependence. Both the external field $\mathbf{h}$, and the mean field $\mathbf{H}$, are assumed spatially homogeneous, as is the action in (25). It would seem at first that the maximization of $\mathcal{F}$ leads to results quite similar to those obtained in the case of a finite symmetry group. Set

$$\mathbf{M} = \nabla u(\mathbf{H} + \mathbf{h}) = \langle \varphi \rangle \quad (27)$$

then

$$\begin{aligned} \mathcal{F}(\mathbf{H}, \mathbf{h}; \beta) &= -V(\mathbf{M}) + \mathbf{h} \cdot \mathbf{M} \\ V(\mathbf{M}) &= \frac{1}{2} (1 - 2d\beta) \mathbf{M}^2 + \frac{1}{4} \frac{n}{n+2} (\mathbf{M}^2)^2 + \dots \end{aligned} \quad (28)$$

With our choice of normalization, the critical value of β remains $\beta_c = 1/2d$, as it was in the case $n = 1$. A spontaneous magnetization appears for $\beta > \beta_c$, and the critical exponents are the same as those found before. In particular the specific heat is discontinuous at the transition, within this approximation.

There exists, however, a crucial difference from the finite case. In the ordered phase with $\mathbf{M} \neq 0$, there remains a continuous invariance group, $O(n-1)$, the little group which leaves the order parameter $\mathbf{M}$ invariant. As a consequence, the pure states are continuously degenerate, in other words pure phases are described by a point on a continuous manifold $O(n)/O(n-1)$, i.e. the unit sphere in n -dimensional space. This leads us to the consideration of the response to external fields with a slow spatial dependence. Let $\mathbf{h} = \mathbf{h}_0 + \mathbf{h}_1 \cos \mathbf{k} \cdot \mathbf{x}_i$, with $\mathbf{h}_0$ and $\mathbf{h}_1$ both very small, and assume $\beta > \beta_c$. The constant (or average) part $\mathbf{h}_0$ defines the mean direction of the magnetization. The susceptibility tensor, i.e. the coefficient of the additional piece of the magnetization proportional to $\mathbf{h}_1$, will have quite a different behaviour according to whether $\mathbf{h}_1$ is parallel or orthogonal to $\mathbf{h}_0$. If $\mathbf{h}_1$ is in these $n-1$ transverse directions, the corresponding variation in free energy vanishes in the limit of small wave numbers, $\mathbf{k} \rightarrow 0$, as it amounts simply to rotating in large spatial regions the global direction of

the magnetization within the manifold of degenerate states. This is the Goldstone phenomenon, corresponding to the existence of $n - 1$ soft excitation modes in the ordered phase. Their possible entropic contribution will have to be considered at a later stage, since they will tend to provide a destabilizing mechanism. In the original σ -model the corresponding group is $O(4)$ (isomorphic to $SU(2) \times SU(2)/Z_2$, a direct product of two unitary unimodular two-dimensional groups, up to a global change in sign of both elements) and the three independent soft modes are associated to the three π -mesons. To be explicit we return to the convexity inequality in the presence of such an external field. Generalizing equations (26)–(27) and (28), we can write

$$\frac{\ln Z}{n} \geq - \inf_{\{M_i\}} \left[\sum_i V(M_i) - \frac{\beta}{2} \sum_i M_i \cdot \sum_{j(i)} (M_j - M_i) - \sum_i h_i \cdot M_i \right] \quad (29)$$

The square of the spontaneous magnetization M_s , in the absence of h , and for β close to β_c , is approximately given by

$$((n+2)/n)(\beta/\beta_c - 1),$$

so that

$$V(M) = \frac{1}{4} \frac{n}{n+2} \left[(M^2 - M_s^2)^2 - (M_s^2)^2 \right] + \dots \quad (30)$$

$$M_s^2 = \frac{n+2}{n} (\beta/\beta_c - 1)$$

The spontaneous magnetization M_s is along h_0 . For h_1 infinitesimal, we write

$$M_i = M_s + \delta M_i \quad (31)$$

To first order in δM_i and h_1 , the above minimization condition yields

$$\frac{2n}{n+2} M_s (M_s \cdot \delta M_i) - \beta \sum_{j(i)} (\delta M_j - \delta M_i) = h_1 \cos k \cdot x_i \quad (32)$$

We look for a plane wave solution

$$\delta M_i = \text{Re} A \exp i k \cdot x_i \quad (33)$$

such that

$$\frac{2n}{n+2} M_s (M_s \cdot A) + 2\beta \sum_{\mu=1}^d (1 - \cos a k_\mu) A = h_1 \quad (34a)$$

Assuming $a k_\mu$ small (a is the lattice spacing), in the long wavelength limit, we can rewrite this amplitude equation as

$$\frac{2n}{n+2} M_s (M_s \cdot A) + \beta a^2 k^2 A = h_1 \quad (34b)$$

We now split h_1 and A into longitudinal and transverse components with respect to M_s , to find

$$A^\perp = \frac{h_1^\perp}{\beta a^2 k^2} \quad A^\parallel = \frac{h_1^\parallel}{\beta a^2 \left(k^2 + \frac{2n}{n+2} \frac{M_s^2}{\beta a^2} \right)} \quad (35)$$

where β can safely be replaced by β_c . In the limit $k = 0$, we observe a typical infrared divergence in the transverse response (or susceptibility) corresponding to massless modes, while the longitudinal amplitude remains finite, the corresponding mass being proportional to the spontaneous magnetization. Of course the number of soft modes is equal to the number of transverse components, i.e. $n - 1$. They are sometimes referred to as spin waves.

The conclusion of this analysis is that a low temperature phase in a system with a spontaneously broken continuous symmetry exhibits massless excitation modes. Their effective dynamics has to be carefully investigated. Potential infrared divergences become more and more severe as the dimension decreases. As a result, we shall see that would-be Goldstone modes are responsible for the absence of ordered phases in dimension two in the case of a noncommutative continuous symmetry group, where they play a pre-eminent role.

As schematic as it is, the mean field approximation reveals, using unsophisticated mathematical tools, a wealth of interesting behaviour. We shall study corrections at the end of this chapter. For the time being, let us look at an alternative point of view particularly useful in the case of discrete variables.

3.1.4 The Bethe approximation

We deal here with a variant of mean field theory, which will be exemplified again on the Ising case. Its main virtue is not that it allows to improve the computation of critical exponents, but that it offers a rather accurate means to locate the critical singularities.

The idea is to think of the system as a superposition of small clusters submitted to an effective field, while imposing a consistency condition by requiring the equality of magnetization, computed in alternative ways. To start with, let us imagine all two neighbouring spin clusters, their number being Nd , if N is the number of lattice sites, acted upon by an effective field H_2 which incorporates both the external field h , as well as the effect of the other interactions. If $\tau = \tanh \beta = e^{-2\beta}$, the corresponding two site partition function is proportional to

$$z_2 = 1 + 2\tau\rho_2 + \rho_2^2 \quad \rho_2 = e^{-2H_2} \quad (36a)$$

Up to a prefactor $[\frac{1}{2} \exp(\beta d + h)]^N$, we could think of replacing the partition function by $(z_2)^{Nd}$, but this would count each spin $2d$ times instead of one. We therefore correct for this overshooting effect by dividing by $z_1^{N(2d-1)}$, where z_1 is a one site partition function of a spin subject to another effective field H_1

$$z_1 = 1 + \rho_1 \quad \rho_1 = e^{-2H_1} \quad (36b)$$

The corresponding approximation is therefore

$$Z = [\frac{1}{2} \exp(\beta d + h)]^N \frac{z_2^{Nd}}{z_1^{N(2d-1)}} \equiv z_B^N \quad (37)$$

We have two parameters at our disposal, H_1 and H_2 . To determine them, we impose two conditions. Those are simply the equality of the magnetization per site in the complete system and in each of its subsystems. Taking into account the fact that z_2 includes two sites and z_1 only one, we find

$$\begin{aligned} M &= \frac{\partial}{\partial h} \ln z_B(h, \beta) \\ &= 1 - \rho_2 \frac{\partial}{\partial \rho_2} \ln z_2(\rho_2, \tau) = 1 - 2\rho_1 \frac{\partial}{\partial \rho_1} \ln z_1(\rho_1) \end{aligned} \quad (38a)$$

Thus from the last equation

$$z_2 = (1 + \rho_2 \tau) z_1 \quad \text{i.e.} \quad \frac{\rho_1}{\rho_2} = \frac{\tau + \rho_2}{1 + \tau \rho_2} \quad (38b)$$

while the first equation furnishes, through (37) and assuming ρ_2 and ρ_1 functions of h and τ , the relation

$$\rho_2 = e^{-2h} \left(\frac{\tau + \rho_2}{1 + \tau \rho_2} \right)^{2d-1} \quad (38c)$$

Table 1. Values of $\tanh \beta_c$ for the Ising model as a function of the dimension d , on a cubic lattice.

dimension	Bethe approximation	exact value
1	1	1
2	$1/3 = 0.33333\dots$	$\sqrt{2} - 1 = 0.41421\dots$
3	$1/5 = 0.2$	$0.218096(3)$
4	$1/7 = 0.14286\dots$	
$d \rightarrow \infty$	$1/(2d-1)$	$1/(2d-1)$

Therefore

$$z_B = \frac{1}{2} e^{h+d\beta} (1 + \rho_2 \tau)^{2d-1} (1 + 2\rho_2 \tau + \rho_2^2)^{1-d} \quad (39)$$

Here ρ_2 is implicitly determined in terms of h and $\tau = e^{-2\beta}$ through equation (38c). When trying to solve for ρ_2 , we find the critical singularity. To see this, use rather than ρ_2 the combination

$$p = \frac{\rho_2 + \tau}{1 + \rho_2 \tau} \quad (40)$$

related to β and h by

$$\tau = e^{-2\beta} = \frac{p^{1-d} e^h - p^{d-1} e^{-h}}{p^{-d} e^h - p^d e^{-h}} \quad (41)$$

exhibiting the symmetry $h \leftrightarrow -h, p \leftrightarrow p^{-1}$. If the external field vanishes, $h = 0$, the branch $p(\beta) \sim e^{-2\beta}$ as $\beta \rightarrow \infty$ will become multivalued for a critical value β_c , given by

$$e^{-2\beta_c} = 1 - d^{-1}, \quad 2d\beta_c = 1 + \frac{1}{2d} + \frac{1}{3d^2} + \frac{1}{4d^3} + \dots \quad (42)$$

If we were to trust this approximation, it would imply a transition for any dimension $d > 1$, with $\beta_c(d) \rightarrow \infty$ as $d \rightarrow 1$. However, we expect that the scheme makes sense only for $d \rightarrow \infty$. We then observe an agreement with the mean field critical value, with an exact $1/2d$ correction to $2d\beta_c$. In the following table we compare the predictions of the Bethe approximation for $\tanh \beta_c$ with the exact value (in three dimensions the value is only known within a small error on the sixth digit).

The prediction for $\tanh \beta_c$ is remarkably accurate and could be improved using larger clusters. On any other lattice, one would have to replace $2d$ by the coordination number q .

For $\beta < \beta_c$, $h = 0$, the appropriate solution of the equation (38) is $p = \rho_2 = 1$, hence

$$\beta < \beta_c \quad z_B = \frac{1}{2} e^{d\beta} (1 + e^{-2\beta})^d 2^{1-d} = (\cosh \beta)^d \quad (43)$$

as expected. At low temperature $\beta > \beta_c$, the magnetization is given by

$$M = \frac{1 - p^2}{1 - 2\tau p + p^2} \quad (44)$$

As $\beta \rightarrow \beta_c + 0$

$$M^2 \sim \frac{2d}{2d-1} 3 \left(\frac{\beta}{\beta_c} - 1 \right) \quad (45)$$

in agreement with the mean field prediction (11) for $d \rightarrow \infty$. The free energy is the logarithm of z_B . If we set $p = e^{-u}$, we have for $\beta > \beta_c$

$$z_B = \left[\cosh \beta (1 - e^{-2\beta}) \frac{\sinh du}{\sinh u} \right]^d (\cosh du)^{1-d}$$

$$e^{-2\beta} = \frac{\sinh(d-1)u}{\sinh du}, \quad u^2 = \frac{12}{2d-1} (\beta - \beta_c) + \dots \quad (46)$$

Close to criticality

$$\beta > \beta_c \quad z_B / (\cosh \beta)^d = 1 + O((\beta - \beta_c)^2) \quad (47)$$

meaning that Bethe's approximation predicts a second order transition with a discontinuous specific heat.

Compute the discontinuity of the specific heat, and compare it with the mean field prediction.

We observe within the same approximation an important property. Namely, if the external field is nonvanishing, for instance $h > 0$, all thermodynamic quantities become regular. In particular no singularity shows up when inverting equation (41) for p as a function of β (or $\tau = e^{-2\beta}$). In figure 3 we plot the inverse function $\tau(p)$, comparing the case $h = 0$ to $h \neq 0$ for $d = 3$.

It is legitimate to ask about the fate of the critical singularity as h takes a nonvanishing value. A bifurcation takes place and, for $\beta < \beta_c$, one finds two complex conjugate singularities on the imaginary h axis. More precisely, if we look for values where

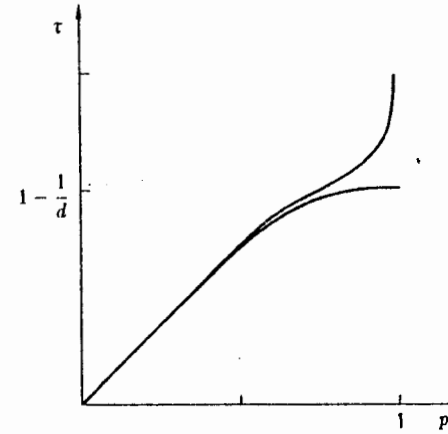


Fig. 3 Bethe approximation. Representation of τ as a function of p for h equal or different to zero, assuming $d = 3$.

the solution $p(\beta, h)$ of equation (41), becomes multivalued as a function of h for real $\beta < \beta_c$, we find that h has to be pure imaginary (and p a phase) with absolute value larger than h_c . As $\tau = e^{-2\beta} \rightarrow \tau_c$, one has

$$h_c \sim a \left(\frac{\tau}{\tau_c} - 1 \right)^{3/2} \quad a = \frac{2}{3} \left(\frac{\tau_c}{1 - \tau_c} \right) \left(\frac{2}{1 - \tau_c^2} \right)^{1/2} \quad (48)$$

At any of the two points $\pm ih_c$, the free energy has a branch point of order 3/2, with a singular part

$$f_{\text{sing.}} \sim \text{cst} (h \pm ih_c)^{3/2} \quad (49)$$

This phenomenon is not an artefact of the approximation, but is the reflection of a singularity discovered by Lee and Yang to be discussed below.

3.1.5 Critical exponents

Let us summarize what has been learnt from these elementary approximations. We are interested in the vicinity of critical points with long-range correlations for the dynamical variables. On this scale, the discontinuous lattice structure can be ignored, as was

verified in the previous chapter in the case of the two-dimensional Ising model. Mean field ideas led us to minimize a trial free energy $V(\mathbf{M}) - \mathbf{h} \cdot \mathbf{M}$, a function of mean magnetization $\mathbf{M}$, which plays the role of an order parameter, of an external field $\mathbf{h}$ coupled linearly to $\mathbf{M}$, and other parameters such as the temperature. For instance, in an $O(n)$ symmetric vector model, we found a typical expression of the form

$$V(\mathbf{M}) = \frac{1}{2}(1 - 2d\beta)\mathbf{M}^2 + \frac{1}{4} \frac{n}{n+2} (\mathbf{M}^2)^2 + \dots$$

Note that $\mathbf{M}$ is no longer a vector of unit length, as was the case for the original variables in the partition function. We recognize in $V(\mathbf{M})$ the potential part of a Lagrangian, with dynamical variables $\mathbf{M}$, in self-interaction. In field theory, in the classical limit, when one neglects the contributions of the kinetic term due to the possible spatial dependence of $\mathbf{M}(\mathbf{x})$, one similarly minimizes the potential term to find the ground state of the system. The coefficient of the quadratic term may be identified as a square mass, however not restricted to be positive in the present context. To take this into account, and recall that this square mass vanishes at criticality (here linearly), we denote this coefficient by θ . Also let us replace the coefficient $n/(n+2)$ of the $(\mathbf{M}^2)^2$ term by a positive coupling constant g to prepare for a more general situation (as was already envisioned in chapter 1). Finally, we neglect higher order terms and write

$$V(\mathbf{M}) = \frac{1}{2}\theta\mathbf{M}^2 + \frac{1}{4}g(\mathbf{M}^2)^2 \quad (50)$$

Let us now retrace the various steps which in one guise or another lead, within the mean field approximation, to the critical exponents. Minimizing $V(\mathbf{M}) - \mathbf{h} \cdot \mathbf{M}$ yields

$$\nabla_{\mathbf{M}} V(\mathbf{M}) \equiv \mathbf{M}(\theta + g\mathbf{M}^2) = \mathbf{h} \quad (51)$$

(i) In the absence of external field, $\mathbf{h} = 0$, the solution is

$$\begin{array}{lll} \theta > 0 & \mathbf{M} = 0 & V_{\min} = 0 \\ \theta < 0 & \mathbf{M}_s^2 = -\theta/g & V_{\min} = -\frac{1}{4}\theta^2/g \end{array} \quad (52)$$

The distinction $\theta > 0, \theta < 0$, corresponds to high and low temperature. The specific heat is proportional to $\partial^2 V_{\min}/\partial\theta^2$, and hence discontinuous. One writes generally

$$C(\theta) \underset{\theta \rightarrow \pm 0}{\sim} C_{\pm} |\theta|^{\alpha_{\pm}} \quad (53a)$$

To encompass the discontinuous mean field result, one agrees to set

$$\alpha = \alpha_{\pm} = 0 \quad (53b)$$

(ii) For $\theta < 0$, the spontaneous magnetization is given by $\mathbf{M}_s^2 = -\theta/g$ with

$$\theta < 0 \quad |\mathbf{M}_s| \sim (-\theta)^{\beta} \quad \beta = \frac{1}{2} \quad (54)$$

(iii) In the high temperature phase $\mathbf{M}(\mathbf{h})$ vanishes linearly with $\mathbf{h}$, $\mathbf{M}(\mathbf{h}) \sim \mathbf{h}/\theta$, the coefficient being the susceptibility, which diverges when $\theta \rightarrow 0$ as

$$\theta > 0 \quad \chi \sim \chi_+ \theta^{-\gamma_+} \quad \gamma_+ = 1 \quad (55)$$

In the low temperature phase we saw that the longitudinal and transverse response have to be distinguished. In tensor form

$$d\mathbf{h}_a = (\chi^{-1})_{ab} d\mathbf{M}_b = [(\theta + 3g\mathbf{M}^2)\delta_{ab} + 2g(\mathbf{M}_a \mathbf{M}_b - \mathbf{M}^2 \delta_{ab})] d\mathbf{M}_b$$

For $\theta > 0$ we recover the previous result, but for $\theta < 0$ and a spontaneous magnetization $\mathbf{M}^2 = -\theta/g$, we have

$$(\chi^{-1})_{ab} \sim -2\theta \frac{\mathbf{M}_{sa} \mathbf{M}_{sb}}{\mathbf{M}_s^2}$$

In the transverse directions, χ^{-1} is strictly zero. This is Goldstone's phenomenon. In the parallel direction

$$\begin{array}{lll} \chi^{\parallel} \sim \frac{1}{2} \frac{1}{(-\theta)} \\ \theta < 0 & \chi^{\parallel} \sim \chi_- (-\theta)^{-\gamma_-}, & \gamma_- = 1 \\ & \chi_+/\chi_- = 2 \end{array} \quad (56)$$

(iv) At criticality, $\theta = 0$, the magnetization varies as the cubic root of the external field. In other words the response to an external field is much larger (as compared to the case when $\theta > 0$), signalling the transition to an ordered state. We write

$$|\mathbf{M}| \sim |\mathbf{h}|^{1/\delta} \quad \delta = 3 \quad (57)$$

(v) Let us now assume that the external field varies from point to point, and reinstate the "kinetic term". Choose $n = 1$ for simplicity (although for $n > 1$ the following expressions will remain valid for the longitudinal components). We have to

minimize the complete expression

$$S = \sum_i \left\{ -\frac{1}{2} M_i \sum_{j(i)} (M_j - M_i) + V(M_i) - h_i M_i \right\} \quad (58)$$

Thus

$$-\sum_{j(i)} (M_j - M_i) + (\theta + g M_i^2) M_i = h_i \quad (59)$$

For $\theta > 0$, this leads to

$$M_i = \sum_j \int \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot (x_i - x_j)}}{\theta + \sum_\mu 2(1 - \cos k_\mu)} h_j + O(gh^3) \quad (60)$$

As h goes to zero and at large distance $|x_i - x_j| \rightarrow \infty$, we find

$$\langle M_i M_j \rangle = \frac{\delta M_i}{\delta h_j} \Big|_{h=0} = \int \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot (x_i - x_j)}}{\theta + \sum_\mu 2(1 - \cos k_\mu)} \frac{\theta^{(d-2)/2}}{2 \left(2\pi \theta^{1/2} |x_i - x_j| \right)^{(d-1)/2}} \exp \left(-\theta^{1/2} |x_i - x_j| \right) \quad (61)$$

The exponential decrease as $\exp[-|x_i - x_j|/\xi]$ defines a correlation length ξ . The latter diverges as one approaches the critical point

$$\theta \rightarrow +0 \quad \xi \sim \theta^{-\nu} \quad \nu = \frac{1}{2} \quad (62)$$

For small negative θ , M_i fluctuates around the spontaneous magnetization M_s ,

$$M_i = M_s + \sum_j \frac{\delta M_i}{\delta h_j} \Big|_{h=0} h_j + \dots \quad (63)$$

$$\theta < 0, \quad \frac{\delta M_i}{\delta h_j} \Big|_{h=0} = \int \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot (x_i - x_j)}}{-2\theta + \sum_\mu 2(1 - \cos k_\mu)}$$

Therefore for $\theta \rightarrow -0$, the correlation length behaves as $(-2\theta)^{-1/2}$ and the corresponding exponent remains $\nu = 1/2$ as above.

For an n -component model, transverse modes have an infinite correlation length below criticality.

(vi) At $\theta = 0$, both (61) and (63) reduce to

$$\frac{\delta M_i}{\delta h_j} \Big|_{h=0} = \int \frac{d^d k}{(2\pi)^d} \frac{e^{ik \cdot (x_i - x_j)}}{\sum_\mu 2(1 - \cos k_\mu)} \frac{\Gamma(d/2 - 1)}{4\pi^{d/2} |x_i - x_j|^{d-2}} \quad (64)$$

a generalization in d -dimensions of the Coulomb potential. Of course all this assumes d large enough. When comparing equations (61) and (64), we note that the inverse powers of $|x_i - x_j|$ differ (except when $d = 3$). One defines the exponent η through

$$\theta = 0 \quad \frac{\delta M_i}{\delta h_j} \Big|_{h=0} = \langle M_i M_j \rangle \Big|_{|x_i - x_j| \rightarrow \infty} \frac{\text{cst}}{|x_i - x_j|^{d-2+\eta}} \quad (65a)$$

and thus in the mean field case

$$\eta = 0 \quad (65b)$$

(vii) Assume now a uniform external field h , and consider the one component case only for simplicity. The minimization equation (51) becomes ambiguous when its derivative with respect to M vanishes, marking the onset of a multivalued solution for M . This corresponds to requiring simultaneously

$$\begin{aligned} \theta M + g M^3 &= h \\ \theta + 3g M^2 &= 0 \end{aligned} \quad (66)$$

For a solution in M , the parameters θ , g and h have to be related by the vanishing of the discriminant

$$4\theta^3 + 27gh^2 = 0 \quad (67)$$

Real solutions with $\theta < 0$ can be left aside, since they ignore spontaneous symmetry breaking (recall that $g > 0$). For positive θ , one finds two complex conjugate values of h on the imaginary axis, $h = \pm i h_c$, with

$$h_c = \left(\frac{4}{27g} \right)^{1/2} \theta^{3/2} \quad (68)$$

This is the Lee-Yang edge singularity, derived in the previous subsection. If we set for small positive θ ,

$$h_c \sim \theta^\Delta \quad (69a)$$

Table II. Critical exponents.

	Ising model $d = 2$	Mean field $d \rightarrow \infty$	Spherical model ($n \rightarrow \infty$) $4 > d > 2$
Specific heat $C \sim C_{\pm} \theta ^{-\alpha_{\pm}}$	$\alpha = 0^{(*)}$	$\alpha = 0^{(**)}$	$\alpha = \left(\frac{d-4}{d-2}\right)$
Spontaneous magnetization $\theta < 0 \quad M_{sp} \sim (-\theta)^{\beta}$	$\beta = \frac{1}{8}$	$\beta = \frac{1}{2}$	$\beta = \frac{1}{2}$
Susceptibility $\chi \sim \theta ^{-\gamma_{\pm}}$	$\gamma = \frac{7}{4}$	$\gamma = 1$	$\gamma = \frac{2}{d-2}$
Magnetization at $\theta = 0$ $M \sim h^{1/\delta}$	$\delta = 15$	$\delta = 3$	$\delta = \frac{d+2}{d-2}$
Correlation length $\xi \sim \theta ^{-\nu_{\pm}}$	$\nu = 1$	$\nu = \frac{1}{2}$	$\nu = \frac{1}{d-2}$
Critical correlation $\langle M(\mathbf{x})M(\mathbf{y}) \rangle \sim \frac{\text{cst}}{ \mathbf{x} - \mathbf{y} ^{d-2+\eta}}$	$\eta = \frac{1}{4}$	$\eta = 0$	$\eta = 0$
Lee-Yang edge singularity $h = \pm ih_c, \theta > 0$ $h_c \sim \theta^{\Delta}, M \sim (h \mp ih_c)^{\sigma}$	$\Delta = \frac{15}{8}$ $\sigma = -\frac{1}{6}$	$\Delta = \frac{3}{2}$ $\sigma = \frac{1}{2}$	

Notes: (*) up to logarithms. (**) discontinuous.

the mean field prediction is

$$\Delta = \frac{3}{2} \quad (69b)$$

With $h_0 = \pm ih_c$, $M_0 = \pm iM_c$, $M_c = \frac{3}{2}h_c/\theta = \frac{3}{2}\theta^{1/2}$, the behaviour of the free energy in the vicinity of these points is

$$V(M) - V(M_0) - h(M - M_0) \approx \left[\frac{4}{27gM_0} (h - h_0) \right]^{3/2} - M_0(h - h_0) \quad (70)$$

with a singular part of the form

$$(h - h_0)^{1+\sigma} \quad (71a)$$

such that the corresponding mean field value is

$$\sigma = \frac{1}{2} \quad (71b)$$

To proceed further, we will have to study the corrections, due to fluctuations, which were neglected up to now. In large

enough dimension, these effects are expected not to modify in an essential way the nature of the transition, in particular the critical exponents. This will be discussed in the last section of this chapter. The largest dimension in which the critical behaviour is effectively modified, as indicated by the breakdown of the perturbative estimate of fluctuations, is called the upper critical dimension. From the preliminary discussion in chapter 1, we guess that its value is $d = 4$ for the conventional transition in spin models. It turns out to be $d = 6$ for the Lee-Yang (complex) edge singularity. Other more powerful tools will be needed for the analysis, which go under the name of the renormalization group. This will be the subject of chapter 5. Before we undertake a full-fledged discussion of these matters, we will take a closer look at the meaning of the Lee-Yang singularity as a precursor of symmetry breaking, and analyse an alternative (large n or spherical model) approximation, which throws some light on the upper critical dimension. For the time being, we summarize in table II the mean field critical exponents and compare them with those of the two-dimensional Ising model, as well as those of the spherical model for $2 < d < 4$, where a departure with respect to mean field predictions is observed, as discussed below.

3.2 Lee-Yang zeroes

In 1952 two famous papers appeared by Lee and Yang where they took a different point of view towards the phenomenon of spontaneous symmetry breaking by studying a lattice gas, or, what amounts to the same, an Ising model, in an arbitrary external field, in a finite (but arbitrarily large) domain. In their approach, dimension or coordination does not play a pre-eminent role at first, but the external field is allowed to take complex values in a search of zeroes of the partition function. Since the latter describes a sum over positive weights in the real (physical) domain, it is certainly nonvanishing there, but, as parameters become complex, this is not the case anymore. In a lattice gas interpretation, spin up or down ($\sigma = +1$ or $\sigma = -1$) correspond to occupied or empty sites and the field h is conjugate to the density. Here we will take the "magnetic" interpretation. The important point is that for real β , up to a nonvanishing factor, the partition function is a finite

volume (i.e. for finitely many sites) is a polynomial in the activity $\exp(-2h)$.

3.2.1 The Lee-Yang theorem

The discovery of Lee and Yang is that the zeroes of the partition function are located on the unit circle of the complex activity plane, or equivalently on the imaginary h -axis. For any finite volume, or for that matter on an arbitrary finite graph, the number of zeroes is finite, as the partition function is a polynomial in the activity, and the degree of this polynomial increases with the size. In the infinite volume limit, the zeros will "condense" on a subset of the unit circle. As long as β is chosen in an appropriate range (high temperature), this singular set will leave a vicinity of unit activity (or zero field) free of zeroes. The free energy will remain analytic there. But, as β is increased, the analyticity gap will eventually close, forbidding analytic continuation from $\text{Re } h > 0$ to $\text{Re } h < 0$, a signal of a transition at $h = 0$. As seen along a path h real, h crossing zero, the transition is a first order one, with a jump in magnetization (the discontinuity) being twice the spontaneous magnetization, while the free energy at zero field might have a continuous derivative in β at β_c as is usual in an ordinary second order transition. The analysis is carried out in the activity plane because the Lee-Yang theorem controls the location of zeroes, but as long as the model allows identification of the partition function relative to a finite domain as a polynomial in some parameter, it can obviously be generalized. Such a study has been carried out for instance using the variable $\exp(-2\beta)$ for the Ising model, and gives an insight into the analytic structure of the free energy. Another extension is relative to models involving a continuous symmetry group. The subject has never been developed in as systematic a fashion as would be desirable, if only because in general these distributions of zeroes may be rather wild and lattice dependent. They enable one, however, to reach some interesting results, as will be shown later. Here we concentrate on the activity variable for a general Ising model on arbitrary graph of N sites (vertices of the graphs) with at most one link joining a pair of vertices, which are then said to be adjacent, or neighbours. The total number of links is L .

Let h_i denote the magnetic field at site i . It is convenient at first to have the field vary from site to site. Define

$$\rho_i = e^{-2h_i} \quad \tau = e^{-2\beta} \quad (72)$$

The partition function is of the form

$$Z_N = \frac{1}{2^N} \exp \left(\beta L + \sum_i h_i \right) P(\tau, \rho_i) \quad (73)$$

with P a polynomial in τ and ρ_i , obtained as

$$P = \sum_{\sigma_i = \pm 1} \exp \left[\beta \sum_{(ij)} (\sigma_i \sigma_j - 1) + \sum_i h_i (\sigma_i - 1) \right] \quad (74)$$

The summation $\sum_{(ij)}$ runs over all links. The polynomial P is of degree one in each ρ_i separately, and globally of degree N in all of them. For τ and ρ_i real and positive, P is obviously positive and hence cannot vanish. We assume in the sequel $0 < \tau < 1$, i.e. β real positive, which is the ferromagnetic regime. As examples, let us write the polynomials P corresponding to the simplest graphs

$$\begin{aligned} & \overset{1}{\bullet} \text{---} \overset{2}{\bullet} \quad P_{12} = 1 + \tau(\rho_1 + \rho_2) + \rho_1 \rho_2 \quad (75) \\ & \overset{1}{\bullet} \text{---} \overset{3}{\bullet} \text{---} \overset{2}{\bullet} \quad P_{123} = (1 + \rho_1 \tau)(1 + \rho_3 \tau) + \rho_2(\tau + \rho_1)(\tau + \rho_3) \end{aligned}$$

Up to a regular function, the (total) free energy is the logarithm of P . Its only singularities are therefore the zeroes of the polynomial. If we have a growing sequence of graphs (as parts of a lattice say), in the limit $N \rightarrow \infty$, the free energy will remain analytic in any bounded region outside the accumulation points of zeroes.

We now look more closely at the way the polynomials can be generated by building the graph step by step. Notice first that on a disjoint union P factorizes, $P = P_1 P_2$, where P_1 and P_2 refer to the two disjoint subsets. Starting from such a situation, we can generate a new one, call it (12), by identifying a site a from subset 1 to a site b from subset 2. Let $P_{(12)}$ be the corresponding polynomial.

Now P_1 is linear in ρ_a , P_2 linear in ρ_b , in such a way that

$$P_1 = A_+ + \rho_a A_- \quad P_2 = B_+ + \rho_b B_- \quad (76)$$

Here $A_+(B_+)$ refers to the contribution with $\sigma_a = +1$ ($\sigma_b = +1$) while $A_-(B_-)$ refers to the one with $\sigma_a = -1$ ($\sigma_b = -1$). When a and b are identified, a new activity variable ρ_{ab} is attached to the

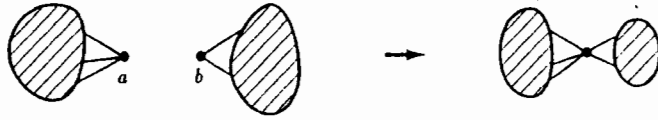


Fig. 4 The contraction process.

site, and we have the following contraction process (figure 4)

$$\begin{aligned} P_1 P_2 &\equiv A_+ B_+ + \rho_a A_- B_+ + \rho_b A_+ B_- + \rho_a \rho_b A_- B_- \rightarrow \\ P_{12} &\equiv A_+ B_+ + \rho_{ab} A_- B_- \end{aligned} \quad (77)$$

Check the contraction process on the examples (75). Construct the polynomial P for a cyclic graph.

Nothing prevents the contraction process from being extended to a single connected part, where at first sites a and b are distinct, then identified as a single site ab with activity variable ρ_{ab} . Thus

$$P \equiv A_{++} + A_{-+}\rho_a + A_{+-}\rho_b + A_{--}\rho_a\rho_b \rightarrow P_{ab} \equiv A_{++} + \rho_{ab}A_{--} \quad (78)$$

It is assumed that in the contracted graph a pair of vertices is joined by at most one link.

The contraction process allows one to obtain the polynomial P for an arbitrary graph starting from the elementary result for the simplest graph of two sites joined by a link. In this case, from equation (75a), the polynomial vanishes for

$$\rho_1 = -\frac{1 + \tau\rho_2}{\tau + \rho_2} \quad (79)$$

This relation defines a global (conformal) one-to-one map from the ρ_2 complex plane to the ρ_1 plane. It obviously leaves the unit circle invariant (for real τ) and for $0 < \tau < 1$, exchanges the interior with the exterior of this circle. It follows therefore that if $|\rho_1| < 1$ and $|\rho_2| < 1$, or $|\rho_1| > 1$, and $|\rho_2| > 1$, the polynomial cannot vanish. This property generalizes in that, for an arbitrary graph, if all ρ_i lie inside, or all ρ_i lie outside the unit circle, P is different from zero. To prove that this is true, it is sufficient to check that this property survives the contraction process. Assume therefore that for a given graph $P(\rho_i) \neq 0$ when $|\rho_i| < 1$ for all i .

When the dependence on ρ_a and ρ_b is made explicit, P is given by (78), while P_{ab} is also a function of the set $\{\rho_i\} - \{\rho_a, \rho_b\}$, through A_{++} and A_{--} , and a new ρ_{ab} . Fix the ρ_i 's distinct from ρ_a and ρ_b inside the unit circle. We want to show that $|A_{++}| > |A_{--}|$, in which case P_{ab} will be nonvanishing for $|\rho_{ab}| < 1$. But $P \neq 0$ for $|\rho_a| < 1, |\rho_b| < 1$. We can take $\rho_a = \rho_b = \rho$. Hence as a function of ρ the second degree polynomial $A_{++} + \rho(A_{+-} + A_{-+}) + \rho^2 A_{--}$ has its two roots of modulus equal or larger than unity, which indeed means that $|A_{++}| \geq |A_{--}|$. Thus P_{ab} is different from zero when all its ρ 's are inside the unit circle, and this property holds for any graph. Now set all $\rho_i \equiv \rho$ and we have a uniform field. From the symmetry property under reversal of the field $h \rightarrow -h, \rho \rightarrow \rho^{-1}$ and $Z(h) = Z(-h)$, we have $e^{+Nh} P(\tau, \rho) = e^{-Nh} P(\tau, \rho^{-1})$ or

$$P(\tau, \rho) = \rho^N P(\tau, \rho^{-1}) \quad (80)$$

Hence if $P \neq 0$ for $|\rho| < 1$, one has also $P \neq 0$ for $|\rho| > 1$. The Lee-Yang theorem follows, stating that the partition function can (and does) only vanish on the unit circle $|\rho| = 1$.

We can check this in some limiting cases. At infinite temperature, $\tau = 1$, and P reduces to $P(1, \rho) = (1 + \rho)^N$, while at zero temperature, $\tau = 0$, $P(0, \rho) = 1 + \rho^N$. Thus decreasing τ , one starts from a degenerate zero with multiplicity N at $\rho = -1$ at infinite temperature, to a uniform distribution $\rho_k = e^{i\pi(2k+1)/N}$. As long as N is finite, no zero is on the real axis ($\rho = 1, h = 0$).

This remarkable result is so general that it does not allow us to find the transition in the infinite volume limit but only suggests the scenario that was described at the beginning of this section. To prove the existence of a transition (or the closure of the analyticity gap) requires more sophisticated means. As was also said before, the same result extends to a φ^4 self-interaction or to models with a continuous symmetry such as the $O(2)$ or rotator model. We have already seen that it shows up in mean field, as well as in the Bethe approximation; in each case it is therefore not an artefact.

3.2.2 The one-dimensional case

As $\tau \rightarrow \tau_c$ we saw that mean field ($d \rightarrow \infty$) predicted an analyticity gap shrinking as $(\tau - \tau_c)^{3/2}$ and a singular behaviour of the free energy as $(h \mp ih_c)^{3/2}$. At the other extreme, let us consider the case $d = 1$. For periodic boundary conditions.

using the transfer matrix formalism of section 1.2.5, the partition function is given by

$$Z_N = \left(\frac{1}{2}e^{\beta+h}\right)^N \text{Tr } T^N \quad T = \begin{pmatrix} 1 & \tau\rho^{1/2} \\ \tau\rho^{1/2} & \rho \end{pmatrix} \quad (81)$$

Thus

$$\text{Tr } T^N = \lambda_+^N + \lambda_-^N \quad (82)$$

where $\lambda_{\pm}$ are the two eigenvalues of T , given by

$$(1-\lambda)(\rho-\lambda) - \rho\tau^2 = 0$$

$$\lambda_{\pm} = \frac{1}{2}(1+\rho) \pm \sqrt{\frac{1}{4}(1-\rho)^2 + \rho\tau^2} \quad (83)$$

In the thermodynamic limit $N \rightarrow \infty$, for $\rho > 0$ (h real), $0 < \tau < 1$, only the largest eigenvalue λ_+ contributes

$$F = \lim_{N \rightarrow \infty} \frac{1}{N} \ln Z_N = \beta + h - \ln 2 + \ln \left\{ \frac{1}{2}(1+\rho) + \sqrt{\frac{1}{4}(1-\rho)^2 + \rho\tau^2} \right\} \quad (84)$$

Singularities in ρ arise from the vanishing of the square root, i.e.

$$1 + 2\rho(2\tau^2 - 1) + \rho^2 = 0$$

$$\rho_c^{(\pm)} = e^{\pm i2h_c} \quad (85)$$

$$\cos 2h_c = 1 - 2\tau^2 = 1 - 2e^{-4\beta}$$

We find a pair of singularities on the unit circle (in ρ) as expected. As τ decreases from 1 to 0, the pair starts at -1 ($\cos 2h_c = -1$) and ends at $+1$, zero temperature playing the role of critical temperature in dimension one. As $\tau \rightarrow 0$, $h_c \sim \tau$ (compare with $h_c \sim (\tau - \tau_c)^{3/2}$ for $d \rightarrow \infty$), while the singular part of the free energy behaves as $(h \mp ih_c)^{1/2}$. Therefore the exponent σ defined in equation (11a) is $-1/2$ for $d = 1$. The upper critical dimension is six for the Lee-Yang edge singularity. As a result σ will vary between $-1/2$ ($d = 1$) to $+1/2$ for $d = 6$, then stay constant above six dimensions. It turns out that $\sigma = -1/6$ in dimension $d = 2$ (chapter 9), in agreement with the idea of smooth behaviour as a function of dimensionality.

3.2.3 General properties

A point that needs emphasis is the claim that the Lee-Yang edge singularity is independent of the temperature above T_c , or in

other words the exponent σ is temperature independent. The one-dimensional calculation supports this claim, as did the mean field approximation. Therefore the question may be studied for vanishingly small β and h close to $i\pi/2$ (i.e. $\rho = e^{-2h}$ close to -1). For h complex, one is not dealing anymore with a genuine statistical model, i.e. admitting a (classical) probabilistic interpretation. Rather we have reached a point where field theory proper is taken seriously. Along the imaginary h axis, the effective theory is described by a self-interacting scalar field φ , with a dominant symmetry breaking and imaginary interaction term $i\varphi^3$ for small fluctuations. This arises from a shift in the field φ designed so as to compensate the linear term induced by the coupling to the external field. This remark enables one to understand in a heuristic way the upper critical dimension, where, as we shall see later, the effective coupling constant is dimensionless. Since from chapter 1 a scalar field has dimension $(d-2)/2$, to insure that a term $\int d^d x \varphi^3(x)$ is dimensionless we need $\frac{3}{2}(d-2) = d$, i.e. $d = 6$.

While for any finite N the Lee-Yang zeroes appear as singularities of the (total) free energy, this need not be the case in the infinite volume limit, as illustrated again by the one-dimensional calculation. Only the boundary points of the accumulation set survive certainly as singularities, hence the name "edge singularity". This phenomenon is familiar when a function has a branch point, but admits analytic continuation around the singularity, even though in some finite approximation it may turn out that there is a preferred locus for a branch cut. This is indeed the case if we return to the general discussion and approximate an infinite lattice of coordination number q by a finite volume of N sites, then take the infinite volume limit. The free energy per site is

$$F = \frac{1}{2}q\beta + h - \ln 2 + \lim_{N \rightarrow \infty} \frac{1}{N} \ln P_N(\tau, \rho) \quad (86)$$

Since P_N is normalised to unity for $\rho = 0$, we have a factorization in terms of zeroes of unit modulus

$$P_N(\rho, \tau) = \prod_{a=1}^N \left(1 - \frac{\rho}{\rho_a(\tau)}\right) \quad (87)$$

Let the zeroes accumulate for $N \rightarrow \infty$ on the unit circle $\rho = e^{i\varphi}$, with a limiting (τ -dependent) density $N\mu(\varphi)$, $\mu(\varphi) = \mu(-\varphi) \geq 0$.

The symmetry property is a reflection of the original symmetry in $h \leftrightarrow -h$. Normalization requires

$$\int_{-\pi}^{+\pi} d\varphi \mu(\varphi) = 1 \quad (88a)$$

This leads to a representation of F as

$$F = \frac{1}{2}q\beta + h - \ln 2 + \frac{1}{2} \int_{-\pi}^{+\pi} d\varphi \mu(\varphi) \ln(1 + \rho^2 - 2\rho \cos \varphi) \quad (88b)$$

where the contributions of conjugate zeroes have been grouped together. The representation (88) is valid in the whole range $0 < \tau < 1$. Below the critical temperature, where the support of $\mu(\varphi)$ is the full circle, it defines in general two distinct analytic functions, one for $|\rho| < 1$, the other for $|\rho| > 1$. At $\rho = 1$ ($h = 0$), F is continuous, but its derivative with respect to h , i.e. the magnetization

$$\begin{aligned} M(\rho) &= 1 + 2 \int_{-\pi}^{+\pi} d\varphi \mu(\varphi) \frac{\rho}{e^{i\varphi} - \rho} = -1 + 2 \int_{-\pi}^{+\pi} d\varphi \mu(\varphi) \frac{1}{1 - \rho e^{-i\varphi}} \\ &= \int_{-\pi}^{+\pi} d\varphi \mu(\varphi) \frac{1 - \rho^2}{1 - 2\rho \cos \varphi + \rho^2} \end{aligned} \quad (89)$$

which vanishes as $h \rightarrow 0$ for $\tau > \tau_c$, becomes discontinuous for $\tau < \tau_c$, when the support of μ covers the circle. One has

$$\tau < \tau_c \quad M_{\pm} = \lim_{h \rightarrow \pm 0} M = \pm 2\pi \mu(0) \quad (90)$$

So $\mu(0) = M_{sp}/2\pi$. From (88), the free energy and the magnetization have an expansion in terms of the (trigonometric) moments of $\mu(\varphi)$. Set

$$\mu_n = \int_{-\pi}^{\pi} d\varphi \mu(\varphi) \cos n\varphi \quad \mu_0 = 1 \quad (91)$$

then

$$\begin{aligned} F &= \frac{1}{2}q\beta + h - \ln 2 - \sum_{n=1}^{\infty} \mu_n \frac{\rho^n}{n} \\ M &= 1 + 2 \sum_{n=1}^{\infty} \mu_n \rho^n \end{aligned} \quad (92)$$

which appear as strong field expansions (small ρ) away from a fully ordered situation (all spins up say, hence $M = 1$). Conversely if such an expansion is known, from the standard moment

problem, one can try to recover the τ -dependent distribution $\mu(\cdot)$. Unfortunately the best that can be done is to obtain finitely many terms in the series (92) and therefore additional information is needed (for instance a parametrization of the singular behaviour of μ near the edge of its support).

In the case $d = 1$, obtain the representation (88) in the form

$$F = \beta + h - \ln 2 + \frac{1}{2} \int_{-\pi}^{+\pi} d\varphi \frac{\theta(\cos \varphi_0 - \cos \varphi)}{2\pi} \quad (93)$$

$$\begin{aligned} &\times \left[\frac{1 - \cos \varphi}{\cos \varphi_0 - \cos \varphi} \right]^{1/2} \ln [1 + \rho^2 - 2\rho \cos \varphi] \\ &\cos \varphi_0 = 1 - 2\tau^2 = 1 - 2e^{-4\beta} \end{aligned} \quad (94)$$

As $\tau \rightarrow 0$ the corresponding distribution $\mu(\varphi)$ tends to be uniform, while as $\tau \rightarrow 1$ it gets concentrated at $\varphi = \pm\pi$. For $0 < \tau < 1$, the edge singularity of $\mu(\varphi)$ is proportional to $(\cos \varphi_0 - \cos \varphi)^{-1/2}$. Compute the moments μ_n .

3.2.4 Zeroes in the temperature plane

No general theorems are known which constrain the zeroes in the complex temperature plane, but it is tempting to analyze their location either numerically or otherwise. We now assume the external field to vanish ($\rho = 1$). Fisher observed that for an Ising model on a square two-dimensional lattice, the zeroes are located in the complex τ plane on two circles, one of them being given by the equation $\bar{\tau} = (1 - \tau)(1 + \tau)^{-1}$, the other obtained by changing τ into τ^{-1} , i.e. inverting with respect to the unit circle. This holds in the thermodynamic limit. For a finite system, some zeroes might violate the above property, depending on the specific boundary conditions. The locus of zeroes is invariant under duality, and, as a consequence, goes through the transition point on the real axis. But more important is the fact that in this case

(i) the complex zeroes close to the critical point accumulate on curves,

(ii) for the two-dimensional model, these curves cross the real axis at a right angle.

Recover the above results from the expressions given in chapter 2.

There seems to be some evidence that, in higher dimension, the complex temperature zeroes accumulate also on curves close to the (real) critical temperature. If we assume this to be true, we can draw some interesting consequences regarding the nature of the singularity of the free energy. Because of the reality properties of the partition function, two curves of zeroes will be incident at T_c , one in the upper half plane and a reflected one in the lower half plane. Let $\pi - \varphi$ be the angle between the positive real T -axis and the upper line of zeroes, and let

$$f_{\text{sing}}^{(\pm)} \sim A_{\pm} \left(\frac{T - T_c}{T_c} \right)^{2-\alpha}$$

be the singular part of the free energy above (plus sign) and below (minus sign) on the real axis. Then both $f_{\text{sing}}^{\pm}$ extend as analytic functions up to the curves of zeroes where their real parts are equal. The curves of zeroes play the role of Stokes lines, where different asymptotic behaviours have to match (in modulus for the partition function). This is consistent only if the following relation holds between the above angle φ , the ratio A_-/A_+ of critical amplitudes and the exponent α

$$\tan[(2-\alpha)\varphi] = \frac{\cos \pi\alpha - A_-/A_+}{\sin \pi\alpha} \quad (95)$$

as a consequence of the above matching condition. If we let $\alpha \rightarrow 0$ and $\varphi \rightarrow \pi/2$ as in $d = 2$, this relation is compatible only if $A_- = A_+$, which is the observed value. The mean field (or large d) approximation gives $\varphi = \pi/4$ (since $\alpha = 0$ and $A_+/A_- = 0$). When $d = 3$, inserting the value of $\alpha \sim 0.11$, and an observed angle φ of the order of 57 degrees, one predicts a ratio A_+/A_- close to 0.5. A more accurate measurement of the angle φ could, however, significantly modify this ratio.

3.3 Large n limit

In this section, we investigate the limit of an n -component vector model as $n \rightarrow \infty$. Of course $O(n)$ symmetry is assumed. It turns out that the solution agrees with the one of the spherical model, solved by Berlin and Kac in the fifties. The $1/n$ corrections

are, however, different. Apart from its technical interest, this solution reveals, in as simple a setting as possible, the failure of mean field theory below the upper critical dimension $d = 4$, for just the reasons which were discussed in the first chapter. The spherical model involves scalar variables φ_i at each site, nearest neighbour interactions, and the constraint $\sum_i \varphi_i^2 = \text{cst}$. The clue to the equivalence between the two models lies in the saddle point treatment to which we turn now.

3.3.1 Saddle point method

As the number of field components n per site goes to infinity, we use a saddle point method to evaluate the path integral. This will be justified by showing that corrections are of order $1/n$. For N sites ($N \rightarrow \infty$) write the partition function as

$$\begin{aligned} Z &= Z_0 \int \prod_j d^n \varphi_j \delta(\varphi_j^2 - 1) \exp n\beta \left\{ \sum_{(ij)} \varphi_i \cdot \varphi_j \right\} \\ &= Z_0 \int \prod_j d^n \varphi_j \frac{n d\alpha_j}{2\pi} \\ &\quad \times \exp n \left\{ \sum_j (i\alpha_j + \lambda_j)(1 - \varphi_j^2) + \beta \sum_{(ij)} \varphi_i \cdot \varphi_j \right\} \\ Z_0 &= \left(\frac{\Gamma(n/2)}{\pi^{n/2}} \right)^{Nd} \end{aligned} \quad (96)$$

The normalizing factor Z_0 is such that $Z \rightarrow 1$ as $\beta \rightarrow 0$. In the second representation we have replaced the constraints $\varphi_j^2 = 1$ by a Fourier integral over a conjugate parameter α_j , while λ_j is arbitrary and will be chosen positive and sufficiently large to ensure convergence of later integrals. As a check, the solution should be independent of these values. The variables φ only occur in the exponential of a quadratic form. Set Δ for the lattice Laplacian, and define the matrix Q through

$$Q_{jk} = -\beta \Delta_{jk} + 2(\lambda_j + i\alpha_j - \beta d)\delta_{jk} \quad (97a)$$

Integrating over φ yields

$$Z = Z_0 \left(\frac{2\pi}{n} \right)^{\frac{1}{2}(n-2)N} \int \prod_j d\alpha_j \exp \frac{1}{2} n \left\{ \sum_j 2(i\alpha_j + \lambda_j) - \text{Tr} \ln Q \right\} \quad (97b)$$

Up to now, no approximation was involved. As $n \rightarrow \infty$ the prefactor in the exponential becomes infinite, while the term between brackets is n -independent. This is the basis for the saddle point method, which (unfortunately) amounts to interchanging the limits $N \rightarrow \infty$ (infinite volume) and $n \rightarrow \infty$ (infinitely many field components). The variational equation with respect to $2(\lambda_k + i\alpha_k)$, with k a space label, leads to the condition

$$1 = \langle k | Q^{-1} | k \rangle \quad (98)$$

We look for a translation invariant solution, hence such that $\lambda_k + i\alpha_k$ be k -independent. This in fact identifies the solution to the one pertaining to the spherical model, with scalar fields φ_k constrained to $\sum_k \varphi_k^2 = \text{cst}$, up to the overall prefactor n . For convenience, define

$$2(\lambda + i\alpha - d\beta) = \beta \xi^{-2} \quad (99)$$

Averaging equation (98) over all sites, and taking the infinite volume limit, gives

$$1 = \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr} Q^{-1} = \frac{1}{\beta} \int \frac{d^d p}{(2\pi)^d} \frac{1}{\xi^{-2} + 2 \sum_1^d (1 - \cos p_\mu)} \quad (100)$$

and to leading order

$$Z = Z_0 \left[\frac{1}{\beta} \left(\frac{2\pi}{\beta n} \right)^{(n-2)/2} \right]^N \times \exp \frac{nN}{2} \left\{ \beta(2d + \xi^{-2}) - \int \frac{d^d p}{(2\pi)^d} \ln \left[\xi^{-2} + 2 \sum_{\mu=1}^d (1 - \cos p_\mu) \right] \right\} \quad (101)$$

In both expressions the momentum p runs over a Brillouin zone, i.e. each component ranges from $-\pi$ to $+\pi$. The quantity ξ plays the role of a correlation length, and is determined self-consistently by equation (100) which is a stationarity condition of $\lim_{N \rightarrow \infty} N^{-1} \ln Z$ with respect to ξ^{-2} .

Let ξ vary from zero to infinity in equation (100). Correspondingly β runs from zero ($\xi^2 \sim \beta$ for $\xi \sim 0$) to some critical value $\beta_c(\xi = \infty)$ with

$$\beta_c = \int \frac{d^d p}{(2\pi)^d} \frac{1}{2 \sum_1^d (1 - \cos p_\mu)} = \frac{1}{2} \int_0^\infty d\alpha e^{-\alpha d} I_0(\alpha)^d \quad (102)$$

This value is related to the probability (denoted $\Pi(0, 1)$) that a Brownian curve returns to the origin, through

$$2d\beta_c = G(0, 1) = \frac{1}{1 - \Pi(0, 1)} \quad (103)$$

The behaviour of $2d\beta_c$ as a function of dimension d has therefore been discussed already, and for $d \rightarrow \infty$

$$2d\beta_c = 1 + \frac{1}{2d} + 3 \left(\frac{1}{2d} \right)^2 + 12 \left(\frac{1}{2d} \right)^3 + \dots \quad (104)$$

The dominant term is the mean field prediction. What is new is the appearance of a *lower critical dimension*, here equal to two, below which no transition occurs. As $d \rightarrow 2$, $\beta_c \rightarrow \infty$. More precisely

$$2d\beta_c \underset{d \rightarrow 2}{\sim} \frac{2}{\pi(d-2)} \quad (105)$$

In two dimensions, there exists only a high temperature symmetric phase. The previous calculation, carried out in a phase with a positive parameter ξ^{-2} , is understood to be in an unbroken symmetry phase, that is in the range $0 \leq \beta \leq \beta_c$. The two-dimensional result is in agreement with the Mermin-Wagner theorem (to be discussed more fully in the next chapter) according to which, for a continuous symmetry group, short-range interactions, and dimension two, there cannot be an ordered phase characterized by a nonvanishing order parameter (the case of an Abelian or commutative symmetry group is set aside, transitions can occur, but again without macroscopic magnetization).

Let us estimate the divergence of the correlation length ξ as $\beta \rightarrow \beta_c - 0$, by computing from (100) and (102)

$$\xi^2(\beta_c - \beta) = \int \frac{d^d p}{(2\pi)^d} \frac{1}{2 \sum_1^d (1 - \cos p_\mu) \left[\xi^{-2} + 2 \sum_1^d (1 - \cos p_\mu) \right]} \quad (106)$$

If $d > 4$, the right-hand side tends to a finite limit as $\xi^{-1} \rightarrow 0$, and $\beta \rightarrow \beta_c$. The sensitive region in the integral is the infrared one, where p goes to zero. The denominator behaves as $p^2(\xi^{-2} + p^2)$, and the integral converges for $d > 4$. Therefore

$$d > 4 \quad \xi \sim (\beta_c - \beta)^{-1/2} \quad \nu = \frac{1}{2} \quad (107)$$

in agreement with mean field. However, for $d \leq 4$ a new phenomenon occurs, and the exponent ν departs from $\frac{1}{2}$. As a result, $d = 4$ is the *upper critical dimension*, an expected property which will remain true for finite n . At $d = 4$, an extra logarithm appears as $\xi^2(\beta_c - \beta) \sim \ln \xi$

$$d = 4 \quad \xi \sim (\beta_c - \beta)^{-1/2} \left[\ln \left(\frac{1}{\beta_c - \beta} \right) \right]^{1/2} \quad \nu = \frac{1}{2} \quad (108)$$

In the range $4 > d > 2$, one finds

$$4 > d > 2 \quad \xi \sim (\beta_c - \beta)^{-1/(d-2)} \quad \nu = 1/(d-2) \quad (109)$$

Finally for $d = 2$, ξ diverges with β exponentially

$$d = 2 \quad \xi \underset{\beta \rightarrow \infty}{\sim} \exp 2\pi\beta \quad (110)$$

In a sense, this means $\nu = \infty$. This behaviour is characterized as asymptotic freedom. If $\beta^{-1} \sim T$ is thought of as a coupling constant, then turning equation (110) around means that β^{-1} behaves like the inverse of the logarithm of the scale.

Returning to $d > 2$, for β larger than β_c , it would seem that the correlation length is infinite for all modes (indeed the massive longitudinal one is of relative weight $1/n \rightarrow 0$). A more careful analysis is required to extract the spontaneous magnetization, since the translational invariant solution is no longer an isolated saddle point. We shall use a slightly different approach below.

The variation of ν as a function of dimension is shown on figure 5, with $\nu = 1$ for $d = 3$. It might be remarked here that a natural extrapolation in d (see for instance equation (102)) is allowed as a result of the structure of the various integral representations, and that the upper critical dimension is signalled by the appearance of logarithms. These ideas will be formalized later on.

The specific heat is discontinuous above four dimensions as in mean field, so we might say that $\alpha = 0$. For $2 < d < 4$, the

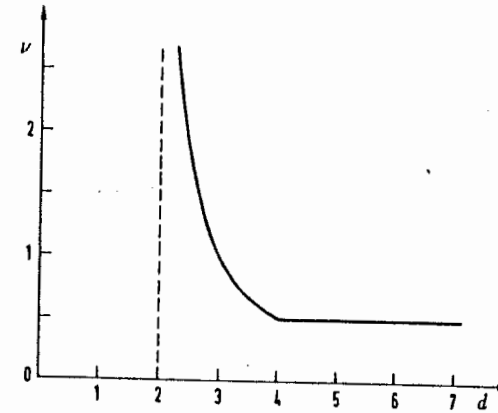


Fig. 5 The exponent ν as a function of dimension d , according to the spherical model.

singular part of the free energy is proportional to ξ^{-d} , hence the singular part of the specific heat is

$$4 > d > 2 \quad C_{\text{sing}} \sim (\beta_c - \beta)^{-(d-4)/(d-2)} \quad \alpha = \frac{d-4}{d-2} \quad (111)$$

Note that α is negative corresponding to a cusp in C . In particular for $d = 3$, $\alpha = -1$. When $d = 4$, there is a logarithmic singularity

$$d = 4 \quad C_{\text{sing}} \sim \frac{1}{\ln[1/(\beta_c - \beta)]} \quad (112)$$

3.3.2 Factorization

The study of the low temperature phase requires some care. As a matter of illustration, we shall use the factorization method of products of invariant observables in the large n limit, already encountered in chapter 1. For the computation of the spontaneous magnetization for instance, we look at the behaviour of the two point function at infinite separation. But first we derive a so-called Ward identity, as part of the equations of motion. This is an exact relation between correlation functions which follows from the structure of the partition function as an integral over

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1989

First published 1989

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Itzykson, Claude.

Statistical field theory / Claude Itzykson, Jean-Michel Drouffe.
p. cm. — (Cambridge monographs on mathematical physics)
Includes bibliographies and indexes.

Contents: v. 1. From Brownian motion to renormalization and
lattice gauge theory — v. 2. Strong coupling, Monte Carlo methods,
conformal field theory, and random systems.

ISBN 0-521-34058-6 (v. 1). ISBN 0-521-37012-4 (v. 2)

1. Statistical physics. 2. Field theory (Physics) I. Drouffe,
Jean-Michel. II. Title. III. Series.

QC174.8.I89 1989

530.1'3—dc19

88-30162

CIP

British Library Cataloging in Publication applied for.

Contents

<i>Contents of Volume 1</i>	<i>ix</i>
<i>Preface</i>	<i>xi</i>
7 Diagrammatic methods	405
7.1 General Techniques	405
7.1.1 Definitions and notations	405
7.1.2 Connected graphs and cumulants	411
7.1.3 Irreducibility and Legendre transformation	415
7.2 Series expansions	421
7.2.1 High temperature expansion	421
7.2.2 The role of symmetries	427
7.2.3 Low temperature expansion — discrete case	430
7.2.4 Low temperature expansion — continuous case	434
7.2.5 Strong field expansions	437
7.2.6 Fermionic fields	438
7.3 Enumeration of graphs	439
7.3.1 Configuration numbers with exclusion constraint	439
7.3.2 Multiply connected graphs	442
7.4 Results and analysis	446
7.4.1 Series analysis	446
7.4.2 An example: the Ising series on a body centered cubic lattice	450
Notes	454
8 Numerical simulations	456
8.1 Algorithms	456
8.1.1 Generalities	456
8.1.2 The classical algorithms	460
8.1.3 Microcanonical simulations	465
8.1.4 Practical considerations	466

8.2	Extraction of results in a simulation	475
8.2.1	Determination of transitions	475
8.2.2	Finite size effects	478
8.2.3	Monte Carlo renormalization group	481
8.2.4	Dynamics and the Langevin equation	486
8.3	Simulating fermions	489
8.3.1	The quenched approximation	491
8.3.2	Dynamical fermions	493
8.3.3	Hadron mass calculation in lattice gauge theory	494
	Notes	499
9	Conformal invariance	501
9.1	Energy-momentum tensor – Virasoro algebra	502
9.1.1	Conformal invariance	502
9.1.2	Energy-momentum tensor	507
9.1.3	Two-dimensional conformal transformations	510
9.1.4	Central charge	515
9.1.5	Virasoro algebra	521
9.1.6	The Kac determinant	532
9.1.7	Unitary and minimal representations	543
9.1.8	Characters of the Virasoro algebra	547
9.2	Examples	549
9.2.1	Gaussian model	550
9.2.2	Ising model	554
9.2.3	Three state Potts model	557
9.3	Finite size effects and modular invariance	564
9.3.1	Partition functions on a torus	564
9.3.2	Kronecker's limit formula	567
9.3.3	Ising model	573
9.3.4	The $A-D-E$ classification of minimal models	579
9.3.5	Frustrations and discrete symmetries	588
9.3.6	Nonminimal models	592
9.3.7	Correlations in a half plane	599
9.3.8	The vicinity of the critical point	603
9.A	Jacobian θ -series and products	609
9.B	Superconformal algebra	613
9.C	Current algebra	619
9.C.1	Simple Lie algebras	619

9.C.2	The Wess–Zumino–Witten model	626
9.C.3	Representations and characters of Kac–Moody algebras	639
	Notes	642
10	Disordered systems and fermionic methods	646
10.1	One-dimensional models	647
10.1.1	Gaussian random potential	647
10.1.2	Fokker–Planck equation	649
10.1.3	The replica trick	658
10.1.4	Random one-dimensional lattice	664
10.2	Two-dimensional electron gas in a strong field	675
10.2.1	Landau levels – Quantum Hall effect	675
10.2.2	One particle spectrum in the presence of impurities	679
10.3	Random matrices	690
10.3.1	Semicircle law	690
10.3.2	The fermionic method	694
10.3.3	Level spacings	696
10.4	The planar approximation	703
10.4.1	Combinatorics	704
10.4.2	The planar approximation in quantum mechanics	712
10.5	Spin systems with random interactions	715
10.5.1	Random external field and dimensional transmutation	716
10.5.2	The two-dimensional Ising model with random bonds	719
10.A	The Hall conductance as a topological invariant	729
	Notes	735
11	Random geometry	738
11.1	Random lattices	739
11.1.1	Poissonian lattices and cell statistics	739
11.1.2	Field equations	755
11.1.3	The spectrum of the Laplacian	760
11.2	Random surfaces	764
11.2.1	Piecewise linear surfaces	765
11.2.2	The conformal anomaly and the Liouville action	773

11.2.3 Sums over smooth surfaces	780
11.2.4 Discretized models	799
Notes	809
<i>Index</i>	

Contents of Volume 1

1	From Brownian motion to Euclidean fields	1
2	Grassmannian integrals and the two-dimensional Ising model	48
3	Spontaneous symmetry breaking, mean field	107
4	Scaling transformations and the XY-model	162
5	Continuous field theory and the renormalization group	233
6	Lattice gauge fields	328

Preface

Some ten years ago, when completing with J.-B. Zuber a previous text on *Quantum Field Theory*, the senior author was painfully aware that little mention was made that methods in statistical physics and Euclidean field theory were coming closer and closer, with common tools based on the use of path integrals and the renormalization group giving insights on global structures. It was partly to fill this gap that the present book was undertaken. Alas, over the five years that it took to come to life, both subjects have undergone a new evolution. Disordered media, growth patterns, complex dynamical systems or spin glasses are among the new important topics in statistical mechanics, while superstring theory has turned to the study of extended systems, Kaluza-Klein theories in higher dimensions, anticommuting coordinates ... in an attempt to formulate a unified model including all known interactions. New and sophisticated techniques have invaded statistical physics, ranging from algebraic methods in integrable systems to fractal sets or random surfaces. Powerful computers or special devices provide "experimental" means for a new brand of theoretical physicists. In quantum field theory, applications of differential topology, geometry, Riemannian manifolds, operator theory ... require a deeper background in mathematics and a knowledge of some of its most recent developments. As a result, when surveying what has been included in the present volume in an attempt to uncover the basic unity of these subjects, the authors have the same unsatisfactory feeling of not being able to bring the reader really up to date. It is presumably the fate of such endeavours to always come short of accomplishing their purpose.

With these shortcomings fully admitted, we have tried to present to the reader an overview of the main themes which justify the title "Statistical field theory." This interpretation of Euclidean field theory offers a new language, effective computing means, as

well as a natural and consistent short-distance cutoff. In other words, it allows one to give a global meaning to path integrals, to discover possible anomalies arising from integration measures, or to understand in simple terms systems with redundant variables such as gauge models. The theory of continuous phase transitions provides a bridge between probabilistic mechanics and continuous field theory, using the renormalization group to filter out relevant operators and interactions. Many authors contributed to these views, culminating in the work of K. Wilson and his collaborators and followers, which promoted the renormalization group as a universal tool to analyse the large-distance behaviour. It still retains its value, while new developments take place, particularly with conformal, or local scale invariance coming to prominence in the study of two-dimensional systems.

The content of this book is naturally divided into two parts. The first six chapters describe in succession Brownian motion, its anti-commutative counterpart in the guise of Onsager's solution to the two-dimensional Ising model, the mean-field or Landau approximation, scaling ideas exemplified by the Kosterlitz-Thouless theory for the XY-transition, the continuous renormalization group applied to the standard φ^4 theory, the simplest typical case, and lattice gauge theory as an attempt to understand quark confinement in chromodynamics.

The next five chapters (in volume 2) cover more diverse subjects. We give an introduction to strong coupling expansions and various means of analyzing them. We then briefly introduce Monte Carlo simulations with an emphasis on the applications to gauge theories. Next we turn to the significant advances in two-dimensional conformal field theory, with a lengthy presentation of the methods as well as early results. A chapter on simple disordered systems includes sample applications of fermionic techniques with no pretence at completeness. The final chapter is devoted to random geometry and an introduction to the Polyakov model of random surfaces which illustrates the relations between string theory and statistical physics.

At the price of being perhaps a bit repetitive, we have tried in the first part to introduce the subject in an elementary fashion. It is, however, assumed that the reader has some familiarity with thermodynamics as well as with quantum field theory. We often switch from one to the other interpretation, assuming that it will

not be disturbing once it is realized that the exponential of the action plays the role of the Boltzmann-Gibbs statistical weight. The last chapters cover subjects still in fast evolution.

Many important subjects could unfortunately not be covered. In random order they include dynamical critical phenomena, renormalization of σ -models or non-Abelian gauge fields except for a mention of lowest order results, topological aspects, classical solutions, instantons, monopoles, anomalies (except for the conformal one). Integrable systems are missing apart from the two-dimensional Ising model. Quantum gravity *à la Regge* is only mentioned. The list could, of course, be made much longer. Our involvement in some of the topics has certainly produced obvious biases and overemphases in certain instances. We have tried, as much as possible, to correct for these defects as well as for the numerous omissions by including at the end of each chapter a section entitled "Notes." Here we quote our sources, original articles, reviews, books and complementary material. These notes are purposely scattered through the volume, as we are sure that our quotations are very incomplete. A fair bibliography in such a large domain is beyond human capacities. Should any one feel that his or her work has not been reported or not properly mentioned, he or she is certainly right and we present our most sincere apologies. On the other hand we did not hesitate to use and sometimes follow very closely some articles or reviews which served our purpose. For instance chapter 5 is built around the definitive contributions of E. Brézin, J.-C. Le Guillou, J. Zinn-Justin and G. Parisi. Except for some further elaboration by the authors themselves, it was futile to try to improve on their work. Further examples are mentioned in the notes. It is the very nature of a survey such as this one to be inspired largely by other people's works. We hope that we did not distort or caricature them.

A book might give the illusion, especially to students, that some knowledge has become definitive and that the authors understand every part of it. This is a completely false view. No one can really fully master even his own subject, and this is luckily a source of progress. It is in the process of learning, of objecting, of finding misprints and errors, in rediscovering for oneself, that one gets the real benefits. It is very likely that, in spite of our care, many errors have crept in here and there. We welcome gladly comments and criticisms.

It was very hard to keep uniform notation throughout the text, even sometimes in the same chapter. This is a standard difficulty, especially when traditional notation in a given domain comes into conflict with those used in another one, and a compromise is necessary. We hope that this will not be a source of confusion for the reader.

We have added appendices which generally gather material in very concise form. They should be supplemented by further reading. For instance appendix C of chapter 9 is obviously insufficient to describe finite and infinite Lie algebras and their representations. This appendix is, rather, meant to induce the interested reader to study the subject further. This is also the nature of several sections where the degree of mathematical sophistication seems to increase beyond the standard background, reflecting recent trends. It was felt difficult to omit these developments but also impossible to give a proper complete introduction.

Included in small type here and there are comments, exercises and short complements ... It was felt inappropriate to develop a scholarly set of problems. In this respect the whole text can be read as a problem book.

One of the "heroes" of the whole subject of statistical physics, in one guise or another, is still to this day our old friend the Ising model. We keep a few bottles of good old French wine for the lucky person who solves it in three dimensions. It would seem appropriate to create in the theoretical physics community a prize for its solution, analogous to the one founded at the beginning of the century for the proof of Fermat's theorem. Both subjects have a similar flavour, being elementary to formulate. While it is to be presumed that the answer itself is to a large extent inessential, they motivated creative efforts (and still do) which go largely beyond the goal of solving the problem itself.

Among the many books which either overlap or amply complement the present one, the foremost are of course those in the series edited by C. Domb and M.S. Green and now J. Lebowitz, entitled *Phase transitions and critical phenomena* and published through the years by Academic Press (New York). We freely refer to this series in the notes. Let us also quote here a few others, again without pretence at exhaustivity. On the statistical side, K. Huang, *Statistical mechanics*, Wiley, New York (1963); H.E. Stanley, *In-*

troduction to phase transitions and critical phenomena, Oxford University Press (1971); S.K. Ma, *Modern theory of critical phenomena*, Benjamin, New York (1976) and *Statistical mechanics*, World Scientific, Singapore (1985); D.J. Amit, *Field theory, the renormalization group and critical phenomena*, 2nd edition, World Scientific, Singapore (1984).

Books on the path integral approach to field theory are by now numerous. Among them, the classical one is R.P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw Hill, New York (1965). Further aspects are covered in C. Itzykson and J.-B. Zuber, *Quantum field theory*, McGraw Hill, New York (1980); P. Ramond, *Field theory, a modern primer*, Benjamin/Cummings, Reading, Mass. (1981); J. Glimm and A. Jaffe, *Quantum physics*, Springer, New York (1981). To fill some gaps on other developments in field theory, see S. Coleman, *Aspects of symmetries*, Cambridge University Press (1985); S. Treiman, R. Jackiw, B. Zumino, E. Witten *Current algebra and anomalies*, World Scientific, Singapore (1985), and to learn about integrable systems, R. Baxter *Exactly solved models in statistical mechanics*, Academic Press, New York (1982), and M. Gaudin *La fonction d'onde de Bethe*, Masson, Paris (1983). Of course, many more books are mentioned in the notes. We are also aware that several important texts are either in preparation or will appear in the near future.

Our knowledge of English remains to this day very primitive and we apologize for our cumbersome use of a foreign language. This lack of fluency has prevented us of any attempt at humour which would have been sometimes more than welcome.

We would have never undertaken writing, were it not for the teaching opportunities that we were given by several universities and schools. One of the authors (C.I.) is grateful to his colleagues from the "Troisième cycle de Suisse Romande" in Lausanne and Geneva, from the "Département de Physique de l'Université de Louvain La Neuve" and from the "Troisième cycle de physique théorique" in Marseille for giving him the possibility to teach what became parts of this text, as well as to the staff of these institutions for providing secretarial help in preparing a French unpublished manuscript. The other author (J.M.D.) acknowledges similar opportunities afforded by the "Troisième cycle de physique théorique" in Paris.

The final and certainly most pleasant duty is, of course, to thank all those, friends, colleagues, collaborators, students and secretaries who have helped us through the years. A complete list should include all the members of the Saclay "Service de physique théorique", together with its numerous visitors and the members of the many departments, institutions and meetings which offered us generous hospitality and stimulation.

Particular thanks go to our very long time friends and colleagues R. Balian, M. Bander, M. Bauer, D. Bessis, E. Brézin, A. Cappelli, A. Coste, F. David, J. des Cloizeaux, C. De Dominicis, E. Gardner, M. Gaudin, B. Derrida, J.-M. Luck, A. Morel, P. Moussa, H. Orland, G. Parisi, Y. Pomeau, R. Lacaze, H. Saleur, R. Stora, J. Zinn-Justin, and J.-B. Zuber for friendly collaborations, endless discussions and generous advice. The final form of the manuscript owes a great deal to Dany Bunel. Let her receive here our warmest thanks for her tireless help. We are also very grateful to M. Porneuf and to the documentation staff, M. Féron, J. Delouvrier and F. Chétivaux.

Last but not least, we thank the Commissariat à l'Energie Atomique, the Institut de Recherche Fondamentale and the Service de Physique Théorique for their support.

Saclay, 1988

7

DIAGRAMMATIC METHODS

This chapter is devoted to technicalities related to various expansions already encountered in volume 1, mostly those that derive from the original lattice formulation of the models, be it high or low temperature, strong coupling expansions and to some extent those arising in the guise of Feynman diagrams in the continuous framework. We shall not try to be exhaustive, but rather illustrative, relying on the reader's interest to investigate in greater depth some aspects inadequately treated. Nor shall we try to explore with great sophistication the vast domain of graph theory. There are, however, a number of common features, mostly of topological nature, which we would like to present as examples of the diversity of what looks at first sight like straightforward procedures.

7.1 General Techniques

7.1.1 Definitions and notations

Let a *labelled graph* $\mathcal{G}$ be a collection of v elements from a set of indices, and l pairs of such elements with possible duplications (i.e. multiple links). We shall also interchangeably use the word *diagram* instead of graph. This abstract object is represented by v points (vertices) and l links. Each vertex is labelled by its index.

The problem under consideration will define a set of *admissible graphs*, with a corresponding *weight* $\omega(\mathcal{G})$ (a real or complex number) according to a well-defined set of rules. We wish to find the sum of weights over all admissible graphs.

Possible constraints on the graphs may be

- i) the *exclusion* constraint, preventing two vertices from carrying the same index

The final and certainly most pleasant duty is, of course, to thank all those, friends, colleagues, collaborators, students and secretaries who have helped us through the years. A complete list should include all the members of the Saclay "Service de physique théorique", together with its numerous visitors and the members of the many departments, institutions and meetings which offered us generous hospitality and stimulation.

Particular thanks go to our very long time friends and colleagues R. Balian, M. Bander, M. Bauer, D. Bessis, E. Brézin, A. Cappelli, A. Coste, F. David, J. des Cloizeaux, C. De Dominicis, E. Gardner, M. Gaudin, B. Derrida, J.-M. Luck, A. Morel, P. Moussa, H. Orland, G. Parisi, Y. Pomeau, R. Lacaze, H. Saleur, R. Stora, J. Zinn-Justin, and J.-B. Zuber for friendly collaborations, endless discussions and generous advice. The final form of the manuscript owes a great deal to Dany Bunel. Let her receive here our warmest thanks for her tireless help. We are also very grateful to M. Porneuf and to the documentation staff, M. Féron, J. Delouvrier and F. Chétivaux.

Last but not least, we thank the Commissariat à l'Energie Atomique, the Institut de Recherche Fondamentale and the Service de Physique Théorique for their support.

Saclay, 1988

7

DIAGRAMMATIC METHODS

This chapter is devoted to technicalities related to various expansions already encountered in volume 1, mostly those that derive from the original lattice formulation of the models, be it high or low temperature, strong coupling expansions and to some extent those arising in the guise of Feynman diagrams in the continuous framework. We shall not try to be exhaustive, but rather illustrative, relying on the reader's interest to investigate in greater depth some aspects inadequately treated. Nor shall we try to explore with great sophistication the vast domain of graph theory. There are, however, a number of common features, mostly of topological nature, which we would like to present as examples of the diversity of what looks at first sight like straightforward procedures.

7.1 General Techniques

7.1.1 Definitions and notations

Let a *labelled graph* $\mathcal{G}$ be a collection of v elements from a set of indices, and l pairs of such elements with possible duplications (i.e. multiple links). We shall also interchangeably use the word *diagram* instead of graph. This abstract object is represented by v points (vertices) and l links. Each vertex is labelled by its index.

The problem under consideration will define a set of *admissible graphs*, with a corresponding *weight* $\omega(\mathcal{G})$ (a real or complex number) according to a well-defined set of rules. We wish to find the sum of weights over all admissible graphs.

Possible constraints on the graphs may be

- i) the *exclusion* constraint, preventing two vertices from carrying the same index

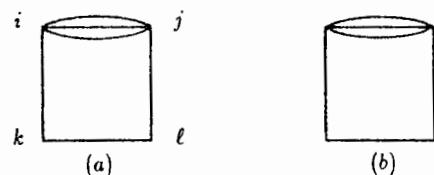


Fig. 1 (a) a labelled graph, (b) the corresponding free graph.

- ii) *simplicity* when two vertices are joined by at most one link (the graph in figure 1(a) is not simple).

Take for instance the straightforward high temperature expansion of the Ising partition function

$$Z = 2^{-N} \sum_{\{\sigma_i = \pm 1\}} \exp \left(\beta \sum_{\langle ij \rangle} \sigma_i \sigma_j \right) \quad (1)$$

$$= 2^{-N} \sum_{\{\sigma_i = \pm 1\}} \sum_{\langle ij \rangle} \left(1 + \sum_{n_{ij}=1}^{\infty} \frac{\beta^{n_{ij}}}{n_{ij}!} \sigma_i \sigma_j \right)$$

Expanding the products, keeping terms with a finite power of β , and averaging over $\sigma_i = \pm 1$, leads to a straightforward high temperature series encountered in volume 1. The successive contributions are associated with graphs defined as follows. A graph has n_{ij} lines joining vertices i and j . Isolated points are not represented as vertices. Since only even powers of σ_i have a nonvanishing unit average, admissible graphs have to obey the following three rules

- i) a line can only join vertices indexed by neighbouring sites, and we may think of the graph as drawn on the lattice,
- ii) an even number of links are incident on a vertex,
- iii) two vertices have distinct labels (the exclusion constraint).

Given an admissible graph, its weight is obtained by associating a factor β to each line, and dividing by the product $\prod_{\langle ij \rangle} n_{ij}!$ i.e. the order of the symmetry group of the graph under permutation of equivalent links.

We can also write

$$Z = (\cosh \beta)^{Nd} \frac{1}{2^N} \sum_{\{\sigma_i = \pm 1\}} \prod_{\langle ij \rangle} (1 + \sigma_i \sigma_j \tanh \beta) \quad (2)$$

which leads for $Z/(\cosh \beta d)^N$ to a different expansion. Admissible graphs are simple with a factor $\tanh \beta$ for each link. Both series are useful in applications.

Two graphs are *isomorphic* when a one-to-one correspondence can be set among vertices and links preserving the incidence relations. The difference lies therefore in the labels of the vertices. Isomorphism leads to equivalence classes called *free graphs* and denoted G . In a pictorial representation, vertices do not carry indices anymore (fig. 1(b)). Conventionally, the corresponding weight $\omega(G)$ will be the average over the equivalent labelled graphs. Call number of configurations $n(G)$ the cardinal of the equivalence class, then

$$\sum_{G \in \mathcal{G}} \omega(G) = n(G) \omega(G) \quad (3)$$

This definition is useful whenever the weight of a graph is independent of the labelling of its vertices. In any case, it allows one to disentangle the part $\omega(G)$ that is specific to the model, from the geometry of the lattice, which yields $n(G)$. The following two sections will treat these problems separately.

The above definitions can be extended in various ways.

- i) Vertices may be of several types.
- ii) Links may have to be oriented.
- iii) A generalization may be envisioned, where instead of dealing with 0 and 1 dimensional simplices (vertices and links), one may be required to consider higher dimensional elements (two dimensional plaquettes in the gauge case).
- iv) Indices may be compound ones, and links may have to carry indices at their extremities.

This list is of course just indicative of possible extensions.

In some applications, the computation of correlations for instance, a subset of vertices carries fixed indices. Equivalence classes of such graphs will be called *rooted graphs*.

Two vertices x and y on G are *linked* if they can be joined by a path along links of the graph $xz_1, z_1z_2, \dots, z_ny$. This provides again an equivalence relation on vertices, and the corresponding classes

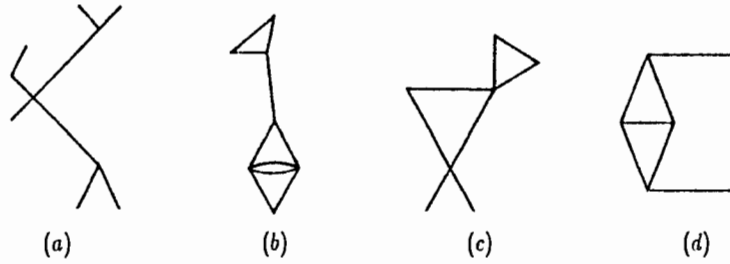


Fig. 2 (a) a tree (b) a graph with four loops (c) a graph with two articulation points (d) a multiply connected graph.

define the connected disjoint parts of the graph. A *connected* graph has a unique connected part.

A *cycle* is a closed path of n links, and n vertices, all distinct, starting and ending at the same vertex. A connected graph without cycles is a *tree* (figure 2(a)). The *number of loops* in a connected graph is the minimum number of links which, when removed, leave a tree (figure 2(b)).

An *articulation point* (figure 2(c)) is such that its omission, together with incident links, increases the number of connected parts. It is therefore a vertex which appears on any path linking certain pairs of vertices. In particular, on a tree, all vertices but the external ones (joined to the graph by only one link) are articulation points. A connected graph without articulation points is a *multiply connected* graph: any two vertices belong to a cycle and can therefore be linked by at least two totally distinct paths.

In terms of the following notation

v_k , number of vertices with k incident links

$v = \sum_k v_k$, total number of vertices

l , number of links

b , number of loops

c , number of connected parts

we have the relation

$$2l = \sum_k k v_k, \quad (4)$$

expressing that each link joins two vertices, thus twice the number of links is equal to the sum over vertices weighted by the number

of incident links. On the other hand, Euler's relation

$$v + b = c + l \quad (5)$$

follows by successively suppressing the links until only v isolated vertices are left. At each stage, the number of loops decreases or the number of connected parts increases by unity.

- i) Compute up to fourth order the Ising partition function on a d -dimensional hypercubic lattice.

Admissible graphs with at most four links are shown on figure 3. The corresponding numbers of configurations on a finite periodic lattice of N sites are respectively Nd , Nd , $Nd(2d-1)$, $\frac{1}{2}Nd(d-1)$ and $\frac{1}{2}Nd(Nd-4d+1)$. Their weights, taking into account the symmetry factor, are respectively $\frac{1}{2}\beta^2$, $\frac{1}{24}\beta^4$, $\frac{1}{4}\beta^4$, $\frac{1}{4}\beta^4$. Summing these contributions yields

$$Z = 1 + \frac{1}{2}Nd\beta^2 + \left[\frac{1}{12}Nd(6d-7) + \frac{1}{8}N^2d^2\right]\beta^4 + \mathcal{O}(\beta^6) \quad (6)$$

To this order, one verifies the extensive character of the free energy. Indeed

$$-\frac{\beta\mathcal{F}}{N} = \frac{F}{N} = \frac{\ln Z}{N} = \frac{1}{2}d\beta^2 + \frac{1}{12}d(6d-7)\beta^4 + \mathcal{O}(\beta^6) \quad (7)$$

is N -independent. The notation $\mathcal{F}$ refers to the traditional thermodynamic definition.

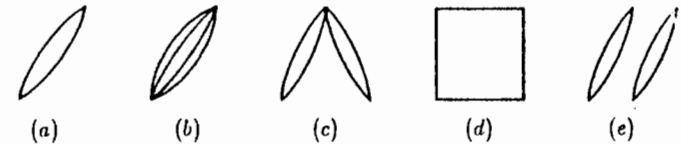


Fig. 3 Graphs for the Ising model up to order β^4 .

This example illustrates the convenience of using $\tanh\beta$ rather than β as a small parameter in a high temperature expansion. Indeed, the corresponding graphs have now to satisfy the constraint of simplicity. Only the graph of figure 3(d) leads to a nonvanishing contribution in an expansion up to order $\tanh^4\beta$, and we find

$$Z = (\cosh\beta)^{Nd} \left[1 + \frac{1}{2}Nd(d-1)\tanh^4\beta + \mathcal{O}(\tanh^6\beta)\right] \quad (8)$$

Equations (6) and (8) are easily shown to agree.

This simplification is useful in avoiding the proliferation of graphs, when investigating the higher order contributions. Other techniques will be discussed below.

- ii) *Kirchoff's theorem.* The previous definitions allow us to recall a theorem due to Kirchoff giving in closed form the number of distinct trees that can be drawn on a connected graph, which uses the same set of vertices, the so called *spanning trees*. To a connected graph G we associate the incidence matrix which is a topological equivalent of (minus) the Laplacian. Along the main diagonal, $(-\Delta)_{ii}$ is equal to the number of links incident on vertex i , while $(-\Delta)_{ij}$ for distinct i and j is minus the number of links joining the vertices i and j . Elements in each line or column of $-\Delta$ add to zero, and therefore $\det(-\Delta)$ is zero, corresponding to the existence of a unique zero mode, a constant, since the graph is connected. The claim is that the determinant of any principal minor $((-1)^{i+j}$ times the determinant obtained by deleting the i th line and j th column) is equal to the number of spanning trees. From the fact that the graph is connected, there exists a unique vector (the zero mode) with equal components corresponding to the zero eigenvalue of $-\Delta$ up to an overall factor. Let M_{ij} be the principal minor of the element ij (including the sign). From

$$\sum_j (-\Delta)_{ij} M_{kj} = \delta_{ik} \det(-\Delta) = 0$$

it follows that, for fixed k , all M_{kj} are equal, and since the matrix is symmetric as is $-\Delta$, all M_{kj} are equal to the same value M . It is therefore sufficient to compute $M = M_{11}$, the determinant of the matrix $-\Delta$ with the first line and column deleted. Let v denote the number of vertices, and $\ell \geq v - 1$ the number of links. Define a $\ell \times v$ matrix $L_{\alpha i}$ where α labels links, i vertices, after giving to each link an arbitrary orientation, through

$$L_{\alpha i} = \begin{cases} +1 & \text{if the link } \alpha, \text{ incident on } i \text{ is oriented off } i \\ -1 & \text{if the link } \alpha, \text{ incident on } i \text{ is oriented towards } i \\ 0 & \text{if the link } \alpha \text{ is not incident on } i \end{cases}$$

Then $(-\Delta) = L^T L$. Call L' the matrix L with its first column deleted, in such a way that

$$M = \det L'^T L' = \sum_{(\alpha_2 < \dots < \alpha_v)} \det L'_{\alpha_2 \dots \alpha_v}^T L'_{\alpha_2 \dots \alpha_v}$$

with the sum running over all $(v-1) \times (v-1)$ matrices $L'_{\alpha_2 \dots \alpha_v}$ obtained by selecting $v-1$ rows in L' labelled $\alpha_2 < \alpha_3 < \dots < \alpha_v$. The above equality is a classical identity in the theory of determinants. Each term in the sum is of the form $(\det L'_{\alpha_2 \dots \alpha_v})^2$, and is zero unless the map $i \rightarrow \alpha_i$ assigns to each vertex $i = 2, \dots, v$ an incident link, in which case the matrix $L'_{\alpha_2 \dots \alpha_v}$ differs from a permutation matrix only by the fact that its entries are ± 1 instead of $+1$. This has two consequences. The first is that $(\det L'_{\alpha_2 \dots \alpha_v})^2$ is equal to unity, and the second that it is in one-to-one correspondence with a spanning tree. This proves Kirchoff's theorem. It gives a topological meaning to the Laplacian which proves useful in percolation and polymer problems. We shall encounter an application in chapter 11.

7.1.2 Connected graphs and cumulants

The fundamental property of exponentiation relies on the following conditions.

- i) The empty graph (no vertex, no link) is admissible, with a weight equal to 1. It has no connected parts ($c = 0$) and is therefore not connected ($c \neq 1$).
- ii) Every union of admissible graphs is admissible.
- iii) The weight of a graph factors into contributions from its connected parts.

If these hypotheses are valid, the sum over all admissible graphs is equal to the exponential of the sum over connected graphs. The exclusion constraint is not compatible with condition (ii). The reader will check on the above example of the β -expansion up to β^4 for the Ising model, that the preceding property is wrong, namely the free energy is not directly given in terms of contributions from connected graphs. It is our present task to modify the rules in order to find a direct expansion for the free energy.

The proof is quite simple. An arbitrary graph is built by drawing successively and independently its c connected parts $\mathcal{G}_1, \dots, \mathcal{G}_c$. The order being immaterial, each disconnected graph is therefore obtained $c!$ times instead of one. Using factorization of weights and summing on c , one finds

$$\sum_{\mathcal{G}} \omega(\mathcal{G}) = \sum_{c=0}^{\infty} \frac{1}{c!} \sum_{\substack{\mathcal{G}_1, \dots, \mathcal{G}_c \\ \text{connected}}} \omega(\mathcal{G}_1) \cdots \omega(\mathcal{G}_c) \quad (9)$$

On the right-hand side, one recognizes the expansion of an exponential, namely

$$\sum_{\mathcal{G}} \omega(\mathcal{G}) = \exp \left(\sum_{\mathcal{G} \text{ connected}} \omega(\mathcal{G}) \right) \quad (10)$$

The above exponential property is crucial in the computation of extensive quantities as well as in the study of various asymptotic behaviours, correlation lengths, or boundary effects. Unfortunately the exclusion rule forbids its direct application. We will now modify these rules, using the *cumulant* method, to restore exponentiation.

Assume that another set of rules can be found avoiding the exclusion constraint. To distinguish them, we represent vertices of the new graphs by open instead of full circles. A contribution of a new graph represents part of the former ones, obtained by identifying vertices with the same labels. If one requires that the new expansion reproduces the previous results, one obtains a set of equations, written symbolically as

$$\begin{aligned} \frac{1}{1!} \circ + \frac{1}{2!} \circ \circ + \frac{1}{3!} \circ \circ \circ + \dots &= \bullet \\ (1 + \bullet)(\circ -) &= - \\ (1 + \bullet)(-\circ + -\circ -) &= - \\ (1 + \bullet)(\text{Y} + \text{Y} + \text{Y} + \text{Y} + \text{Y}) &= \text{Y} \\ (1 + \bullet)(\text{X} + \sum_{4 \text{ terms}} \text{X} + \sum_{3 \text{ terms}} \text{X} + \sum_{6 \text{ terms}} \text{X} + \dots) &= \text{X} \end{aligned} \quad (11)$$

The factor $(1 + \bullet)$ follows from the possibility of identifying as many isolated vertices as one may wish. In this relation, all vertices carry the same index, which has been omitted for clarity.

To solve these equations, one must of course define the associated weights. Take the Ising case as a typical example. In the standard expansion which enforces the exclusion constraint, we have

i) a factor z_k for a vertex with k incident links

- ii) a factor β for each link
- iii) the product of these factors is divided by the order of the symmetry group of the graph.

The expansion without the exclusion constraint follows from similar rules, except for the fact that the factors z_k are to be replaced by cumulants u_k . The equations (11) read explicitly

$$\begin{aligned} e^{u_0} - 1 &= z_0 - 1 \\ u_1 &= z_1/z_0 \\ u_2 + u_1^2 &= z_2/z_0 \\ u_3 + 3u_2u_1 + u_1^3 &= z_3/z_0 \\ u_4 + 4u_3u_1 + 3u_2^2 + 6u_2u_1^2 + u_1^4 &= z_4/z_0 \\ &\dots \end{aligned} \quad (12)$$

The cumulants u_k are obtained by inverting this system as

$$\begin{aligned} u_0 &= \ln z_0 \\ u_1 &= \frac{z_1}{z_0} \\ u_2 &= \frac{z_2}{z_0} - \left(\frac{z_1}{z_0}\right)^2 \\ u_3 &= \frac{z_3}{z_0} - 3\frac{z_2}{z_0}\frac{z_1}{z_0} + 2\left(\frac{z_1}{z_0}\right)^3 \\ u_4 &= \frac{z_4}{z_0} - 4\frac{z_3}{z_0}\frac{z_1}{z_0} - 3\left(\frac{z_2}{z_0}\right)^2 + 12\frac{z_2}{z_0}\left(\frac{z_1}{z_0}\right)^2 - 6\left(\frac{z_1}{z_0}\right)^4 \\ &\dots \end{aligned} \quad (13)$$

To obtain a compact form, write $z(h)$ and $u(h)$ for the generating functions

$$z(h) = \sum_{k=0}^{\infty} \frac{z_k}{k!} h^k \quad (14)$$

$$u(h) = \sum_{k=0}^{\infty} \frac{u_k}{k!} h^k \quad (15)$$

The above relations are then simply

$$u(h) = \ln z(h) \quad (16)$$

- i) Justify in more detail the cumulant procedure and check equation (16) up to order four against (13).
- ii) Rederive the high temperature expansion in the Ising case using cumulants. The initial rules assume $z_{2k} = 1$, $z_{2k+1} = 0$, thus

$$z(h) = \cosh h \quad (17)$$

$$u(h) = \ln \cosh h = \frac{1}{2}h^2 - \frac{1}{12}h^4 + \dots \quad (18)$$

The connected graphs are the first four on figure 3 up to fourth order. Since the exclusion rule no longer applies, the configuration numbers are modified to Nd , Nd , $\frac{1}{2}N(2d)^2$, $\frac{3}{4}Nd(2d-1)$ respectively. These numbers are all proportional to the number of sites N , because of translational invariance, and this insures that the free energy is extensive. Since $u_2 = 1$, $u_4 = -2$, the new weights are $\frac{1}{2}\beta^2$, $\frac{1}{6}\beta^4$, $-\frac{1}{2}\beta^4$, β^4 . The weights are no longer positive, reflecting a similar loss of positivity of the cumulants. Summation over connected graphs yields immediately formula (7) and confirms the general property. The same method applies to the expansion in $\tanh \beta$. One has to introduce factors z_k which depend on the relative direction of incident links, and the graphs obtained in this way are no longer simple, which limits the interest of the method. The choice between simplicity and connectivity depends on the problem at hand. For gauge theories, we shall see that simplicity is more advantageous.

For rooted graphs, the distribution of fixed indices over connected parts leads to equations generalizing (11). Let $Z\langle i_1 \dots i_k \rangle$ (Z is the partition function) denote the sum over all graphs (connected or disconnected) with roots $i_1, \dots, i_k$, and $\langle i_1 \dots i_k \rangle_c$ the sum over connected graphs using cumulants, one has in general

$$\begin{aligned} \langle i_1 \rangle &= \langle i_1 \rangle_c \\ \langle i_1 i_2 \rangle &= \langle i_1 i_2 \rangle_c + \langle i_1 \rangle_c \langle i_2 \rangle_c \\ \langle i_1 i_2 i_3 \rangle &= \langle i_1 i_2 i_3 \rangle_c + \langle i_1 i_2 \rangle_c \langle i_3 \rangle_c + \langle i_1 i_3 \rangle_c \langle i_2 \rangle_c \\ &\quad + \langle i_2 i_3 \rangle_c \langle i_1 \rangle_c + \langle i_1 \rangle_c \langle i_2 \rangle_c \langle i_3 \rangle_c \\ &\dots \end{aligned} \quad (19)$$

Disconnected graphs factorize partly into connected parts involving no roots, the sum of which yields the factor Z , and partly into connected rooted graphs which will realize all possible partitions of the set of indices $i_1, \dots, i_k$. These properties become more

obvious when one introduces a source j . Rooted graphs appear in the expansion of correlation functions, corresponding to

$$\langle \varphi_{i_1} \dots \varphi_{i_k} \rangle = \frac{1}{Z} \frac{\partial^k Z}{\partial j_{i_1} \dots \partial j_{i_k}} \quad (20)$$

or their connected counterparts

$$\langle \varphi_{i_1} \dots \varphi_{i_k} \rangle_c = \frac{\partial^k \ln Z}{\partial j_{i_1} \dots \partial j_{i_k}} \quad (21)$$

The relations (19) follow easily as a consequence of (20) and (21).

7.1.3 Irreducibility and Legendre transformation

To decrease further the number of graphs, we now decompose the connected graphs into irreducible parts. This notion is flexible and depends to some degree on the context. It can for instance be extended to multiply connected graphs (see exercise below). A compromise has to be found in order to obtain manageable calculations. Here we shall introduce irreducibility with respect to links.

A connected graph is *reducible* if it contains at least one link, the omission of which would disconnect it. This is shown on figure 4(a), where we have indicated the links with respect to which the graph can be reduced. We observe that a tree graph can be reduced with respect to each of its links. Having suppressed all these links, the graph is replaced by a collection of irreducible parts, which range from the simple vertex to "cactus" parts built from multiply connected parts, connected at articulation points. If these blocks are represented as hatched blobs, any graph takes the form of a tree with vertices replaced by blobs. Figure 4(b) illustrates symbolically this decomposition. The left-hand side is the sum over connected graphs represented by a white blob.

Assuming the conditions listed in the previous paragraph to be valid, we will show that it is possible to give diagrammatic rules for these generalized tree graphs, and then present a method of resummation.

Let $C(\{h_i\})$ be the sum over the weights of labelled connected graphs, where the vertex labelled i , with k incident links, is assigned a factor $\partial^k u(h_i) / \partial h_i^k$. Later on we shall let the h_i 's go to zero to recover the initial graphs. The derivative operation

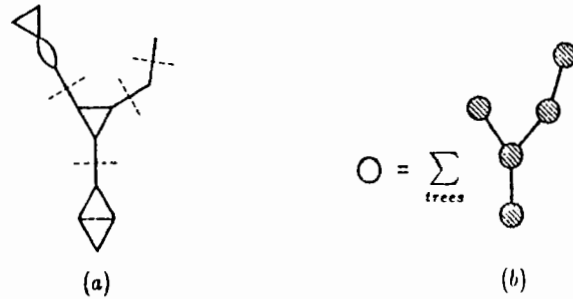


Fig. 4 (a) a reducible graph and its decomposition into irreducible parts (b) general type of connected graphs.

$\partial/\partial h_i$ acting on C amounts to choosing successively in all graphs the vertices with index i , and replacing the factor $\partial^k u(h_i)/\partial h_i^k$ by $\partial^{k+1} u(h_i)/\partial h_i^{k+1}$. Diagrammatically this means adding an external link emanating from vertex i . Symbolically

$$m_i = \frac{\partial C(\{h_i\})}{\partial h_i} = \frac{\partial}{\partial h_i} \bigcirc = \bigcirc_i \quad (22)$$

Similarly we introduce a function (or rather a functional) $I(h_i)$ corresponding to the hatched blob, i.e. the sum over irreducible graphs with similar weights $\partial^k u(h_i)/\partial h_i^k$ for the labelled vertex i . The interpretation of the derivative $\partial/\partial h_i$ as in (22) remains valid when acting on I . In the case of a uniform field $h_i = h$, the contribution of hatched blobs in the trees of figure 4(b) is given by the generating function $I(h) = I(\{h_i\})|_{h_i=h}$. These graphs are unlabelled. It is necessary to include in the diagrammatic rules a convention of dividing out by a symmetry factor to recover once, and only once, each labelled connected graph of the original expansion.

In order to find a resummation procedure for trees, we shall first generate all of them with an incorrect weight, then amend the result using the topological relations discussed in section 1.1. Let us first consider the sum over all graphs, with one vertex singled out. This contribution, denoted C_* , is obtained by weighting each tree by its number of vertices. We build the tree starting from the specified vertex, from which $k = 0, 1, 2, \dots$ links can

emerge. The generalized vertex, represented by the hatched blob, contains generally several simple vertices. It is therefore necessary to "dress" each of them individually. The operation of adding k_i lines incident on the vertex i is represented by the derivative $\partial^{k_i}/\partial h_i^{k_i}$. Each of these lines carries a factor β and has to be joined to a neighbour of the site i . In other words, the dressed vertex is obtained from the bare one by the operator

$$\prod_i \sum_{k_i=0}^{\infty} \frac{\beta^{k_i}}{k_i!} \left(\sum_{j(i)} m_j \right)^{k_i} \frac{\partial^{k_i}}{\partial h_i^{k_i}}$$

which amounts to the replacement $h_i \rightarrow h_i + \beta \sum_{j(i)} m_j$. Assuming now a uniform situation with all $h_i \equiv h$, $m_i \equiv m$, this means that the generalized vertex is (recall that q is the coordination number).

$$C_* = \bigcirc + \bigcirc_i + \bigcirc_{ii} + \bigcirc_{iii} + \dots \quad (23)$$

or, more explicitly

$$C_* = \sum_{k=0}^{\infty} \frac{d^k I(h)}{dh^k} (\beta m)^k = I(h + q\beta m). \quad (24)$$

One can perform a similar calculation if instead one singles out one link. Each graph is then weighted by its number of links. The corresponding quantity is C_l with

$$C_l = \bigcirc - \bigcirc = \sum_{links} \beta m^2 = N \frac{q}{2} m^2 \quad (25)$$

Equation (5) reads in the present case (a tree)

$$v - l = 1 \quad (26)$$

so that each graph is counted exactly once in the difference $C_* - C_l$. The final result is therefore

$$C = C_* - C_l = I(h + q\beta m) - \sum_{links} \beta m^2 \quad (27)$$

where the two terms are proportional to the total number of sites N and m is given by the uniform solution of

$$m_i = \frac{\partial C(\{h\})}{\partial h_i} \quad (28)$$

or equivalently, according to equation (27)

$$m = \frac{1}{N} \frac{\partial I(h + q\beta m)}{\partial h} \quad (29)$$

Equations (27) and (29) state that the functions C and I are Legendre transforms of each other. Equation (29) implies that in (27) the right-hand side is stationary with respect to variations in m for fixed h . This leads to the extremization procedure

$$C(h) = \text{Extr}_m \left(I(h + q\beta m) - \sum_{\text{links}} \beta m^2 \right) \quad (30)$$

If we wish, we can now set the external field equal to zero. This Legendre transformation enables one to extend the mean field procedure to all orders (in principle). Written in the above form of an extremum principle and properly interpreted, it exhibits the convex character of the effective action $C(h)$.

- i) It would seem at first sight that in the example of the Ising model in the absence of an external field, the above technique is useless, since all finite graphs are irreducible, given the fact that an even number of links are incident on each site. The magnetization per site

$$\langle \sigma \rangle = \frac{1}{N} \frac{\partial \ln Z}{\partial h} \Big|_{h=0} = m$$

vanishes indeed to all finite orders. Of course, as we know, this result is wrong at sufficiently low temperature (for $d \geq 2$). It is interesting therefore to note that formula (30), which amounts to a resummation over an infinite number of graphs, is capable of reproducing the phenomenon of spontaneous symmetry breaking by replacing the external field $h = 0$ by an effective mean field $q\beta m$. To see this, we observe that the series expressing I can be reorganized using other parameters, $1/d$ for instance. It is then possible to approximate $I(h)$ at any temperature by the contribution from its leading order graph corresponding to the lowest order graph, i.e. a vertex with a weight $\sum_i u(h_i) = \sum_i \ln \cosh h_i$. Formula (30) becomes

$$\frac{\ln Z}{N} = \text{Extr}_m (\ln \cosh 2d\beta m - \beta d m^2) \quad (31)$$

when all m_i 's are equal, with a unique maximum at $m = 0$ for small enough β . As β increases, this maximum becomes a local

minimum, and two new maxima develop symmetrically at a value corresponding to $\pm m_{sp}$, where m_{sp} is the spontaneous magnetization. Of course, the diagrammatic method is not adapted to provide a choice between extrema; the latter results from a global study, as discussed in volume 1. The value of β where the second derivative in m of the r.h.s. in (30) vanishes yields the critical temperature $\beta_c d = 1$ for a second order transition, and we recover the classical mean field expressions. The interest of the method is that it enables one to study the corrections systematically.

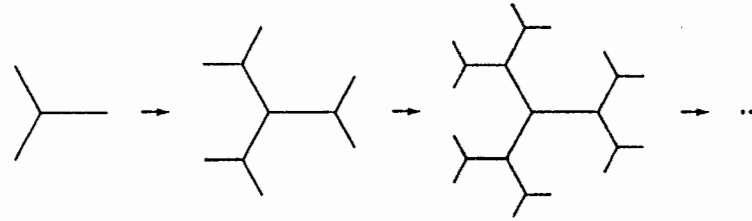


Fig. 5 Recursive construction of a Bethe lattice.

- ii) Number of trees on a Bethe lattice. A Bethe lattice is an infinite tree where each site has q neighbours constructed recursively as shown on figure 5. To be precise, a recursively constructed tree – a Cayley tree – has the unfortunate property that in the “thermodynamic” limit, the number of boundary sites is comparable to the total number of sites. As a result, surface effects cannot be ignored in general, unless one considers only those quantities pertaining to sites deep “inside” the tree (meaning that the infinite volume limit has to be taken prior to any computation). This is what is meant by a Bethe lattice. Many problems admit of an exact solution on such a lattice, devoid of loops, where one can set up recursion relations. Let us count here the number N_k of trees with k branches, per site. We shall use a generating function $G(t) = \sum_{k=1}^{\infty} N_k t^k$. These trees have of course at most q links incident on a site. The equation corresponding to (27) which expresses this fact is

$$\begin{aligned} G(t) &= \sum_{k=1}^q \binom{q}{k} (mt)^k - \frac{1}{2} q t m^2 \\ &= (1 + mt)^q - 1 - \frac{1}{2} q t m^2 \end{aligned} \quad (32)$$

The quantity m represents the contribution of trees growing out a given link

$$m = \sum_{k=0}^{q-1} \binom{q-1}{k} (mt)^k = (1+mt)^{q-1} \quad (33)$$

The solution

$$G(t) = \text{Extr}_m \left\{ (1+tm)^q - 1 - \frac{1}{2} qtm^2 \right\} \quad (34)$$

becomes singular at $t_c = (q-2)^{q-2}/(q-1)^{q-1}$. The singularity arises from the weight of large trees and is related to a percolation threshold on this lattice.

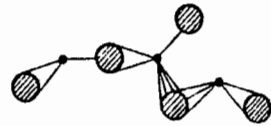


Fig. 6 Decomposition of a graph into multiply connected parts.

- iii) Reduction to multiply connected graphs. To implement such a reduction, the number of variables on which to perform a Legendre transform becomes infinite. We want to resum "cactus" graphs (figure 6) with labelled articulation points. The relevant topological relation is now

$$1 = n_b + n_s - n_m \quad (35)$$

where n_b is the number of blobs, n_s the number of articulation points and n_m the sum over articulation points weighted by the number of incident blobs. Let $B(\{u_i^{(n)}\})$ denote the contribution from multiply connected graphs, such that, to each vertex i with n incident links, is assigned a factor $u_i^{(n)}$ instead of the $\partial^n u(h_i)/\partial h_i^n$ assigned previously. To go from a blob to a cactus, we want to replace each vertex by a "dressed" one

$$u_i^{(n)} = \left[\exp \left(\sum_{k=1}^{\infty} g_i^{(k)} \frac{\partial^k}{\partial h_i^k} \right) \right] \frac{\partial^n u(h_i)}{\partial h_i^n} \quad (36)$$

where $g_i^{(k)}$ is the contribution from all connected graphs such that k links are incident on vertex i

$$g_i^{(k)} = \frac{\partial B(\{u_j^{(n)}\})}{\partial u_i^{(k)}} \quad (37)$$

As above, we use relation (35) to obtain the sum C over all connected graphs in the form

$$C = B(\{u_i^{(n)}\}) + \sum_i \left[\exp \left(\sum_{k=1}^{\infty} g_i^{(k)} \frac{\partial^k}{\partial h_i^k} \right) \right] u(h_i) - \sum_i \sum_{k=1}^{\infty} u_i^{(k)} g_i^{(k)} \quad (38)$$

Again equations (36) and (37) yield conditions under which the right-hand side, considered as a functional of the independent variables $u_i^{(n)}$ and $g_i^{(n)}$, is extremal.

Although these reduction techniques look very attractive in principle, the corresponding expressions are so cumbersome in lattice models that their practical use is very limited. However, in the continuous field theoretic case, analogous irreducible kernels are necessary in several instances. They occur in the Schwinger-Dyson integral equations, in the analysis of bound states using the Bethe-Salpeter equation, as well as in the formal renormalization theory.

7.2 Series expansions

7.2.1 High temperature expansion

Consider a general model with an action (the Boltzmann weight is $\exp S$) of the form

$$S(J) = \beta \sum_{k \geq 2} \frac{1}{k!} V_{i_1 \dots i_k}^{\alpha_1 \dots \alpha_k} \varphi_{i_1}^{\alpha_1} \dots \varphi_{i_k}^{\alpha_k} + \sum_{i, \alpha} J_i^{\alpha} \varphi_i^{\alpha} \quad (39)$$

The fields φ_i with components φ_i^{α} can be either bosonic or fermionic (in the latter case for instance the source J would also include Grassmannian variables). The *a priori* measure is denoted $\prod_i D\varphi_i$. We want to evaluate the generating function of correlations

$$Z(J) = \int \prod D\varphi_i \exp S(J) \quad (40)$$

In the high temperature expansion the integrand is expanded in a series in the variable β . Each term is then integrated over the fields φ_i . With J restricted to its value at a given site, we have to evaluate the local integrals

$$\begin{aligned} z^{\alpha_1 \dots \alpha_n}(J) &= z^{\alpha_1 \dots \alpha_n}(J) = \int \varphi^{\alpha_1} \dots \varphi^{\alpha_n} \exp J \cdot \varphi \, d\varphi \\ &= \frac{\partial^n}{\partial J^{\alpha_1} \dots \partial J^{\alpha_n}} \int e^{J \cdot \varphi} d\varphi = \frac{\partial^n z(J)}{\partial J^{\alpha_1} \dots \partial J^{\alpha_n}} \end{aligned} \quad (41a)$$

Thus we require the value of the function

$$z(J) = \int D\varphi \exp J \cdot \varphi \quad (41b)$$

associated to a site. Once this is known, let us list the rules to compute the partition function Z as an expansion in β . We consider all diagrams involving the following ingredients.

- Vertices corresponding to sites, with $n \geq 1$ branches, indexed by the site i and representing a product of fields $\varphi_i^{\alpha_1} \dots \varphi_i^{\alpha_n}$ with a contribution to the weight

$$z^{\alpha_1 \dots \alpha_n}(J_i)/z(J_i)$$

- Vertices corresponding to interactions with $k \geq 2$ branches, without an index and contributing a factor

$$\beta V_{i_1 \dots i_k}^{\alpha_1 \dots \alpha_k}$$

where $i_1 \dots i_k$ denote the site indices of those previous vertices to which this vertex is linked.

- Links joining an interaction vertex to a site vertex, with a label α , the same index α which appears at both ends. Summation over those α 's will have to be performed.

One has to sum over the contributions of all possible diagrams, connected or not, with an exclusion constraint on sites vertices. Applying the methods of the previous section, we give below some examples.

In the particular case of a homogeneous (i.e. translationally invariant) quadratic action, interaction vertices with two branches can be ignored, and one finds the typical graphs of spin models, in particular those of the Ising model used as an illustration in the previous section.

i) Establish an expansion valid in large dimension for the $O(n)$ -spin model using irreducibility.

The fields φ_i are unit n -component vectors, and the corresponding action reads $\beta \sum_{(ij)} \varphi_i \cdot \varphi_j$. The generating function of cumulants is therefore

$$z(h) \equiv \exp u(h) = \int d\varphi \exp h \cdot \varphi = \Gamma(n/2)(n/2)^{1-n/2} I_{n/2-1}(h) \quad (42)$$

Here $I_k(h)$ is the modified Bessel function, not to be confused with the generating functional of irreducible graphs $I(\{h\})$. In equation (42), the measure over φ is normalized to unity. Setting $H = h + 2d\beta m$ in relation (30), we find

$$\frac{\ln Z}{N} = I(H) - \frac{H^2}{4\beta d}, \quad H = 2\beta d \frac{dI}{dH} \quad (43)$$

As was stated before, the solution $H = 0$ ceases to be the physical one when it corresponds to a minimum of the right-hand side of (43). The location of the transition is given by requiring the coefficient of the second derivative in H at $H = 0$ to vanish. To lowest order $I(H) = u(H)$, thus, in the above units, $\beta_c = n/2d$ to this order, which vanishes like $1/d$ for large dimension. On figure 7, we show all diagrams up to order 6, as well as their second derivative in H at the origin. Since β_c behaves like d^{-1} , a diagram of order k gives a contribution behaving as $d^{-k+[k/2]}$. This is due to the fact that irreducibility forces the diagram to involve only cycles, and a cycle of length $2k$ can only use k distinct directions on a lattice. The configuration number is a polynomial in d of degree k . If the series for I is ordered in powers of βd and d^{-1} , only a finite number of diagrams will contribute to a given order in d^{-1} . Thus one derives an equation giving $\beta_c d$ as a series in d^{-1} , of the form

$$\frac{n}{2\beta_c d} = 1 - \frac{1}{2d} - \frac{2-2/(n+2)}{4d^2} - \frac{7-8/(n+2)}{8d^3} + \dots \quad (44)$$

to be compared with the large n limit (with β_c/n rescaled as β_c) given in (3.104).

ii) It is interesting to rederive in this framework the large n limit. From the above discussion, one should keep β/n finite as $n \rightarrow \infty$. Inspecting figure 7, we see that the only contributions in the limit $n \rightarrow \infty$ arise from diagrams involving a single loop. This suggests the use of the formalism introduced in exercise

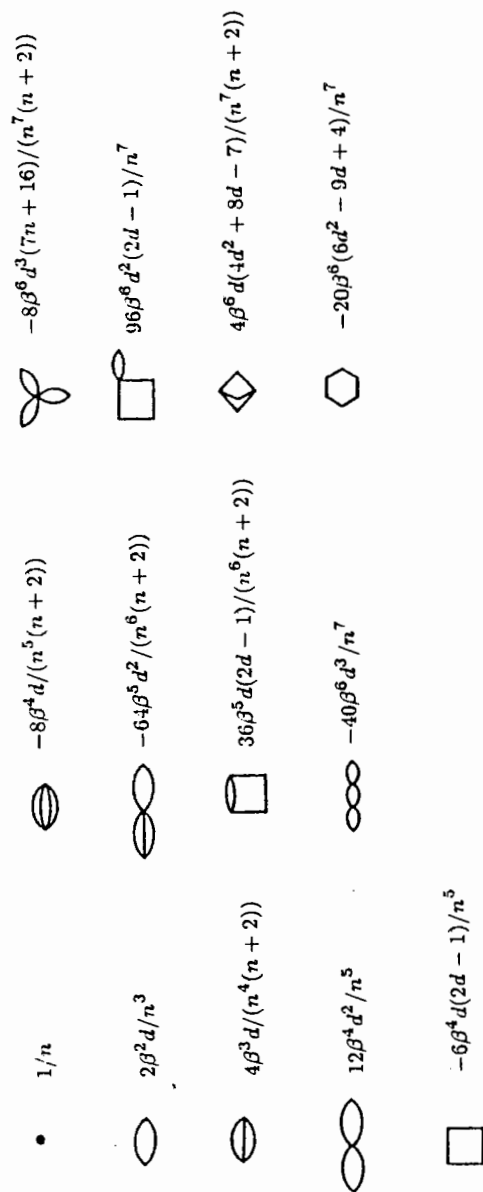


Fig. 7. Diagrams for the computation of the irreducible kernel $I(H)$ up to sixth order, and their contribution to $d^2 I(H)/dH^2|_{H=0}$.

(iii) of section 1.3. The function $u(h)$ behaves as

$$u(h) \sim \frac{1}{2}n \left\{ \sqrt{1 + 4h^2/n^2} - 1 - \ln \left(\frac{1}{2} + \frac{1}{2}\sqrt{1 + 4h^2/n^2} \right) \right\} \quad (45)$$

showing that the functions $g^{(3)}, g^{(4)} \dots$ do not appear in equation (36); only $g^{(1)}$ (a vector) and $\bar{g}^{(2)}$ (a matrix) play a role. The right-hand side of equation (36) reads

$$\exp \left(\frac{\partial}{\partial h_\alpha} \bar{g}_{\alpha\beta}^{(2)} \frac{\partial}{\partial h_\beta} \right) u^{(n)} \left(|h + g^{(1)}| \right) \simeq u^{(n)} \left(\sqrt{|h + g^{(1)}| + 2 \text{Tr} \bar{g}^{(2)}} \right) \quad (46)$$

As for the quantity B , we have only to keep the one link diagram, as well as those involving a single loop. It is shown in section 3 how to compute the number of loops with p links drawn on a lattice. This leads to the expression

$$B(u^{(1)}, \bar{u}^{(2)}) = \beta d u^{(1)^2} + \frac{1}{2} \text{Tr} \left\{ \int_0^\infty \frac{ds}{s} \left[I_0(2\beta \bar{u}^{(2)})^d - 1 \right] e^{-s} \right\} \quad (47)$$

Imposing stationarity with respect to $u^{(1)}$ and $\bar{u}^{(2)}$, one finds

$$\frac{\ln Z}{N} = \beta \nu - \frac{1}{2}n - \frac{1}{2}n \ln \frac{2\beta \nu}{n} + \int_0^\infty \frac{ds}{s} [I_0^d(s) - 1] e^{-s\nu} \quad (48)$$

where ν satisfies

$$\frac{2\beta}{n} = \int_0^\infty ds I_0^d(s) e^{-s\nu} \quad (49)$$

The critical value of β corresponds to the lowest possible value of ν , namely $\nu = d$. One verifies that the $1/d$ expansion of β_c coincides with the one of formula (44) when $n \rightarrow \infty$.

iii) Mean field equations and systematic corrections in the general case, using irreducible kernels.

Formula (31) can be generalized to arbitrary interactions. However the method is slightly cumbersome, due to the existence of two types of vertices (sites and interactions). Referring to the discussion in section (1.3), let us spell out some details. First let us define irreducibility with respect to interaction vertices. A connected diagram is said to be reducible with respect to a k -vertex if omission of the latter disconnects it into exactly k parts. This is not as general a definition as strong irreducibility, where it was only required that the diagram splits into connected parts. Diagrammatically,

we distinguish interaction vertices using black dots, site vertices by white ones. Equation (22)

$$m_i = \text{diagram of a site vertex with a loop} = \frac{\partial C(\{h_i\})}{\partial h_i} \quad (50)$$

in which a site vertex i is distinguished, is to be complemented by a corresponding formula for the interaction vertex

$$H_i = \text{diagram of an interaction vertex with a loop} = \frac{\partial S}{\partial \varphi_i} \Big|_{\varphi_i = m_i} \quad (51)$$

Equations (23) and (24) generalize as

$$C_\bullet = \text{diagram of a blob} + \text{diagram of a blob with a loop} + \text{diagram of a blob with two loops} + \dots = I(\{h_i + H_i\}) \quad (52)$$

with a similar relation for an interaction vertex. The topological relation

$$v - \sum (k-1)n_k = 1 \quad (53)$$

where v is the number of blobs and n_k the number of k -interaction vertices, leads to

$$C(\{h\}) = I(\{h + H\}) + S(\{m(i)\}) - \sum_i m_i H_i \quad (54)$$

Again (50) and (51) express stationarity with respect to H and m considered as independent variables. The expansion of I involves diagrams such that any interaction vertex belongs at least to one loop. It is natural to organize this series according to the number of loops.

$$I(h) = \text{diagram of a single loop} + \text{diagram of a double loop} + \dots \quad (55)$$

The first term corresponds to mean field, in complete generality (see formulas (6.61)-(6.66)). Again, a diagrammatic approach (as

opposed to a global method) does not allow competing extrema to be distinguished.

7.2.2 The role of symmetries

For models invariant with respect to a group of transformations, it is possible to achieve simplifications analogous to the use of the variable $\tanh \beta$ in the Ising case.

A *chiral* model involves fields which are elements of a compact group G . A natural basis to express any function of φ uses matrix elements $D_{\alpha\beta}^r(\varphi)$ of irreducible representations (r labels such representations). This is the content of the Peter-Weyl theorem in the case of square integrable functions on a compact semisimple Lie group. *A fortiori*, this is true for a finite group. The measure $d\varphi$ is the invariant, or Haar measure on the group. We assume that it is normalized to one. Then we have the familiar relations expressing orthogonality and completeness

$$\int d\varphi D_{\alpha\beta}^r(\varphi) \overline{D_{\gamma\delta}^s(\varphi)} = \frac{1}{d_r} \delta_{rs} \delta_{\alpha\gamma} \delta_{\beta\delta} \quad (56)$$

$$\sum_{r,\alpha,\beta} d_r D_{\alpha\beta}^r(\varphi) \overline{D_{\alpha\beta}^r(\varphi')} = \delta(\varphi, \varphi') \quad (57)$$

where d_r is the dimension of the r -representation, and $\delta(\varphi, \varphi')$ the appropriate Dirac measure (δ -function) on G . It is convenient to expand the Boltzmann weight $\exp S$ in terms of the various $D^r(\varphi)$ and to perform the integrations over the φ_i 's using group theoretic techniques, in particular equation (56), the simplest one of this kind.

As an illustration, let us return to the n -vector model. We denote in this case the group elements by $\mathcal{R} \in SO(n)$ and, considering the spin manifold as the homogeneous space $SO(n)/SO(n-1)$, we express the action using group variables as

$$S = \beta \sum_{\langle ij \rangle} \Phi_0 \mathcal{R}_i^{-1} \mathcal{R}_j \Phi_0 = \sum_{\langle ij \rangle} S_{ij}(\mathcal{R}_i^{-1} \mathcal{R}_j) \quad (58)$$

with Φ_0 a fixed (isotopic) vector. Using equation (56) one has

$$\exp S = \prod_{\langle ij \rangle} \sum_r d_r \text{Tr} \left(\beta_r D^r(\mathcal{R}_i^{-1} \mathcal{R}_j) \right) \quad (59)$$

with coefficients β_r ($r \times r$ matrices) given by

$$\beta_r = \int d\mathcal{R} \overline{\mathcal{D}^r(\mathcal{R})} \exp S_{ij}(\mathcal{R}) \quad (60)$$

and β_0 is a simple scalar. Each term in the product (59) is represented diagrammatically as follows. For each pair of neighbours (ij), choose a nontrivial representation $r \neq 0$, and draw a link from i to j indexed by r . The corresponding contribution is $d_r \beta_r / \beta_0$ after integration over the group variables. It is understood that in computing the partition function, an overall power of the scalar β_0 is factored out. The graphs will be simple with at most one (oriented) link joining two vertices. Changing the orientations of links amounts to substituting for the representation r its adjoint $\bar{r}$ with $\mathcal{D}^{\bar{r}} = (\mathcal{D}^r)^*$. According to equation (60), we have $\beta_{\bar{r}} = \overline{\beta_r}$.

In the Ising case, the group Z_2 has a single nontrivial representation, isomorphic to the group itself. Therefore $\beta_0 = \cosh \beta$, $\beta_1 = \sinh \beta$ and one recovers the standard rules of the $\tanh \beta$ expansion.

If several links, labelled by $r_1, \dots, r_n$, are incident on a vertex, the integration over the site field $\mathcal{R}_i$ yields a contribution of the form

$$\int d\mathcal{R} \mathcal{D}^{r_1}(\mathcal{R}) \dots \mathcal{D}^{r_n}(\mathcal{R})$$

The case of a single incident link, $n = 1$, is excluded since the representations r are assumed nontrivial and the corresponding integral vanishes. In the case $n = 2$, reversing one of the orientations, one finds that the two successive links have to carry identical representations. More precisely

$$\begin{aligned} \int d\mathcal{R}_j [d_r \text{Tr} \beta_r \mathcal{D}^r(\mathcal{R}_i^{-1} \mathcal{R}_j)] [d_s \text{Tr} \beta_s \mathcal{D}^s(\mathcal{R}_j^{-1} \mathcal{R}_k)] &= \\ &= \delta_{rs} [d_r \text{Tr} \beta_r^2 \mathcal{D}^r(\mathcal{R}_i^{-1} \mathcal{R}_k)] \end{aligned} \quad (61)$$

Thus, between sites i and k , we have a contribution similar to the one between two neighbours, with β_r replaced by β_r^2 . Repeating the process around a k -rung loop yields a factor

$$d_r \text{Tr} \beta_r^k \quad (62)$$

When more than two links are incident on a vertex, the computations become more cumbersome and involve coupling coefficients between several representations, as well as the corresponding β'_r 's.

If instead of the n -vector model, we consider the chiral case with an action

$$S = \beta \sum_{\langle i,j \rangle} \text{Tr}(\mathcal{R}_i^{-1} \mathcal{R}_j) \quad (63)$$

a slight simplification arises from the fact that the β'_r 's are scalars.

This is also the case for the plaquette action of lattice gauge models

$$S = \beta \sum_p \text{Tr}(\mathcal{R}_{ij} \mathcal{R}_{jk} \mathcal{R}_{kl} \mathcal{R}_{li}) \quad (64)$$

Rather than using the graphical prescriptions of section 2.1, it is more useful to represent the interaction by the plaquette $ijkl$, the sides of which carry the fields (figure 8). The corresponding diagrams have the topology of surfaces.

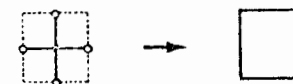


Fig. 8 Representation of the interaction used in gauge models.

It is also possible in this instance to obtain a restriction to connected diagrams by adapting the cumulant method. Unfortunately this has the disadvantage of generating a proliferation of graphs. For instance, if one enumerates the fourth order graphs arising on a lattice reduced to a single plaquette, one finds 16 distinct topological types including the sphere, the torus, the projective plane, the Klein bottle... Only one of the graphs is not connected. The method is obviously useful only in the case where one attempts to re-sum families of graphs with a fixed topology.

On the other hand, group theoretic integration techniques lead to an economy in the calculations. As above, the expansion

$$\exp S = \sum_p \sum_r \beta_r \chi_r(\mathcal{R}_{ij} \mathcal{R}_{jk} \mathcal{R}_{kl} \mathcal{R}_{li}) \quad (65)$$

in terms of the characters $\chi_r = \text{Tr } \mathcal{D}^r$, allows integration over the $\mathcal{R}$'s using formula (56). Diagrams are expressed as closed surfaces made of plaquettes. The simplest one is the boundary of a three-dimensional cube, with six faces, which may occupy $Nd(d-1)(d-2)/6$ positions on the lattice. This means that the partition functions admits of an expansion, which starts as

$$Z = \beta_0^{Nd} \left(1 + \frac{1}{6} Nd(d-1)(d-2) \sum_{r \neq 0} d_r^2 \left(\frac{\beta_r}{\beta_0} \right)^6 + \dots \right) \quad (66)$$

7.2.3 Low temperature expansion - discrete case

In the high temperature expansion, we selected one-by-one a configuration of interactions (depicted by a diagram and summed over field configurations). We now adopt the opposite view to obtain a low temperature series. Starting from a field configuration, chosen in general as an ordered state minimizing the (physical) energy (or maximizing the Boltzmann weight), we modify it successively by taking into account excitations. By necessity, the techniques are different when these excitations have a continuous or discrete spectrum. We consider here the easiest discrete case.

By assumption, the energy (or action) has a discrete spectrum $E_0, E_1, \dots$, each value with a degeneracy $n_0, n_1, \dots$, in any finite volume. Let us pretend that the ground state is nondegenerate ($n_0 = 1$) and assume that excited ones arise from a finite modification of the ground state configuration. We return to this point later. The partition function reads

$$Z = \sum n_i e^{-\beta E_i} \quad (67)$$

and the task is to classify states according to increasing energy. Configurations are identified by the set of fields φ_i , taking a different value as compared to the ones $\varphi_i^{(0)}$ in the ground state. Let us split the action as a sum $S_0 + S_1 + S_2 + \dots$ involving increasing numbers of modified fields

$$S_n \equiv \sum_{i_1, \dots, i_n} S_n(\varphi_{i_1}, \dots, \varphi_{i_n}) \quad (68)$$

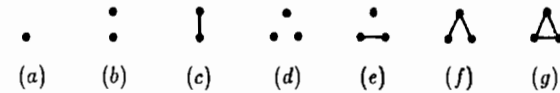


Fig. 9 The first few terms in the low temperature series for the Ising model.

with $S_n(\varphi_i^{(0)}, \dots, \varphi_{i_n}^{(0)}) = 0$ for $n > 0$, and $S_0 \equiv S(\{\varphi_i^{(0)}\})$. Thus we have

$$Z = e^{S_0} \sum_{\text{configurations}} \prod_i e^{S_1(\varphi_i)} \left[\prod_{ij} (1 + f_{ij}) \prod_{ijk} (1 + f_{ijk}) \dots \right] \quad (69)$$

with

$$f_{i_1 i_2 \dots i_n} = e^{S_n(\varphi_{i_1}, \dots, \varphi_{i_n})} - 1 \quad (70)$$

The quantity in brackets in (69) is then expanded, and represented diagrammatically. Vertices are labelled by the fields $\varphi_i \neq \varphi_i^{(0)}$ which differ from their ground state value. There are also vertices corresponding to "interactions", related to these sites and representing the various terms $f_{ij}, f_{ijk}, \dots$ chosen in the expansion of (69). The corresponding graphs are simple, in the sense that n sites are linked at most with one n -interaction vertex. The corresponding expansion is closely related to the low density expansion in the theory of imperfect gases.

As an illustration, let us return once again to the Ising model in the presence of a positive external field which lifts the ground state degeneracy by a quantity $2Nh$. We take the zero field limit $h \rightarrow +0$ only after the thermodynamic one. Only those states differing from the ground state, with all spins aligned parallel to the field, $\sigma_i^{(0)} = +1$, by finitely many spin flips, will contribute to the previous expansion.

In the zero field limit, for a lattice with N sites having q neighbours, we write the action as

$$S = \frac{1}{2} \beta N q + \beta q \sum_i (\sigma_i - 1) + \beta \sum_{(i,j)} (\sigma_i - 1)(\sigma_j - 1) \quad (71)$$

The first term yields the ground state contribution ($\exp \frac{1}{2} \beta q N$). The second term yields a factor $\exp(-2\beta q)$ when a spin is reversed. The final term, when expanded as in (69), attributes to every link joining neighbouring sites a factor $(\exp 4\beta - 1)$.

The first few diagrams are shown on figure 9. Take a hypercubic d -dimensional lattice, with $q = 2d$. The configuration numbers are respectively N , $\frac{1}{2}N(N-1)$, Nd , $\frac{1}{6}N(N-1)(N-2)$, $Nd(N-2)$, $Nd(2d-1)$ and 0 (on a cubic lattice). Correspondingly

$$\begin{aligned} \ln Z/N = & \beta d + e^{-4\beta d} + d e^{-8\beta d + 4\beta} - (d + \frac{1}{2})e^{-8\beta d} \\ & + d(2d-1)e^{-12\beta d + 8\beta} - 4d^2 e^{-12\beta d + 4\beta} \\ & + (2d^2 + d + \frac{1}{3})e^{-12\beta d} + \dots \end{aligned} \quad (72)$$

The effective expansion parameter is $\exp(-4\beta)$. The first part of the interaction term, namely $\exp(4\beta)$ tends to decrease the order of a given term in the expansion. In fact, it will exactly compensate the contribution of sites located inside a cluster of reversed spin which interact with their q neighbours. Thus only "surface" terms will be unbalanced. Since the surface increases with the size of the cluster, one is assured of being able to arrange the series in increasing powers of $\exp(-4\beta)$, with only a finite number of diagrams contributing to a given order. The case $d = 1$ is exceptional. No matter how large a sequence of reversed spins, its contribution will remain $\exp(-8\beta)$. We find here a phenomenon discussed in chapter 1 (volume 1), with $d = 1$ being the lower critical dimension where the ordered phase disappears. That the low temperature expansion involves a balance between energetic versus entropic contributions from surfaces bounding flipped spins, is the origin of the celebrated analysis of Peierls (1936), leading to a proof of the existence of ordered phases at low temperature. Peierls' argument has been the source of inspiration for many generalizations.

We observe in the series (72) a typical loss of positivity in the coefficients, as compared with the high temperature expansion for the partition function. As a consequence, the analysis of such low temperature series frequently reveals complex singularities which might limit their usefulness.

Low temperature expansions require a knowledge of the ground state which minimizes energy. The latter is in general not unique in the absence of an external field. But in ferromagnetic spin models, an infinitesimal external field is sufficient to lift the degeneracy by a term proportional to the volume. As temperature decreases and the limit of zero field is taken, degenerate ground states are separated by an infinite energy barrier and an expansion around one of them (describing a pure phase) is therefore justified.



Fig. 10 Equivalent low temperature diagrams in a lattice gauge theory.

One can encounter more complex situations. Take for instance an *antiferromagnetic* Ising model, with a reversed sign for the interaction term. At low temperature minimization of energy tends to favor antiparallel neighbouring spins. On certain lattices, which do not admit a bipartition, such configurations are not allowed. A simple example is the triangular two-dimensional lattice. On each elementary triangular face of the lattice, at least one pair of spins cannot be antiparallel. This is the phenomenon of *frustration*: a frustrated link is one for which spins are parallel. The distribution of frustrated links can produce a thermodynamic degeneracy of the ground states and leads to the possibility of transitions among them through excited states.

- i) Study the ground state degeneracy of the antiferromagnetic Ising model on a triangular lattice using the expressions given in chapter 2.
- ii) Obtain the low temperature expansion of a Z_2 -gauge model with an action

$$S = \beta \sum_p \sigma_{ij} \sigma_{jk} \sigma_{kl} \sigma_{li} \quad (73a)$$

with the field σ_{ij} taking values ± 1 . The ground state has all plaquette terms equal to +1, meaning that σ_{ij} is a pure gauge

$$\sigma_{ij} = s_i s_j \quad (s_i = \pm 1) \quad (73b)$$

whatever the values of the variables s_i are. The corresponding degeneracy is therefore 2^N .

Low temperature diagrams are generated by reversing the link variables σ_{ij} . But if all variables pertaining to links incident on a site are simultaneously reversed, this amounts to a gauge transformation (figure 10). One is consequently led to identify all diagrams corresponding to a reversal of all link variables incident on a site.

7.2.4 Low temperature expansion - continuous case

The previous method is to be modified in the continuous case. One has collective excitations (generalized spin waves) and formula (67) is replaced by a genuine functional integral, to which we can only apply the saddle point method and its appended perturbative - possibly renormalized - corrections. This technique is by now familiar to the reader. Let us summarize the main steps. Given a ground state configuration $\{\varphi_i^0\}$ satisfying

$$\left. \frac{\delta S}{\delta \varphi_i} \right|_{\varphi_i = \varphi_i^0} = 0 \quad (74)$$

the action is then expanded around this configuration, starting with a (hopefully) negative definite quadratic form corresponding to a genuine maximum

$$S(\varphi_i) = S(\varphi_i^0) + \frac{1}{2} \sum_{i,j} \Delta_{ij} (\varphi_i - \varphi_i^0)(\varphi_j - \varphi_j^0) + V(\{\varphi_i - \varphi_i^0\}) \quad (75)$$

The partition function is expressed as

$$\begin{aligned} Z &= \int \prod_i d\varphi_i \exp S(\varphi_i) \\ &= \exp(S(\varphi_i^0)) \exp(V(\{\delta/\delta j_i\})) \int \prod_i d\varphi_i \times \\ &\quad \exp \left\{ \sum_{i,j} (\varphi_i - \varphi_i^0) \Delta_{ij} (\varphi_j - \varphi_j^0) + \sum_i j_i (\varphi_i - \varphi_i^0) \right\} \Big|_{j=0} \\ &= \det^{-1/2} \Delta \exp(S(\varphi_i^0)) \exp(V(\{\delta/\delta j_i\})) \exp \left(\frac{1}{2} j_i (\Delta^{-1})_{ij} j_j \right) \Big|_{j=0} \end{aligned} \quad (76)$$

A diagrammatic expansion follows when $\exp V(\{\delta/\delta j_i\})$ is expanded in a power series. Each monomial in V is associated to a vertex. The effect of the functional derivatives $\delta/\delta j_i$ when we then set the source j to 0 is to contract in all possible ways these vertices $(\Delta^{-1})_{ij}$ (Wick's theorem). A difference with the discrete case, is that one obtains a series in powers of β^{-1} , rather than in $\exp(-\beta)$. Indeed Δ^{-1} is proportional to β^{-1} and, although V contains a factor β , the number of propagators overwhelms the number of interaction vertices. This is the standard counting of the number of loops, with β^{-1} playing the role of $\hbar$. Note that the stationary configuration φ^0 is β -independent.

A crucial difficulty is the possible existence of zero modes, in particular those originating from Goldstone's theorem. As a result, Δ is only semidefinite and cannot be inverted in a straightforward way. It is necessary to isolate these degrees of freedom and to treat them by a direct integration and not by the saddle point method, if possible. An example was worked out in volume 1 (chapter 6, section 2.4) to compute mean field corrections for a lattice continuous gauge field system. These zero modes arise generally from the existence of a continuous transformation group G , leaving the action invariant, and thus producing a continuum of ground state configurations starting from any given one. A general method for dealing with these problems is to introduce a symmetry breaking term which is then averaged out, in the following form. Let a noninvariant *gauge fixing* function $f(\varphi)$ be introduced. The term *gauge fixing* is borrowed for gauge theories, where the method was introduced. It is not meant here to imply necessarily a local invariance group. Let $^g\varphi$ denote the action of G on the configuration φ , and assume that the equation $f(^g\varphi) = c$ has a unique solution $g \in G$. There might not exist such a global gauge fixing, this difficulty is generally bypassed by requiring the above property to hold in a small neighborhood of the identity. For a global invariance group we need only require that $f(^g\varphi) = c$ fixes g only up to a transformation of the little group G^φ of φ^0 (the subgroup of G leaving φ^0 invariant, and as a consequence all terms in the action expanded around φ^0). This is an inessential complication in the following, and we leave it to the reader to modify accordingly the formulae.

Write

$$\det \mathcal{M}(\varphi) \int dg \delta(f(^g\varphi) - c) = 1 \quad (77)$$

where

$$\mathcal{M}(\varphi) = \frac{\delta f(^g\varphi)}{\delta g} \quad (78)$$

In principle, the Jacobian should be written $|\det \mathcal{M}(\varphi)|$ but according to the assumptions, the absolute value sign, can be omitted since $\det(\mathcal{M}(\varphi))$ does not vanish. Introducing the identity (77) into the functional integral for Z and using the invariance of the action S , the measure, as well as $\det \mathcal{M}(\varphi)$, one

gets

$$Z = \int e^{S(\varphi)} \delta(f(\varphi) - c) \det \mathcal{M}(\varphi) d\varphi dg \quad (79a)$$

One can now integrate over g . Moreover, since the result is c -independent, we can also integrate over c with an arbitrary (normalized) weight function $\exp F(c)$ (for instance proportional to $\exp -c^2$), and obtain

$$Z = \int \exp [S(\varphi) + F(f(\varphi)) + \text{Tr} \ln \mathcal{M}(\varphi)] d\varphi \quad (79b)$$

In this highly symbolic form, invariance is apparently lost (due to the presence of $F(f(\varphi))$) and the saddle point method applies without difficulty. The nonlocal term $\det \mathcal{M}(\varphi)$ can be replaced by an equivalent Grassmannian integral, introducing auxiliary (Faddeev-Popov ghost) fields, $\bar{\psi}, \psi$

$$Z = \int \exp [S(\varphi) + F(f(\varphi)) + \bar{\psi} \mathcal{M}(\varphi) \psi] d\varphi [d\psi d\bar{\psi}] \quad (79c)$$

A typical illustration is afforded by the nonlinear σ -model (or $O(n)$ model) in two dimensions (the lower critical dimension) at low temperature. This system with its global symmetry has a behaviour reminiscent of nonabelian gauge fields in four dimensions, namely it is asymptotically ultraviolet free. Since it is much simpler than its gauge counterpart, it has been under very active scrutiny to understand how to obtain reliable information on the spontaneous mass generation from a clever use of perturbative calculations. Moreover, interesting conjectures have been made (Zamolodchikov) concerning an exact expression for the scattering matrix. We refer the reader to the literature for a detailed treatment of the σ -model and various generalizations involving an arbitrary homogeneous manifold instead of a sphere for the target space. Such generalizations have been in particular extremely fruitful in the study of two-dimensional random systems using the replica trick (chapter 10) or in string field theory.

- i) Starting from an original G -invariant model, study the remnant symmetry in the form (79c) after gauge fixing.
- ii) For a two-dimensional n -vector model, with an action written as

$$S = -\frac{1}{2g_0^2} \int d^2x (\partial \Phi)^2 \quad \Phi \cdot \Phi = 1 \quad (80)$$

the renormalization of the coupling constant is given to lowest order by

$$\begin{aligned} \frac{1}{g^2} &= \frac{1}{g_0^2} + \frac{n-2}{2\pi} \ln \mu a + \dots \\ \mu \frac{\partial}{\partial \mu} g &\equiv \beta(g) = -\frac{n-2}{4\pi} g^3 + \dots \end{aligned} \quad (81)$$

with μ a renormalization scale and a the inverse (ultraviolet) cut-off. This expresses the ultraviolet asymptotic freedom property, stating that, as the mass scale μ gets large, the renormalized coupling g is driven to 0 provided that $n > 2$. The factor $(n-2)$ is in agreement with the fact that the property holds only for $n > 2$, the borderline being the XY -model with $n = 2$. An analytic continuation in n below 2 would imply reversing the conclusions.

7.2.5 Strong field expansions

Consider the case where a large external field is present. This does not modify in an essential way the derivation of low temperature series. Rather it suppresses the problem of degenerate extrema. Take for instance the Ising model, the vertex contribution becomes

$$z = \exp(-2H - 2\beta q) \quad (82)$$

where H denotes the external field, while the link contribution remains

$$u = \exp(4\beta) - 1 \quad (83)$$

We get a power series in two variables z and u , and the problem encountered in section 2.3 was to order it in the low temperature variable $(1+u)^{-1}$. Alternatively, we may look for an expansion valid when z goes to zero, i.e. for strong field. Powers of z count the number k of vertices. For fixed k , the number of links is at most $\frac{1}{2}k(k-1)$, and $\frac{1}{2}qk$ for a finite coordination number q . Each coefficient is a polynomial in u .

Thus equation (72) can be recast in the form

$$\frac{1}{N} \ln Z = H + \beta d + z + z^2 \left(du - \frac{1}{2} \right) + z^3 \left\{ d(2d-1)u^2 - 2du + \frac{1}{3} \right\} + \mathcal{O}(z^4) \quad (84)$$

in direct correspondence with the graphs of figure 8.

An application of these series is the study of the Lee–Yang edge singularity, as discussed in Chapter 3 (volume 1).

7.2.6 Fermionic fields

The extension of the diagrammatic rules to Grassmannian integrals is straightforward, with due care paid to antisymmetry properties (chapter 2, volume 1). Let us give an elementary example by rederiving a known result on Gaussian integrals through perturbative means.

Consider the integral over $\{q_i, \bar{q}_i\}$ anticommuting variables with a source term $\{\eta_i, \bar{\eta}_i\}$, itself anticommutative

$$\begin{aligned} Z(\bar{\eta}, \eta) &= \int \exp \left(m \sum_{i,\alpha} \bar{q}_i^\alpha q_i^\alpha \right) \left[\frac{dq d\bar{q}}{m} \right] \\ &\quad \times \exp \left(-\bar{q}_i^\alpha \Delta_{ij}^{\alpha\beta} q_j^\beta + \bar{\eta}_i^\alpha q_i^\alpha + \bar{q}_i^\alpha \eta_i^\alpha \right) \\ &= \det(m - \Delta) \exp \bar{\eta} (m - \Delta)^{-1} \eta \end{aligned} \quad (85)$$

To obtain the corresponding expansion, the mass term is essential and is kept as part of the measure. The second term in the exponential is expanded as an infinite series, and the resulting integrals are interpreted diagrammatically. Each term $\Delta_{ij}^{\alpha\beta}$ is represented by an ordered link, from site i to site j say. When integrating over $q_i^\alpha, \bar{q}_i^\beta$ the only nonvanishing terms are those with a factor 1 or a factor $\bar{q}_i^\alpha q_i^\alpha$. As a consequence the diagrams are either closed fermionic oriented loops, or arcs extending from a “source” $\bar{\eta}$ to a “sink” η . A loop $(i_1 i_2 \cdots i_n i_1)$ yields a factor $-\text{Tr}(\Delta_{i_1 i_2} \cdots \Delta_{i_n i_1})/nm^n$. The trace is over the internal indices $(\alpha\beta)$ of the matrices $\Delta_{ij}^{\alpha\beta}$. A minus sign is assigned to each fermionic loop, and arises from the anticommutativity of the variables $q, \bar{q}$ when they are reordered in the standard fashion $\bar{q}_i^\alpha q_i^\alpha$ to obtain a unit integral. The additional factor $1/n$ is a consequence of an incomplete cancellation among symmetry factors.

In principle, we have to reject those diagrams which intersect at a site i with the same internal index. This exclusion rule can in fact be ignored using the cumulant method as we saw previously. Due to the identity

$$\exp m \bar{q} q = 1 + m \bar{q} q$$

specific to Grassmannian variables, the diagrammatic rules for cumulants are identical to the rules in the (nonconnected) initial expansion.

Check that the sum over diagrams with more than one line (with the same internal index α) incident on site i vanishes.

Another consequence of the trivial transformation to cumulants is the immediate exponentiation of the expansion. In the absence of sources, $\ln Z$ is the sum over all diagrams with a unique loop. The only connected terms involving sources are those with a single arc from $\bar{\eta}_i$ to η_j .

We have thus reconstructed the expansion in Δ/m of

$$\ln Z(\bar{\eta}, \eta) = \text{Tr} \ln \left(1 - \frac{\Delta}{m} \right) + \frac{1}{m} \bar{\eta} \left(1 - \frac{\Delta}{m} \right)^{-1} \eta \quad (86)$$

Further examples will be presented in chapter 10. We also quote here that in three dimensions the Ising model admits a nontrivial Grassmannian presentation, with quartic interaction terms, which leads to elaborate expansion rules, loosely referred to as fermionic surfaces.

7.3 Enumeration of graphs

7.3.1 Configuration numbers with exclusion constraint

We have already observed that the notion of free graph enables one to factorize a contribution $n(G)$ involving only the geometry of the lattice. This a function of the lattice size N . If one uses, as is the most frequent case, periodic boundary conditions, a connected diagram will have a configuration number $n(G)$ proportional to N because of translational invariance if N is large enough to exclude special configurations which would take advantage of periodicity to include nontrivial loops. Even though we do not include these terms in the present context, their study is a legitimate part of the theory, and might reveal interesting finite size effects. More generally $n(G)$ will be a polynomial in N , the degree of which is equal to the number of disconnected parts of the graph. The term of degree 1 in N is of particular importance, since it is the

only one that survives in the computation of $\ln Z$, an extensive quantity proportional to N .

To compute these configuration numbers, we can restrict ourselves to simple graphs, since a multiple link does not affect the number of configurations. Moreover the computations can be reduced to those pertaining to certain multiply connected diagrams using the following property. Consider the set of graphs $G_1 \cup G_2 \equiv \{G\}$ that can be obtained starting from two given ones G_1 and G_2 by identifying certain of their elements. If $k(G; G_1, G_2)$ denotes the number of ways in which G can be split into G_1 and G_2 , the following formula holds

$$n(G_1)n(G_2) = \sum_{G=G_1 \cup G_2} k(G; G_1, G_2)n(G) \quad (87)$$

We now apply this formula in several instances.

(a) *Reduction to connected diagrams*

If G_1 and G_2 are free graphs with c_1 and c_2 connected parts respectively, G will have at most $c_1 + c_2$ connected parts, and only one of the graphs in $G_1 \cup G_2$ will have exactly this number, namely the graph where one does not identify any element in both graphs. Hence for instance the equation with two triangular graphs

$$n(\Delta \Delta)^2 = 2n(\Delta \Delta) + 2n(\Delta \Delta) + 2n(\Delta \Delta) + n(\Delta) \quad (88)$$

allows one to compute $n(\Delta \Delta)$ in terms of configuration numbers for connected graphs. Similar formulas are useful for gauge theories where graphs have the topology of surfaces. As an example for two cubes

$$n(\square \square)^2 = 2n(\square \square) + 2n(\square \square) + n(\square) \quad (89)$$

(b) *Reduction to multiply connected graphs*

A similar method applies when G_1 and G_2 have a common root i . In the graphs occurring on the r.h.s. of equation (87), a unique one will have an articulation point i . Thus

$$n(\Delta_i)^2 = 2n(\Delta_i) + 2n(\Delta_i) + n(\Delta_i) \quad (90)$$

When one sums the number of configurations of a graph over the position of the root i , one obtains Np times the number of configurations of the free graph, where p is the number of vertices equivalent to the rooted one under automorphisms of the graph.

Table I. Reduced number of configurations $\bar{n}(G)$ for loop graphs on triangular, square, cubic, body centered cubic and face centered cubic lattices.

ring	triangular	square	cubic	b.c.c.	f.c.c.
3	2	---	---	---	8
4	3	1	3	12	33
5	6	---	---	---	168
6	15	2	22	148	970
7	42	---	---	---	6168
8	123	7	207	2736	42069
9	380	---	---	---	301376
10	1212	28	2412	61896	2241420
11	3966	---	---	---	17173224
12	13265	124	31754	1579324	134806918
13	45144	---	---	---	1079802216
14	155955	588	452640	43702920	8798329680
15	545690	---	---	---	...
16	1930635	2938	6840774	1282524918	...
17	6897210	---	---	---	...
18	24852576	15268	108088232	---	...
20	...	81826	1768560270	---	...
22	---	449572	---	---	...
24	---	2521270	---	---	...
26	---	14385376	---	---	...

The previous formula in example (90) becomes

$$[3Nn(\Delta)]^2 = 2Nn(\Delta) + 4Nn(\Delta) + 3Nn(\Delta) \quad (91)$$

(c) *Reduction along a link*

One can furthermore envision the case where G_1 and G_2 have a common link ij , which allows a reduction of graphs which can be split along a link. Thus for example

$$n(\Delta_{ij})^2 = 2n(\Delta_{ij}) + n(\Delta_{ij}) \quad (92)$$

Since for coordination q , there are $\frac{1}{2}Nq$ links, one obtains

$$[3Nqn(\Delta)]^2 = 2Nqn(\Delta) + 3Nqn(\Delta) \quad (93)$$

These techniques allow one to derive configuration numbers from those of a smaller set of graphs for which simple and general methods are not available. A case by case study is then unavoidable. In the following tables, we show some of these enumerations. Table I counts closed loops. Table II

gives configuration numbers for some irreducible diagrams of a more complex structure on standard lattices (triangular and square two dimensional lattices, simple cubic, body centered cubic, and face centered cubic tridimensional ones). Finally table III gives configuration numbers for lattice gauge theories on a d -dimensional hypercubic lattice.

7.3.2 Multiply connected graphs

The preceding method reveals itself very quickly rather untractable because it generates a large number of terms. It is in particular necessary to compute configuration numbers for all graphs, multiply connected or not. It is of course possible to use general techniques such as cumulants or Legendre transforms. But the price is the loss of simplicity for graphs, even if the original ones admitted this restriction. Let us present a third technique, which requires suitable modifications according to the lattice and the model.

Let $A(G)$ and $B(G)$ be two families of weights satisfying

$$A(G) = \sum_{g \subseteq G} B(g) \quad (94)$$

and such that if G is the disjoint union $G_1 \cup G_2$, one has

$$g_1 \subseteq G_1, g_2 \subseteq G_2 \Rightarrow A(g_1 \cup g_2) = A(g_1) + A(g_2) \quad (95)$$

It then follows that the corresponding $B(G) = 0$, for G disconnected.

The proof relies on the inversion formula

$$(-1)^{p(G)} B(G) = \sum_{g \subseteq G} (-1)^{p(g)} A(g) \quad (96)$$

which is proved by induction on the number of links $p(G)$ of G .

To prove formula (96), one uses relations (94) and (96) for any graph g strictly included in G (as a recursive hypothesis). The coefficient of $B(g)$ is then $\sum_{G \supset g' \supseteq g} (-1)^{p(g')}$. Every pair g', g'' which differ only by a single fixed link not belonging to g , gives no contribution in this sum. The only remaining contribution is the one of the only graph g' equal to G with the fixed link deleted. This proves recursively equation (96).

Table II. Reduced number of configurations $\bar{n}(G)$ for some low order irreducible diagrams.

Diagram	triangular	square	cubic	b.c.c.	f.c.c.
	—	—	—	12	36
	—	—	—	—	2
	—	—	—	—	600
	—	—	—	—	48
	6	—	24	480	3888
	3	—	—	—	3132
	—	—	—	24	—
	—	—	—	3	9
	—	—	—	—	192
	—	20	436	—	—

The graphs g contributing to (96) can be subdivided into three categories: $g \subseteq G_1, g \subseteq G_2$ and those of the form $g = g_1 \cup g_2$ with $g_1 \subseteq G_1, g_2 \subseteq G_2$ both nonempty. The proof of the above proposition is then straightforward.

Let us apply this result to a lattice limited to the sites and links of a given graph. Choose $A(G) = \ln Z(G)$. The hypotheses are

Table III. Some graphs in lattice gauge theories and their configuration numbers.

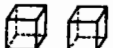
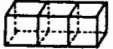
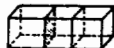
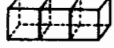
6.1		$\frac{1}{8}d(d-1)(d-2)$
10.1		$\frac{1}{2}d(d-1)(d-2)(2d-5)$
11.1		$\frac{1}{2}d(d-1)(d-2)(2d-5)$
12.1		$-\frac{1}{12}d(d-1)(d-2)(12d-29)$
12.2		$\frac{4}{3}d(d-1)(d-2)(d-3)$
12.3		$\frac{1}{3}d(d-1)(d-2)(d-3)$
14.1		$\frac{1}{2}d(d-1)(d-2)(2d-5)^2$
14.2		$2d(d-1)(d-2)(4d^2-22d+31)$
14.3		$d(d-1)(d-2)(d-3)$
14.4		$4d(d-1)(d-2)(d-3)$
15.1		$\frac{1}{3}d(d-1)(d-2)(d-3)(2d-5)$
15.2		$\frac{1}{3}d(d-1)(d-2)(d-3)$
15.3		$\frac{4}{3}d(d-1)(d-2)(d-3)$

Table III. (continued)

15.4		$d(d-1)(d-2)(2d-5)^2$
15.5		$4d(d-1)(d-2)(4d^2-22d+51)$
16.1		$\frac{1}{2}d(d-1)(d-2)(4d^2-24d+37)$
16.2		$\frac{1}{8}d(d-1)(d-2)(d-3)$
16.3		$8d(d-1)(d-2)(d-3)^2$
16.4		$16d(d-1)(d-2)(d-3)^2$
16.5		$16d(d-1)(d-2)(d-3)^2$
16.6		$\frac{1}{3}d(d-1)(d-2)(d-3)(2d-5)$
16.7		$-d(d-1)(d-2)(22d^2-113d+147)$
16.8		$2d(d-1)(d-2)(d-3)$
16.9		$\frac{1}{2}d(d-1)(d-2)(2d-5)^2$
16.10		$2d(d-1)(d-2)(4d^2-22d+31)$
16.11		$4d(d-1)(d-2)(d-3)$

justified when G is disconnected. As a result, for the complete lattice one has

$$\ln Z = \sum_{g \text{ connected}} n(g)B(g) \quad (97)$$

where the contributions $B(g)$ follow directly from equation (94).

To lowest orders with graphs up to three links,

$$\begin{aligned} \ln Z(/) &= B(/) \\ \ln Z(\vee) &= B(\vee) + 2B(/) \\ \ln Z(\mathbf{Z}) &= B(\mathbf{Z}) + 2B(\vee) + 3B(/) \\ \ln Z(\Upsilon) &= B(\Upsilon) + 3B(\vee) + 3B(/) \\ \ln Z(\triangle) &= B(\triangle) + 3B(\vee) + 3B(/) \\ &\dots \end{aligned} \quad (98)$$

This procedure is well adapted to automatic implementation on a computer. The inversion formula (96), valid for labelled graphs, is not true for free graphs.

In a second stage one may apply this same proposition, for articulated graphs, where the property indicated in equation (95) remains valid.

7.4 Results and analysis

7.4.1 Series analysis

The expansion techniques described in this chapter provide us with truncated series, up to an order n , of the form

$$f(z) = \sum_{k=0}^n f_k z^k + \mathcal{O}(z^{n+1}) \quad (99)$$

One is interested in two types of questions.

(1) Find the singularities z_i of the function, and study the behaviour of f in their vicinity. In particular one would like to know whether it is likely that the behaviour is of the form

$$f(z) \sim A_i(z - z_i)^{\gamma_i} \quad (100)$$

and, if this is the case, what is the exponent γ_i , and if possible the amplitude A_i .

(2) Extrapolate the function outside its circle of convergence. For instance, one may be interested by the limiting value as z reaches infinity in a given direction. It is clear that these are *a priori* unsolvable problems, unless one is given further information: regularity of the coefficients f_k , knowledge of the analyticity domain... in which case one will try to exploit it, using a method appropriate to each specific instance. Unfortunately these data are frequently missing, and one can only proceed blindly, using different techniques and comparing their results. We shall only give here a brief survey, referring the reader to specialized treatments.

It must be pointed out parenthetically that a vast knowledge has been accumulated on the large order behaviour of the continuous field theoretic perturbation series, originating in the seminal work of Bender and Wu (for the anharmonic oscillator) and Lipatov (for its transposition to genuine field theories). This information has provided a very important tool in the analysis of such series. Nothing as systematic is known in the case of lattice series, although isolated results exist in some instances, reflecting in particular the nature of the lattice.

The most straightforward technique is the *ratio method* and its refinements. If the coefficients f_k are real and the nearest singularity z_0 is unique, isolated, and "algebraic" (i.e. of the form (100)), the ratio of successive coefficients behaves as

$$\frac{f_k}{f_{k+1}} = z_0 \left[1 + \frac{1+\gamma}{k} + \mathcal{O}\left(\frac{1}{k^2}\right) \right] \quad (101)$$

The study of the sequence of ratios yields both z_0 and γ . One can profitably use such accelerators of convergence as the Neville tables in this context.

The Neville tables arise from a practical display of Aitken's iteration formulas. Let $f(x | x_0, \dots, x_k)$ denote the unique polynomial of degree k which coincides with $f(x)$ at the points $x_0, \dots, x_k$, the formulas

$$f(x | x_i, \dots, x_{i+n}) = \left| \begin{array}{cc} f(x | x_i, \dots, x_{i+n-1}) & x - x_i \\ f(x | x_{i+1}, \dots, x_{i+n}) & x - x_{i+n} \end{array} \right| / (x_i - x_{i+n}) \quad (102)$$

allow a recursive construction of interpolating expressions on 2, 3, ... points. When this is applied to a sequence $\{u_n\}$ with an asymptotic behaviour which is assumed (or known) to be polynomial in $1/n$, one is led to consider the set of sequences

$$\begin{cases} u_n^{(1)} = (n+1)u_{n+1} - nu_n = \lim u_n + \mathcal{O}\left(\frac{1}{n^2}\right) \\ u_n^{(2)} = \frac{1}{2} \left\{ (n+2)u_{n+1}^{(1)} - nu_n^{(1)} \right\} = \lim u_n + \mathcal{O}\left(\frac{1}{n^3}\right) \\ u_n^{(3)} = \frac{1}{3} \left\{ (n+3)u_{n+1}^{(2)} - nu_n^{(2)} \right\} = \lim u_n + \mathcal{O}\left(\frac{1}{n^4}\right) \\ \dots \end{cases} \quad (103)$$

In practice, one generally observes first an improved convergence, then, as the order increases, an instability (the correction terms with an increasing power in $1/n$, might have increasing coefficients).

Problems occur when the number of available terms in the series is insufficient for the singularity to impose its behaviour on the coefficients. Also such a method is tailored to deal with a unique singularity.

Padé approximants provide a powerful means of investigating several singularities at once, with larger convergence domains. The idea is to approximate an analytic function by a sequence of rational ones. More precisely the Padé approximant $[L/M](z) \equiv P_L(z)/Q_M(z)$ is the unique rational fraction, a ratio of two polynomials of degree L and M respectively, defined up to a common multiple, such that its Taylor series up to degree $L+M$ agrees with the known expansion of $f(z)$

$$f(z) - P_L(z)/Q_M(z) = \mathcal{O}(z^{L+M+1}) \quad (104)$$

Equation (104) yields $(L+M+1)$ equations on $(L+1)+(M+1)-1$ unknowns. Padé approximants are by definition particularly suitable for reconstructing meromorphic functions. Experimentally, one observes that they simulate branch points quite well, by accumulating alternatively zeroes and poles on the branch cut. They are well suited for obtaining analytic continuation beyond the natural circle of convergence of Taylor series. Difficulties may sometimes occur with an erratic convergence, in addition to the occurrence of spurious singularities.

When one assumes an isolated power law singularity, meaning that $(z-z_0)^{-\gamma}f(z)$ is regular at $z=z_0$, it is better to apply Padé approximants to the logarithmic derivative

$$\frac{1}{f} \frac{df}{dz} = \frac{\gamma}{z-z_0} + \text{regular part}$$

with an isolated pole. In practice, nonfactorizable singularities might occur $(A(z)(z-z_0)^\gamma + B(z))$ as well as confluent singularities of the form $(A(z-z_0)^\gamma + B(z-z_0)^\beta + \dots)$, and even worse. The method loses its accuracy, leading possibly to a reasonable location of z_0 , with an unstable estimate of the exponent γ .

Integral approximants generalize the preceding technique. These are solutions of a differential equation of a suitable type, with polynomial coefficients determined in such a way that, up to the order to which it is known, $f(z)$ satisfies the equation. For instance, the first order linear differential equation

$$Q_K(z)f'(z) + P_M(z)f(z) + R_L(z) = \mathcal{O}(z^{K+L+M+2}) \quad (105)$$

yields standard Padé approximants for $Q_k \equiv 0$ ($k \equiv -1$), and those of the logarithmic derivative, when $R_L \equiv 0$. These approximants are singular at the zeroes of Q , and behave as

$$\begin{aligned} & A(z-z_0)^\gamma + B \\ \gamma &= -\frac{P(z_0)}{Q'(z_0)} \quad B = -\frac{R(z_0)}{P(z_0)} \end{aligned}$$

These approximants are in general suited to take into account subdominant terms in the vicinity of singularities, and yield better values for the exponents. The trouble is the proliferation in the number of possible approximants, with large deviations between various ones, leading to a difficult choice between several estimates. These integral approximants admit a further generalization to several variables, and hence allow the study of multicritical points.

Information on the *analytic properties* of the function f , can be usefully combined with the preceding techniques. Let us give an example where it is assumed that at a *known* value z_c , the function $f(z)$ has an accumulation of singularities of the type

$$f(z) = \sum_{n,p=0}^{\infty} C_{np} \left(1 - \frac{z}{z_c}\right)^{\gamma_n} \ln^p \left(1 - \frac{z}{z_c}\right) \quad (106)$$

Then $f(z)$ admits an integral representation

$$f(z) = \int \left(1 - \frac{z}{z_c}\right)^\lambda \sigma(\lambda) d\lambda \quad (107)$$

and its transform

$$g(t) = \int \frac{\sigma(\lambda)}{1 - \lambda t} d\lambda = \sum_{n,p=0}^{\infty} C_{np} \frac{\gamma_n^p p!}{(1 - t\gamma_n)^{p+1}} \quad (108)$$

is a meromorphic function in t . It is not difficult to relate the Taylor coefficients of $g(t)$ to those of $f(z)$. Padé approximants, ill-adapted to the confluent singularities of $f(z)$, are on the contrary very well suited to the analysis of the meromorphic structure of $g(t)$. This is the so-called *Padé-Mellin* method.

In short, it is clear that extracting exponents and amplitudes from a limited perturbative series is an art form. Except in a few simple cases, it requires sophisticated expertise.

7.4.2 An example: the Ising series on a body centered cubic lattice

We cannot of course show here all the series obtained in a great variety of models. We have given in chapter 6 (volume 1) examples for lattice gauge models. Here we limit ourselves to quoting a few results pertaining to the three-dimensional Ising model on a body centered cubic lattice, a system for which rather long series are known. Other results will be found in the references.

The high temperature series in the variable $t \equiv \tanh \beta$ for the free energy per site is

$$\begin{aligned} \frac{1}{N} \ln Z = & \ln 2 - 4 \ln(1+t) + 12t^4 + 148t^6 + 2496t^8 + 52168t^{10} + \\ & + 1242078t^{12} + 32262852t^{14} + 892367762t^{16} + \mathcal{O}(t^{18}) \end{aligned} \quad (109)$$

By taking successive derivatives, this series allows the internal energy and the specific heat to be found. One can in this way obtain the exponent α .

The correlation function $\langle \sigma_0 \sigma_{\mathbf{x}} \rangle$ depends on the (vector) variable $\mathbf{x}$. The corresponding high temperature diagrams have odd vertices in $\mathbf{0}$ and $\mathbf{x}$. One is interested in the behaviour at large distance, or, equivalently, for small wavenumber of its Fourier

transform,

$$\tilde{\chi}(\mathbf{k}) = \sum e^{i\mathbf{k} \cdot \mathbf{x}} \langle \sigma_0 \sigma_{\mathbf{x}} \rangle = \chi - \frac{\mu_2}{2d} \mathbf{k}^2 + \mathcal{O}(\mathbf{k}^4) \quad (110)$$

Here χ is the static susceptibility. The lattice anisotropy shows up in fourth order terms in $\mathbf{k}$. If one ignores these anisotropies, one can study alternatively the spherical moments

$$\mu_n = \sum_{\mathbf{x}} |\mathbf{x}|^n \langle \sigma_0 \sigma_{\mathbf{x}} \rangle \quad (111)$$

which are obviously convergent in the high temperature phase, where the correlation function decreases exponentially. In particular $\chi = \mu_0$ diverges at the critical temperature with an exponent γ . The effective correlation length $\xi_{eff}^2 = 2d\chi/\mu_2$ diverges with the exponent ν . For the body centered cubic lattice, the following series are known

$$\begin{aligned} \chi = & 1 + 8t + 56t^2 + 392t^3 + 2648t^4 + 17864t^5 + 118760t^6 \\ & + 789032t^7 + 5201048t^8 + 34268104t^9 + 224679864t^{10} \\ & + 1472595144t^{11} + 9619740648t^{12} + 62823141192t^{13} \\ & + 409297617672t^{14} + 2665987056200t^{15} + 17333875251192t^{16} \\ & + 112680746646856t^{17} + 731466943653464t^{18} \\ & + 4747546469665832t^{19} + 30779106675700312t^{20} \\ & + 199518218638233896t^{21} + \mathcal{O}(t^{22}) \end{aligned} \quad (112)$$

and

$$\begin{aligned} \mu_2 = & 8t + 128t^2 + 1416t^3 + 13568t^4 + 119240t^5 \\ & + 992768t^6 + 7948840t^7 + 61865216t^8 + 470875848t^9 \\ & + 3521954816t^{10} + 25965652936t^{11} + 189180221184t^{12} \\ & + 1364489291848t^{13} + 9757802417152t^{14} + 69262083278152t^{15} \\ & + 488463065172736t^{16} + 3425131086090312t^{17} \\ & + 23896020585393152t^{18} + 165958239005454632t^{19} \\ & + 1147904794262960384t^{20} + 7910579661767454248t^{21} \\ & + \mathcal{O}(t^{22}) \end{aligned} \quad (113)$$

Observe that both series have integral positive coefficients, so that we are sure of the existence of a nearest singularity for positive t .

Let us illustrate series analysis using the example of the susceptibility. We face two difficulties. One arises from the

symmetries of the lattice which admits a bipartition in two simple cubic sublattices, with the spins on one sublattice interacting only with those of the second. If one changes the sign of all spins on one sublattice and β into $-\beta$ (or t into $-t$), the partition function is invariant, as is clear from equation (109) apart from a normalization term in $\ln(1+t)$. This symmetry leads to an antiferromagnetic transition at $-t_c$. The susceptibility can be split into even and odd parts, according to the contributions of pairs of spins located on the same or different sublattices. The odd part is related to the even one by increasing each spatial contribution by a unit step. This amounts loosely to multiplying it by the internal energy, with an exponent $\alpha - 1$. One thus expects that the coefficients χ_n in the series (112) behave as

$$\chi_n \approx A n^{\gamma-1} t_c^{-n} [1 + B(-1)^n n^{\alpha-\gamma+1}] \quad (114)$$

It is then natural to split the series into even and odd parts in t , and to consider the quantities

$$y_n = - \left[\ln \left(\frac{\chi_n \chi_{n-4}}{\chi_{n-2}^2} \right) \right]^{-1} \approx \frac{n^2}{\gamma - 1} \quad (115)$$

The exponent γ is the limit of the sequence

$$\gamma_n = 1 + 2 \frac{y_n + y_{n-1}}{(y_n - y_{n-1})^2} \quad (116)$$

The second difficulty is the existence of a confluent singularity predicted by the renormalization group. The expected correction to formula (114) is of the form $(1 + cn^{-\Delta_1})$, where $\Delta_1 = \omega\nu$. As a consequence, one cannot efficiently use the Neville tables to accelerate the convergence of the sequences. Indeed one observes strong instabilities if one attempts to use these techniques. However, if t_c is sufficiently accurately determined, one can use the Padé-Mellin method. Yet another possibility is to use the approximate value $\Delta_1 \simeq 0.50$ to suppress this singularity. In the analysis carried out by Zinn-Justin, one introduces the average $\bar{y}_n = (y_n + 2y_{n-1} + y_{n-2})/4$ and considers the quantity

$$Y_n = \left(\frac{\bar{y}_n^{1/2\Delta_1} - \bar{y}_{n-2}^{1/2\Delta_1}}{\Delta_1} \right)^{2/(\Delta_1-1)} \quad (117)$$

which replaces (115) by eliminating entirely the correction due to the first confluent singularity.

From the series, a consistent set of data on the singularities can be extracted in the form

$$\begin{aligned} t_c^{-1} &= 6.3543 \\ \gamma &= 1.239 \pm 0.0025 \\ \nu &= 0.6305 \pm 0.0015 \\ \eta &= 0.035 \pm 0.003 \end{aligned} \quad (118)$$

The effective correlation length computed as indicated above using the second moment μ_2 , does not coincide with the "true" correlation length defined as

$$\xi(\hat{\mathbf{x}})^{-1} = - \lim_{|\mathbf{x}| \rightarrow \infty} \frac{\ln \langle \sigma_0 \sigma_{\mathbf{x}} \rangle}{|\mathbf{x}|} \quad (119)$$

for which one can also obtain high temperature series. The pitfalls in analysing this quantity have been already discussed in chapter 6 (volume 1, section 3.5).

The low temperature series are more irregular on this same b.c.c. lattice, due to the existence of spurious complex singularities, and hence lead to a less trustworthy determination of the critical exponents.

Using as a variable $u = \exp -4\beta$, the free energy per site is given by

$$\begin{aligned} \ln Z &= u^4 + 4u^7 - 4\frac{1}{2}u^8 + 28u^{10} - 64u^{11} + 48\frac{1}{3}u^{12} + 204u^{13} - 786u^{14} \\ &+ 1164u^{15} + 922\frac{3}{4}u^{16} - 8760u^{17} + 20032u^{18} - 9164u^{19} \\ &- 84215\frac{4}{5}u^{20} + 294677\frac{1}{3}u^{21} - 378996u^{22} - 569704u^{23} \\ &- 54012882u^{24} + 112640896u^{25} - 5164464u^{26} \\ &- 694845120u^{27} + 2160781086u^{28} + \mathcal{O}(u^{29}) \end{aligned} \quad (120)$$

The pattern of minus signs is suggestive of the complex nearby singularities mentioned above. We use here an abbreviated notation, for instance $4\frac{1}{2}$ instead of $(4 + \frac{1}{2})$, to avoid huge numerators in the fractions.

The spontaneous magnetization $\langle \sigma \rangle$ is given by the series

$$\begin{aligned} \langle \sigma \rangle &= 1 - 2u^4 - 16u^7 + 18u^8 - 168u^{10} + 384u^{11} - 314u^{12} - 1632u^{13} \\ &+ 6264u^{14} - 9744u^{15} - 10014u^{16} + 86976u^{17} - 205344u^{18} \\ &+ 80176u^{19} + 1009338u^{20} - 3579568u^{21} + 4575296u^{22} \\ &+ 8301024u^{23} + 3832961\frac{1}{2}u^{24} - 7942796u^{25} + 1118118u^{26} \\ &+ 43016052u^{27} - 133595088\frac{6}{7}u^{28} + \mathcal{O}(u^{29}) \end{aligned} \quad (121)$$

Using $\langle \sigma \rangle$, it is possible to compute the connected correlation function $\langle \sigma_0 \sigma_{\mathbf{x}} \rangle_c = \langle \sigma_0 \sigma_{\mathbf{x}} \rangle - \langle \sigma \rangle^2$, and to obtain the corresponding low temperature susceptibility

$$\begin{aligned} \frac{1}{4}\chi = & u^4 + 16u^7 - 18u^8 + 252u^{10} - 576u^{11} + 519u^{12} + 3264u^{13} \\ & - 12468u^{14} + 20568u^{15} + 26662u^{16} - 215568u^{17} + 528576u^{18} \\ & - 164616u^{19} - 3014889u^{20} + 10894920u^{21} - 13796840u^{22} \\ & - 2990961u^{23} + 190423962u^{24} - 399739840u^{25} - 22768752u^{26} \\ & + 2803402560u^{27} - 8743064909u^{28} + \mathcal{O}(u^{29}) \end{aligned} \quad (122)$$

Analyzing the poles and residues of the Padé approximants of the logarithmic derivative of these series allows one to obtain the critical exponents within the ordered phase. However, these results are less reliable than those obtained from high temperature series, due to more important fluctuations in the estimates. One finds that the equality $\gamma' = \gamma$ is approximately satisfied. At this stage the low temperature exponents are given by

$$\begin{cases} \alpha' \simeq 0 \\ \beta = 0.312 \pm 0.05 \\ \gamma' = 1.25 \pm 0.05 \end{cases} \quad (123)$$

Let us conclude by giving the strong field expansion of the free energy in the variables $\mu = \exp(-2\beta H)$ and $u = \exp(-4\beta)$

$$\begin{aligned} \ln Z = & u^4 \mu + (4u^7 - 4\frac{1}{2}u^8)\mu^2 + (28u^{10} - 64u^{11} + 36\frac{1}{3}u^{12})\mu^3 + \\ & (12u^{12} + 204u^{13} - 798u^{14} + 948u^{15} - 366\frac{1}{4}u^{16})\mu^4 \\ & + (12u^{14} + 216u^{15} + 1262u^{16} - 9072u^{17} + 17592u^{18} \\ & - 14184u^{19} + 4174\frac{1}{5}u^{20})\mu^5 + (27u^{16} + 312u^{17} \\ & + 2368u^{18} + 4312u^{19} - 92992u^{20} + 275021\frac{1}{3}u^{21} \\ & - 353640u^{22} + 216036u^{23} - 51444\frac{1}{2}u^{24})\mu^6 \\ & + (72u^{18} + 704u^{19} + 4404u^{20} + 17616u^{21} - 36348u^{22} \\ & - 833064u^{23} + 3795726u^{24} - 7072736u^{25} + 6798900u^{26} \\ & - 3344712u^{27} + 669438\frac{1}{7}u^{28})\mu^7 + \mathcal{O}(\mu^8) \end{aligned} \quad (124)$$

Notes

The subjects treated in this chapter are rather technical and are

found in the series *Phase Transitions and Critical Phenomena*, vol. 3 entitled *Series Expansions for Lattice Models*, C. Domb and M.S. Green eds, Academic Press (1974), and more recently in the *Proceedings of the 1980 Cargèse Summer Institute on Phase Transitions*, M. Levy et al. eds, Plenum (1982).

General graph theory is treated in C. Berge, *The Theory of Graphs*, Methuen (1962). The graph reduction technique is discussed in F. Englert, *Phys. Rev.* **129** 567 (1963), which contains additional references.

Series expansions, dealing mostly with the Ising model, are presented in C. Domb, *Adv. Phys.* **19**, 339 (1970) and in a sequence of papers by M.F. Sykes and collaborators, *J. Math. Phys.* **6** 283 (1965); *J. Phys.* **A5** 624, 640, 653, 661, 667 (1972); *J. Phys.* **A6** 1517 (1973); *J. Math. Phys.* **14** 1060, 1066, 1071 (1973); *J. Phys.* **A12** L25 (1979). See also S. McKenzie, *J. Phys.* **A8** L102 (1975) and **A12** L185 (1979), and B.G. Nickel in the proceedings of 1980 Cargèse summer institute quoted above. The problems specific to the spin-spin correlation functions are discussed in M.E. Fisher and R.J. Burford, *Phys. Rev.* **156** 583 (1967).

Padé approximants are described in a book by G.A. Baker, Jr, *Essentials of Padé Approximants*, Academic Press (1975). The analysis of confluent singularities using the Padé-Mellin method is due to G.A. Baker and D.L. Hunter, *Phys. Rev.* **B7** 3377 (1973). Examples of other techniques are found in J. Zinn-Justin, *J. Physique* **42** 783 (1981). For a generalization to several variables, see M.E. Fisher and J.H. Chen in the Proceedings of 1980 Cargèse summer institute quoted above, and J.H. Chen, M.E. Fisher and B.G. Nickel, *Phys. Rev. Lett.* **48** 630 (1982).

The work of R. Peierls on the existence of ordered phases at low temperature has been quoted in chapter 2 (volume 1). For the asymptotic freedom of the σ -model, see A.M. Polyakov, *Phys. Lett.* **B59**, 79 (1975), E. Brézin and J. Zinn-Justin, *Phys. Rev.* **B14**, 3110 (1976). Exact S-matrices for two dimensional models are due to A.B. Zamolodchikov and Al.B. Zamolodchikov, *Nucl. Phys.* **B133**, 525 (1978).

NUMERICAL SIMULATIONS

A new field opened when modern computers offered the possibility of performing extensive simulations of large systems. This allows known behaviours to be checked and provides an exploratory guide in circumstances where theoretical tools are absent. Measurements of observables can be compared both to existing theoretical expectations – providing a crucial test for their validity – and to experiments – checking the modelling of a physical system –. This chapter presents the background material needed to design simulations on a (relatively) large scale. Some numerical examples have already been presented in previous chapters, and we only give a few further illustrations, pertaining mainly to lattice gauge theory, the usefulness of which relies extensively on this method. We also describe a practical implementation of real space renormalization, known as the *Monte Carlo Renormalization Group* method (Ma, Swendsen, Wilson). Finally, we discuss specific issues relevant to the extension of the simulations to fermionic systems.

8.1 Algorithms

8.1.1 Generalities

Systems with up to 10^6 to 10^7 variables can be handled by computers, and these numbers may soon be significantly increased. The measurements can be sufficiently numerous to allow statistical accuracy. Although simulated systems still have a modest size as compared to macroscopic systems, collective effects already clearly appear and accurate results about critical phenomena emerge from the numerical simulations. It turns out, when investigating more closely the available numerical methods, that one gets a better insight into the foundations of equilibrium statistical physics, the

ergodicity problems, the meaning of probabilities and, last but not least, ways to approach equilibrium. Although we seem to deal in this chapter with technical matters, the alert reader will realize that the subject in fact deserves the closest scrutiny and that experience can only be gained through practice.

The large number of variables forbids any *exact* numerical treatment, for instance by enumerating all possible configurations ($\sim 10^{8000}$ for a small Ising system $30 \times 30 \times 30$) and computing their contributions. There exists however a few attempts at deriving explicit expressions for very small lattices (with a $4 \times 4 \times 4$ Ising system, the number of configurations is already of macroscopic order $2 \cdot 10^{19}$! Luckily, symmetries simplify considerably the task of enumerating them). In practice, it is therefore necessary to use a representative sampling of the system. On each sample, one performs the desired measurements and the statistical analysis is done exactly as for an ordinary experiment. One has to face two additional difficulties. The first is to simulate nature, which freely provides a representative sampling of a system, and the second is to estimate the biases arising from small size.

A direct method would consist, using a random number generator, in choosing successively and independently the different degrees of freedom of the system and to assign to the resulting configuration its statistical weight. This weight is the product of the Boltzmann factor by the measure on the fields. A well-known application is the roulette wheel which ensures a positive income (without any risk) for casinos at the expense of players who ignore the laws of statistics, or want to ignore them. The application of this idea is at the origin of the name *Monte Carlo method* given to such simulations. The simple sampling is currently used to investigate percolation and fluid flows for instance, and yields accurate results in these cases.

However, simple random sampling is not practical for the type of systems we have in mind, generalized spin systems, gauge theories, random surfaces, etc., for which the Boltzmann weights relate several degrees of freedom and may generate nontrivial collective effects. The energy fluctuations are of order $1/\sqrt{N}$, where N is the system size. Configurations which are not energetically very close to the mean energy (or mean action), have a negligible effect in the computation of mean values. As the simple random sampling does not take into account this fact, it is necessary to introduce

a bias. Let $\mathbf{x}$ represent the set of all variables characterizing the configuration of a system. A numerical simulation generates a sequence $\{\mathbf{x}_i\}$ of independent configurations randomly chosen according to the same probability distribution $P(\mathbf{x}_i)$. Using the notation of statistical physics, one replaces the computation of an average observable

$$\langle f \rangle = \frac{\sum_{\mathbf{x}} f(\mathbf{x}) \exp(-\beta \mathcal{H}(\mathbf{x}))}{\sum_{\mathbf{x}} \exp(-\beta \mathcal{H}(\mathbf{x}))} \quad (1)$$

by the evaluation of a sequence

$$f_n = \frac{\sum_{i=1}^n f(\mathbf{x}_i) \exp(-\beta \mathcal{H}(\mathbf{x}_i)) P^{-1}(\mathbf{x}_i)}{\sum_{i=1}^n \exp(-\beta \mathcal{H}(\mathbf{x}_i)) P^{-1}(\mathbf{x}_i)} \quad (2a)$$

and one checks that, in the infinite "time" limit $n \rightarrow \infty$, the random variables f_n tend towards the required quantity

$$\lim_{n \rightarrow \infty} f_n = \langle f \rangle \quad (2b)$$

up to errors of order $1/\sqrt{n}$.

Verify that if all configurations $\mathbf{x}_i$ are chosen independently with the same probability law $P(\mathbf{x})$, the time average $(1/n) \sum_{i=1}^n f(\mathbf{x}_i)$ tends to $\sum_{\mathbf{x}} P(\mathbf{x}) f(\mathbf{x})$. As a consequence check the validity of the limit (2b).

The Boltzmann factor can vary over several orders of magnitude. If $P(\mathbf{x})$ is almost constant, very few samples contribute to the sum in (2a) and the number n of iterations, i.e. the computing time, becomes enormous if a reasonable accuracy is to be achieved. However, it is possible to correct this misfortune by adjusting the probability $P(\mathbf{x})$. This leads to the idea of *importance sampling*, namely to adapt the bias probability to the quantities to be computed. The optimal choice is obviously the Boltzmann weight itself, $P(\mathbf{x}) = \exp(-\beta \mathcal{H}(\mathbf{x}))$.

Due to the interactions, each field (component of $\mathbf{x}$) cannot be selected independently from the others. A possibility is to use a Markov process which generates $\mathbf{x}_{i+1}$ from the knowledge of $\mathbf{x}_i$. Two successive configurations $\mathbf{x}_i$ and $\mathbf{x}_{i+p}$, separated by a finite "time" interval p , are thus correlated, but this is unimportant if this correlation decreases fast enough with p so that the probability distribution of $\mathbf{x}_{i+p}$ approaches, independently of

$\mathbf{x}_i$, the desired probability $P(\mathbf{x})$. Under these conditions, relation (2b) remains of course valid. Let $W(\mathbf{x}_i, \mathbf{x}_{i+1})$ be the conditional probability to select $\mathbf{x}_{i+1}$ starting from $\mathbf{x}_i$, thus defining the Markov process. It should satisfy the following properties.

- i) *Normalization* $\sum_{\mathbf{y}} W(\mathbf{x}, \mathbf{y}) = 1$
- ii) *Ergodicity* $W(\mathbf{x}, \mathbf{y}) > 0$. Indeed, one should be able to reach all configurations.
- iii) *Limiting probability*

$$\sum_{\mathbf{x}} P(\mathbf{x}) W(\mathbf{x}, \mathbf{y}) = P(\mathbf{y}) \quad (3)$$

The distribution W , considered as a matrix in the space spanned by configurations, has a unique maximal eigenvalue equal to 1 (according to conditions i and ii). Equation (3) indicates that the corresponding positive eigenvector is $P(\mathbf{x})$. Hence W^p applied on any state distribution probability transforms it into the equilibrium probability $P(\mathbf{x})$ as p increases indefinitely. The limiting distribution of the estimator f_n of $\langle f \rangle$ can also be easily obtained. For large n , this is of course the normal Gaussian probability $\exp[-(f_n - \langle f \rangle)^2 / 2\sigma_n^2]$, with a standard deviation

$$\lim_{n \rightarrow \infty} n\sigma_n^2 = \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^n (f(\mathbf{x}_i) - \langle f \rangle)^2 \right) \quad (4)$$

Check this result, using Fourier transforms of probabilities.

The three conditions are far from being sufficient to determine in a unique way the conditional probability $W(\mathbf{x}, \mathbf{y})$. In practice, it is convenient to decompose W into elementary steps

$$W(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{z}_1} W_1(\mathbf{x}, \mathbf{z}_1) W_2(\mathbf{z}_1, \mathbf{z}_2) \cdots W_k(\mathbf{z}_{k-1}, \mathbf{y}) \quad (5)$$

Condition (3) is replaced by the more severe *detailed balance* condition

$$P(\mathbf{x}) W_j(\mathbf{x}, \mathbf{y}) = P(\mathbf{y}) W_j(\mathbf{y}, \mathbf{x}) \quad (6)$$

Relation (3) is then automatically fulfilled. It is easier to check (6) than (3), and the symmetry of this relation is often useful. Nevertheless, even with this additional condition, the

indeterminacy in the choice of W remains large and leads to various possible algorithms.

In practical realizations, one proceeds in elementary steps which affect only a small number of degrees of freedom in the system. For instance, W_j just updates the field at location j and leaves the remaining fields unchanged. This is of interest because, for systems with short-range interactions, the amount of computation needed for this update remains limited and independent of the size of the system. The total computer time for a complete Monte Carlo "sweep" (i.e. a step of the Markov chain affecting the whole configuration) is therefore proportional to the size of the system. To be impartial, one should also mention the main defect of this process. It is difficult to excite long wavelength collective modes at low temperature and near criticality. The thermalization time needed to erase the memory of the initial configuration becomes relatively important. Moreover, in the critical domain, there are important correlations between successive steps (*critical slowing down*). This is equivalent to saying that there exist dynamical critical phenomena in the approach to equilibrium, as a counterpart to the static ones. However, this technique is presently the only well understood approach and one of the most efficient.

8.1.2 The classical algorithms

A very simple way to satisfy condition (6) is to make the choice

$$W_j(\mathbf{x}, \mathbf{y}) = P(\mathbf{y}) \propto \exp(-\beta \mathcal{H}(\mathbf{y})) \quad (7)$$

This choice can be interpreted physically as a thermalization of a spin with a source at the desired temperature, the other parts of the system being frozen. The target spin chooses its final configuration according to the equilibrium temperature, taking into account the interactions with its neighbours which do not evolve. This procedure is known as the *heat bath* algorithm.

Compared to the initial problem of choosing the whole configuration with a probability proportional to the Boltzmann weight, the process is considerably simplified, since the probability (7) depends only on the neighborhood of one site. In a practical implementation for the Ising model, one computes the local field $h_j = \sum_{k(j)} \sigma_k$ seen by the spin at site j . The spin is then randomly

reset to ± 1 (independently of its previous value), with a probability $p_j = [1 + \exp(-2\beta h_j)]^{-1}$ to take the value $+1$ (in practice, one generates a random number uniformly distributed between 0 and 1 and the spin is set to $+1$ if this number is less than p_j , to -1 otherwise). The heat bath algorithm leads also to rapid calculations with more general discrete models.

However, and in particular for continuous systems, generating random numbers according to the law (7) is not straightforward, and may lead to rather long calculations which should be avoided. Let us describe a general method for continuous distributions. Define the function $F(y)$ according to

$$y = \int^{F(y)} P(x) \, dx \quad (8)$$

One checks easily that, if y is a random variable uniformly distributed on $[0, 1]$, $x = F(y)$ is a random variable with a probability distribution $P(x)$. The function F is unfortunately difficult to obtain numerically, except in some very simple cases such as the normal distribution, the exponential law, etc. The inversion procedure required in computing F can be avoided using a classical algorithm which has numerous refinements. Through a simple transformation, the variable x can be assumed to take its value in the interval $[0, 1]$. The following crudest version of the algorithm generates x according to the probability $P(x)$.

- 1) Generate two independent random numbers x and y uniformly distributed on $[0, 1]$.
- 2) If $P(x) < (\text{Sup}P)y$, return to step 1. Otherwise, x is the correctly distributed answer.

Check the validity of this method and compute the frequency of rejections in step 2.

This algorithm generates two random numbers, computes $P(x)$, makes one comparison and performs these operations in a mean number of steps equal to $\text{Sup}P$. Its efficiency increases as the distribution law becomes flatter. Numerous improvements allow its use with full efficiency for any distribution law; the interested reader will find a general reference in the notes.

A generalization of the above ideas has led Metropolis to propose an alternative algorithm to the heat bath method. The configuration $\mathbf{x}$ is again updated site by site according to the following rules.

- M1) Generate a trial value for the field at a given site, chosen at random, leading to a possible configuration $\mathbf{y}$. The probability law $Q(\mathbf{y})$ governing this choice is inessential (except for efficiency considerations) and may even depend on the previous value of the field.
- M2) If $\mathcal{H}(\mathbf{y}) \leq \mathcal{H}(\mathbf{x})$, the trial value $\mathbf{y}$ is accepted and the algorithm terminates.
- M3) If on the contrary the energy increases, $\mathcal{H}(\mathbf{y}) > \mathcal{H}(\mathbf{x})$, generate a random number λ uniformly distributed on $]0, 1[$. The trial configuration $\mathbf{y}$ is accepted if $\lambda < \exp -\beta(\mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{x}))$. Otherwise, the old configuration $\mathbf{x}$ is kept unchanged.

We observe that step M2 would be the same, if we tried to find the ground state of the system. It is in step M3 that an increase in energy which would have been rejected is accepted with a temperature dependent probability. Check that the corresponding transition probability

$$W_j(\mathbf{x}, \mathbf{y}) \propto \begin{cases} Q(\mathbf{y}) & \text{if } \mathcal{H}(\mathbf{y}) \leq \mathcal{H}(\mathbf{x}), \mathbf{x} \neq \mathbf{y} \\ Q(\mathbf{y}) \exp -\beta(\mathcal{H}(\mathbf{y}) - \mathcal{H}(\mathbf{x})) & \text{if } \mathcal{H}(\mathbf{y}) > \mathcal{H}(\mathbf{x}) \\ \sum_{\mathbf{z}, \mathcal{H}(\mathbf{z}) > \mathcal{H}(\mathbf{x})} Q(\mathbf{z})(1 - \exp -\beta(\mathcal{H}(\mathbf{z}) - \mathcal{H}(\mathbf{x}))) & \text{if } \mathbf{x} \equiv \mathbf{y} \end{cases} \quad (9)$$

obeys the conditions required for W as soon as the choice for $Q(\mathbf{y})$ satisfies the ergodicity condition. A weaker condition is to require that a finite power of W satisfies this criterion.

The probability $Q(\mathbf{y})$ is chosen in order to maximize (empirically) the efficiency of the method. For instance, a continuous field φ will be modified by a random quantity $\delta\varphi$ generated in such a way that the mean value $\langle \delta\varphi \rangle$ remains of the order of the fluctuation of the field in the entire system. If one were to generate frequently too large a shift, one would be led to a large rejection rate. Conversely, too small a shift would lead to a small rejection rate, but the system would only evolve very slowly. In both

cases, numerous sweeps of the system would be necessary to obtain independent configurations and thermalization. In practice, a rejection rate of about 50% seems to yield good results.

It is possible to apply the algorithm to the same field several times before going to the next one. If the number of applications is large enough, one creates a secondary Markov chain which converges towards the probability (7). It is thus possible to interpolate between the Metropolis and the heat bath algorithm. This technique may provide a better efficiency.

(1) Computers produce a substitute for random numbers as a deterministic sequence with ergodic properties. In this respect, they deserve the name "pseudorandom" numbers. The classical generators are based on a linear congruential method. They yield a sequence x_n between 0 and 1 using the auxiliary integral dynamical system u_n defined through

$$\begin{cases} x_n = 2^{-p} u_n \\ u_{n+1} = A u_n \end{cases} \quad [\text{mod } 2^p] \quad (10)$$

The numbers $\{x_n\}$ obtained in this fashion provide a "reasonable" pseudorandom sequence if the odd integer A is appropriately chosen. This algorithm is very easy to implement and is furthermore adapted to the structure of computers. However, some biases might appear. In large-scale simulations, it is a good idea to check the random number generators. For instance, the previous one may produce correlations between three consecutive numbers, so that a three-dimensional configuration may be considerably biased. It is therefore necessary to use a more sophisticated procedure. For instance, one extracts at random (using a sequence such as (10)) a number in a table which is renewed using another independent generator (the sequence (10) can again be used, with different A and u_0). In any case, a serious numerical simulation requires statistical tests checking the independency and the distribution of the configurations. Obviously, the replacement of a generator by another one should not modify the results.

(2) In certain instances, the fields can be cast into families such that two fields in the same family do not interact. As an example, for a model with nearest neighbour interactions on a (hyper)cubic lattice, sites can be classified according to the parity of the sum of their coordinates. It is essential to keep this in mind when programming vectorial computers or array processors. Indeed, these machines run very quickly on iterative loops such that

successive operations are independent. This is the case if one splits the loop over fields into two independent ones over even or odd sites. The organization of the storage should also be carefully studied. For instance, a spin of an Ising model can be represented as a single bit. These can be arranged into words (the unit of information in the computer) in such a way that they can be simultaneously treated by single logical instructions. These considerations are important in simulations, where time, storage and cost are the essential limitations.

(3) Another example of difficulty arises when sweeping the fields in succession leading in certain cases to a spurious stabilization of some states. Let us give one example. Consider a two-dimensional gauge theory with an action

$$S = \beta_1 \sum \sigma_1 \sigma_2 \sigma_3 \sigma_4 + \beta_2 \sum \sigma_1 \sigma_2 \sigma_3 \sigma_4 \sigma_5 \sigma_6 \quad (11)$$

which uses rectangular loops with two adjacent plaquettes, and assume β_2 large and negative. This model is equivalent to an antiferromagnetic spin system in an external field. A ground state takes a chessboard aspect (figure 1) in which one out of two plaquette variables is set to -1 . The reader can convince himself that, starting from a configuration with all plaquette variables equal to $+1$, a regular sweep of the link variables done line by line keeps the configuration unchanged. The same phenomenon appears if one starts from the chessboard configuration with β_2 large and positive (a ferromagnetic system). In both cases, the original state remains stable and it is very difficult to analyze the transition line (a first order one) between the two regions. Such a behaviour also appears in simulations of the eight vertex model (or equivalently, the one-dimensional quantum Heisenberg model, or also the two-dimensional Ising model with a four spin interaction). An improvement uses a random choice of the link to be updated, instead of sweeping regularly through the lattice. The results are, however, not yet satisfactory since the plane is partitioned into the two possible degenerate chessboard (staggered) configurations with boundaries which evolve very slowly. It is necessary to have an understanding of the stabilization process in order to overcome this difficulty. The transition between the pure gauge phase (ordered ferromagnetic) and the staggered phase (ordered antiferromagnetic) is expected to occur near $\beta_1 + 4\beta_2 \approx 0$. For definiteness, we choose β_1 positive and therefore β_2 negative. Flipping one link starting from one of the two ordered states costs respectively the energies $-2\beta_2$ and $-6\beta_2$. The process is therefore very improbable. On the other hand, flipping simultaneously two

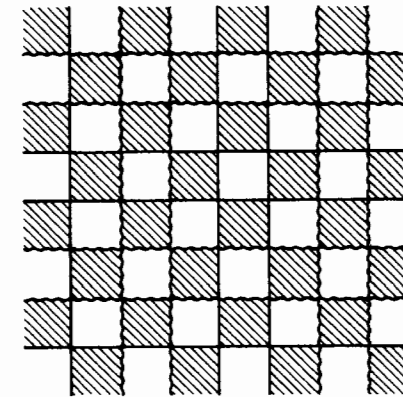


Fig. 1 The fundamental configuration of a gauge system described by the action given in equation (11) for $\beta_2 \ll 0$. Wavy links and shaded plaquettes are set to -1 .

parallel neighbouring links is almost energetically costless. This is the fundamental excitation in this domain, and it is wise to build a simulation based on this double flip updating. Unfortunately, such an algorithm would violate the ergodicity condition if one were not to alternate with a link per link sweeping. A simulation taking into account all these considerations will be considerably more efficient than the original one.

8.1.3 Microcanonical simulations

A technique has been proposed which creates pseudorandom sequences of configurations without using the usual random number generators. The idea is to run along a deterministic path in configuration space, with the constraint that the energy lies in a small band of width δE . This restriction ensures that this path is a random one as in the usual analysis of the ergodic theorem. The process should be compared to a random number generator. In both cases, the ergodicity properties are dependent on the specific procedure used to generate the motion. To rely on this microcanonical technique, it is necessary to perform severe statistical tests to check randomness experimentally. The following method seems to satisfy the above criteria.

Consider a *demon* with a reservoir containing positive energy up to δE . As it visits all sites in succession, it flips each field (or spin) *if it can*. If the energy required to change the local configuration is greater than is available in its reservoir, or if this change yields more energy than the reservoir can contain, the field remains unchanged. If not, the field is flipped and the amount of energy in the tank is correspondingly updated. Such a process can be efficiently implemented, in particular for the Ising model, with due attention paid to the remark on parallel processing of independent families. In practice, one uses a "battalion of demons" visiting simultaneously spins which do not interact.

This procedure should be considered as a pseudorandom generator of configurations, on the same footing as the usual algorithms. At present, it seems that Ising systems are well simulated. The interest of the method is its great efficiency with respect to computing time. Empirical improvement tricks are also used to force randomness. For instance, the demon reservoirs are redistributed randomly at fixed time intervals.

In contradistinction with the previous algorithms, in microcanonical simulations, the energy is kept fixed (up to δE) rather than the temperature. It is thus necessary to measure the temperature. The mean energy available to the demons can be used as a thermometer. Indeed, at thermal equilibrium with the system, their energy distribution should follow the Boltzmann law $\exp(-\beta E_{\text{demon}})$. The reservoir has just one degree of freedom and its mean energy is easy to compute exactly as a function of β . For the Ising system, E_{demon} is a multiple of the flipping energy by an integral factor up to δE . Hence

$$\langle E_{\text{demon}} \rangle = \frac{\sum_{j=0}^{\delta E} j e^{-\beta j}}{\sum_{j=0}^{\delta E} e^{-\beta j}} \quad (12)$$

a relation which can be inverted to yield the temperature β .

8.1.4 Practical considerations

As is clear from the previous considerations, Monte Carlo simulations are in fact an art. The specifics of the implementation depend on the goals (and capabilities) of the performer. We collect here further technical details.

Boundary conditions As the tractable lattice sizes are rather small, boundary conditions play an essential role. For instance, a 10^3 three-dimensional lattice with free boundary conditions has 48.8% of its sites lying on its boundary. This effect increases with dimension and the measurement of bulk quantities becomes problematic. The most common way to handle this problem is to use periodic boundary conditions. From the programmer's point of view, this is easily done using a *modulo* operation, an elementary instruction for computers. Moreover, one takes advantage of the binary structure of data in computers by choosing as linear size a power of 2, replacing this operation by an even more rapid *logical and* instruction. There exists an even more efficient trick which simultaneously treats all dimensions, by choosing "twisted" boundary conditions. The identification of sites in a given direction is conveniently shifted in orthogonal directions, wrapping the lattice as a helix. To be specific, consider a two-dimensional system. In the computer memory, sites are arranged into a one-dimensional array $(1, 1), (1, 2), \dots, (1, L), (2, 1), \dots, (2, L), (3, 1), \dots$, so that the addresses of the neighbours of a given site are obtained by shifting its own address by fixed quantities (± 1 and $\pm L$). The twisted boundary conditions extend this rule to the initial and final sites, the addresses being defined modulo L^2 . This trick greatly simplifies the programming and is very efficient.

One should be aware that periodicity may introduce spurious effects. In some cases, this might be a disaster if one does not pay sufficient attention to its implications. Let us quote some examples.

- i) Mass measurements are performed by fitting two-point correlations $\langle \varphi_0 \varphi_x \rangle$ at large distances, by an exponential falloff $\text{Cst} \exp(-x/\xi)$ if the medium is infinite. With periodic boundary conditions, the correlation function becomes periodic. Taking into account only the first duplication φ_L of φ_0 , the previous formula is replaced by

$$\langle \varphi_0 \varphi_x \rangle \underset{\substack{x \text{ large} \\ 0 < x < L}}{\sim} \text{Cst} \cosh \left(\frac{x - \frac{1}{2}L}{\xi} \right) \quad (13)$$

A typical example of such a behaviour is shown in figure 2.

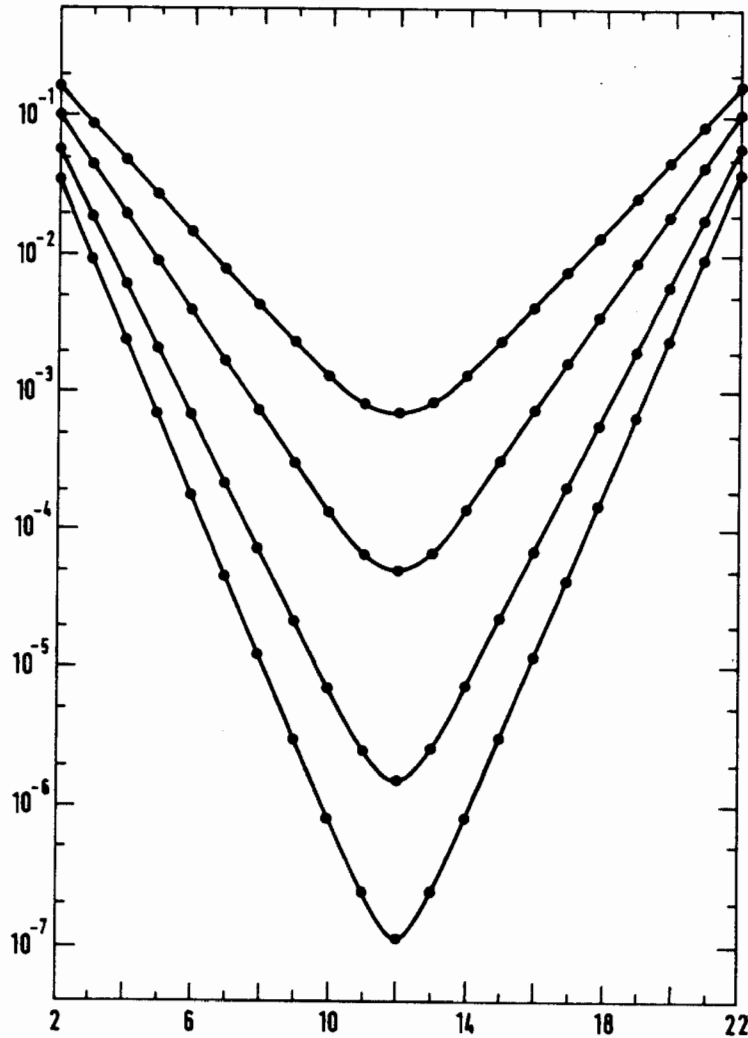


Fig. 2 Typical hyperbolic cosine behaviours in the computation of the pion propagator (from A. Billoire, R. Lacaze, E. Marinari and A. Morel, *Nucl. Phys. B*251 [FS13], 581 (1985)).

- ii) In lattice gauge theory, the static quark potential is derived from the observation of the average of a Wilson operator on a loop of size R in the spatial direction multiplied by T in the time direction, in the limit where $T \rightarrow \infty$ at fixed R . If $T = L$ in a periodic box of size L , the loop breaks into

two independent closed lines, due to the boundary conditions. These are called Polyakov loops. Their use suppresses the problem of the limit $T \rightarrow \infty$ as well as finite size effects arising from the existence of corners in Wilson loops.

- iii) When studying field theories at finite temperature, the partition function $Z = \text{Tr} \exp -\mathcal{H}/T$ at temperature T is expressed in terms of a functional integral with an additional Euclidean “temporal” dimension, for instance

$$Z = \int d\Phi \exp \left\{ -\frac{1}{T} \int_0^{1/T} dt \int d^d \mathbf{x} \mathcal{L}(\Phi) \right\} \quad (14)$$

and a periodic constraint $\Phi(0, \mathbf{x}) = \Phi(1/T, \mathbf{x})$ for the field Φ in the simplest case of a scalar bosonic field. Numerical simulations for such problems are easily performed, using an asymmetric lattice with $N_s \times N_t$ sites, $N_s a_s \gg N_t a_t$, periodic boundary conditions and possibly unequal spacings a_s and a_t . In the continuous limit, the temperature is related to $N_t a_t$ through

$$T = \frac{1}{N_t a_t} \quad (15)$$

- iv) Finally, antiperiodic boundary conditions may also lead to interesting effects. For instance, this allows the measurement of interface energies at low temperature.

Lattice size The choice of the lattice linear size L is obviously related to the effective cost of a simulation, especially in high dimensions, and is often the main limitation to numerical accuracy. However, simulations on lattices of rather small size are not irrelevant. The corresponding calculations are rapid and yield exploratory qualitative runs useful in determining the range of parameters to be chosen in a larger scale simulation. Furthermore, the evaluation of some observables (internal energy and some other bulk quantities) are experimentally found to be rather accurate even for small lattice sizes.

On the other hand, large sizes obviously reduce finite size effects. For technical reasons (initialization of loops, vectorization,...), the computing time needed for the upgrading of a single dynamical variable is also reduced with increasing size. Statistical accuracy is of course better for bulk quantities, since the measurements are

made over more lattice sites. Large sizes are in any case needed to compute correlation functions, but the cost can be reduced by increasing the size only along the field separation.

Note that a simulation is relevant only if the correlation length is less than the linear size. The detailed study of critical phenomena would therefore seem to require huge sizes.

Thermalization time and choice of initial configuration It is always important to estimate the thermalization time, first to reach equilibrium, secondly to know the number of sweeps which should separate two configurations in the generating Markov chain in order that they can be considered as independent. Let us recall the definition of relaxation functions. Starting from an off-equilibrium state characterized by a variation $\delta\theta$ of some thermodynamic quantity θ (temperature, external field, etc.) from its equilibrium value, a nonlinear relaxation function for an observable A can be defined as

$$\phi_A^{\delta\theta}(t) = \frac{\langle A(t) \rangle - \langle A(\infty) \rangle}{\langle A(0) \rangle - \langle A(\infty) \rangle} \quad (16)$$

and allows the introduction of an associated relaxation time

$$\tau_A^{\delta\theta} = \int_0^\infty dt \phi_A^{\delta\theta}(t) \quad (17)$$

Similar quantities can be also computed at equilibrium, using linear response theory

$$\phi_{AB}(t) = \frac{\langle A(0)B(t) \rangle - \langle A \rangle \langle B \rangle}{\langle AB \rangle - \langle A \rangle \langle B \rangle} \quad (18)$$

$$\tau_{AB} = \int_0^\infty dt \phi_{AB}(t)$$

The usual assumption is that the two functions are related through

$$\lim_{\delta\theta \rightarrow 0} \phi_A^{\delta\theta}(t) = \phi_{AB}(t) \quad (19)$$

where B is here the quantity conjugate to θ (e.g. the internal energy if θ is the temperature, the magnetization if it is the external field).

It is expected that the thermalization time is of the order of equilibrium fluctuation times. This is of course useful only if one has some *a priori* knowledge of the latter, which is seldom the

case. Similarly, one can also make some guesses when metastable states are present. Assume that an energy barrier ΔE separates an unstable from a stable state. From the theory of thermally activated processes, one may estimate that the relaxation time behaves as

$$\tau \sim \tau_0 \exp(\beta \Delta E) \quad (\beta \Delta E \gg 1) \quad (20)$$

where τ_0 is a typical time within the metastable phase. Such phenomena occur in first order transitions and are at the origin of hysteresis loops to be discussed later. The energy barrier is here the energy required for the “nucleation” of the stable phase. Similar energy barriers occur in other problems such as spin glasses and lead to extremely long relaxation times, decreasing the efficiency of the Monte Carlo method.

Close to a second order transition, the relaxation time diverges, since fields are correlated in volumes of order ξ^d . One should thus be aware that a critical slowing down occurs near T_c , with its own dynamical critical exponent

$$\tau \sim \left(1 - \frac{T}{T_c}\right)^{-\zeta} \quad (21)$$

For the relaxation function ϕ_{AA} , the relaxation time is nearly proportional to the susceptibility $\chi_{AA} = \beta N[\langle A^2 \rangle - \langle A \rangle^2]$. For a finite system, the divergence is of course rounded (see below). In most cases, in the absence of reliable estimates, it is preferable to fit τ by observing, for instance, during thermalization, how fast an observable reaches its (apparent) equilibrium value. Furthermore, one should be aware that relaxation times may vary considerably according to the observable.

Measuring observables The measurement of certain observables may take a significant amount of computing time as compared to the one used to generate equilibrium configurations. We shall see for instance that fermionic averages imply the inversion of huge $L^d \times L^d$ matrices. The magnitude of statistical fluctuations gives rise to another type of limitation, and measurements become meaningless if average values are too small when compared with the background noise. For instance, derivatives cannot be obtained by finite differences between *independent* measurements.

This is the case of the specific heat $C = dE/dT$. We present in the following a technique which permits access to this quantity. In general, bulk quantities are more easily measured, but their interest is restricted since they are less directly related to renormalized observables of the continuous limit. Masses or correlation lengths are physically more interesting, but require much more care and longer simulations since their determination is based on the way a signal decreases with distance, just as it disappears in the noise.

Statistical errors An estimator $\bar{A}$ for an observable A is the arithmetic average over $\mathcal{N}$ measures. The corresponding statistical error is estimated by

$$\delta A = \sqrt{\frac{1}{\mathcal{N}(\mathcal{N}-1)} \sum_i (A_i - \bar{A})^2} \quad (22)$$

This formula is meaningful only if the measures A_i are uncorrelated, which is not the case when they are generated by a Markov chain. Hence it is important to perform various tests to check statistical independence of the measurements. An empirical method consists in computing $\mathcal{N}(\delta A)^2$ for various sets of measures and verifying that the results are $\mathcal{N}$ -independent.

The value of the relaxation time τ_{AA} (equation (18)) is of course related to this question. One shows that the previous formula should be multiplied by a factor $\sqrt{1 + 2\tau_{AA}}$.

Check this result.

Parametrization of the fields Fields may take their values in a great variety of domains, such as Lie groups (especially in lattice gauge systems). The nontrivial structure of the manifold lengthens the computing time in a significant way. Moreover, the parametrization may have an influence on the required storage capacity.

As an example, let us consider fields taking their values in the group $SU(2)$. We are faced with the following alternatives.

- i) We can parametrize the fields using a 2×2 complex matrix. Storage requires 8 words (floating point numbers) per field. The

multiplication is easily implemented and involves 32 elementary floating point multiplications (in fact, this number can be slightly reduced using some programming tricks).

- ii) Alternatively one can parametrize the fields as points on the unit sphere S_3 in four-dimensional space as $U = a_0 + i\mathbf{a} \cdot \vec{\sigma}$, with $a_0^2 + \mathbf{a}^2 = 1$, a_0 and $\mathbf{a}$ real. Now, only 4 words are required. The group multiplication is a bit lengthier to program ($UU' = (a_0 a'_0 - \mathbf{a} \cdot \mathbf{a}') + i(a_0 \mathbf{a}' + a'_0 \mathbf{a} + \mathbf{a} \wedge \mathbf{a}') \cdot \vec{\sigma}$), but requires only 16 elementary multiplications.

The second parametrization is better from all points of view and should be preferred. It has moreover the additional interest (Creutz, 1980) of permitting implementation of the heat bath algorithm. The Boltzmann weight for a field U appears as $\exp \beta \text{Tr} UV$ where V is a sum of $SU(2)$ matrices and admits the same parametrization $V = b_0 + i\mathbf{b} \cdot \vec{\sigma}$ with no constraint on $b^2 = b_0^2 + \mathbf{b}^2$. The draught of U is taken according to the measure

$$dP(U' = UV^{-1}) \propto d\mathbf{a}' \sqrt{1 - a_0'^2} \exp(\beta b a_0') da_0$$

The random choice of a_0' is governed by an exponential and is performed using the usual algorithms, corrected by a rejection procedure to take into account the square root factor. Then the direction of $\mathbf{a}'$ is selected randomly and the group element can be reconstructed. This yields a very efficient implementation of the heat bath algorithm for the $SU(2)$ group.

The structure of higher $SU(n)$ groups ($n \geq 3$) does not allow such simplifications. The manifold structure of the group is more intricate, and it is presently better to use a direct parametrization in terms of $n \times n$ complex matrices, although the required storage capacity is more important.

A sufficiently dense discretization of the domain on which the fields take their values might seem an interesting possibility for decreasing the storage capacity (one integer being sufficient to describe the field) and the computing time (using an internal multiplication table). However, one should be sure that the properties of the system are unaffected by this simplification.

Consider as an example the $U(1)$ group, well represented by Z_n for large n . Spin systems (with an action $\beta \sum_{(i,j)} \cos[2\pi(m_i - m_j)/n]$, $m_i = 0, \dots, n-1$) as well as their gauge counterparts with Z_n symmetry have been extensively studied and exhibit two transi-

tions for $n > 4$ (figure 3). The interpretation of the intermediate phase is well understood in terms of massless spin wave excitations, as in the XY -model (chapter 4 in volume 1). The transition between the high temperature and the intermediate phases looks like the $U(1)$ transition and belongs to the same universality class. As the temperature decreases, the discrete structure of Z_n becomes apparent and the system undergoes a second transition. As n increases, the lower critical point is shifted towards lower and lower temperature. As a consequence, the $U(1)$ model can be approximated by a discrete Z_n clock model provided the temperature is greater than this lower critical temperature.

This technique would become very interesting if it could be efficiently extended to nonabelian continuous groups. Unfortunately, the corresponding discrete subgroups are only in finite number and might be not sufficiently dense.

Consider again $SU(2)$. From the symmetry groups of regular polyhedra (reduction to $SO(3)$ of the desired discrete subgroups of $SU(2)$ by factoring out the center Z_2), one has the following subgroups (chapter 11).

- 1) Q quaternionic group, 8 elements. This is the multiplicative group of $\pm 1, \pm i\sigma_k$, a double covering of the four group realized as rotations of π around three orthogonal axes.
- 2) $\bar{T}$ double covering of the tetrahedral group, 24 elements.
- 3) $\bar{O}$ double covering of the cubic or octahedral group, 48 elements.
- 4) $\bar{Y}$ double covering of the dodecahedral or icosahedral group, 120 elements.

The groups in this list, which of course omits cyclic and dihedral groups (except for case 1), have been studied in four-dimensional lattice gauge theory and exhibit a unique first order transition at a coupling β respectively equal to 1.23, 2.175, 3.21 and 5.9. Recalling that the crossover region occurs near $\beta \approx 2$ in the $SU(2)$ gauge model, one sees that the last two subgroups are satisfactory in a limited range for simulation of the continuous model.

Turning now to $SU(3)$, among discrete subgroups some are known with 108, 216, 648 and 1080 elements. The respective first order transitions lie at $\beta = 2.5, 3.2, 3.43$ and 3.58 . Hence it seems that none of these groups is useful as a substitute for the continuous group, since the interesting crossover region lies near $\beta \approx 5.9$ for the $SU(3)$ system.

8.2 Extraction of results in a simulation

Once a simulation has been programmed, one has at one's disposal a system where parameters (temperature, external field, etc.) can be adjusted and on which various measurements can be made. Beyond the standard treatment of data, we focus on some peculiarities arising in the study of critical properties.

8.2.1 Determination of transitions

The first task is to determine the phase diagram – if it is not already known – and to localize the critical regions. A preliminary exploration is performed by moving back and forth along a path in the phase diagram, while measuring some observable. As an example, consider a lattice gauge model with Z_6 symmetry. The initial temperature is chosen at a rather high value, and the initial configuration is fully disordered, all fields being chosen independently at random. The system is then allowed to evolve during a sufficiently long time to reach thermal equilibrium (heuristically determined by the fact that observables fluctuate without drift around their mean value). Then the Markov process continues as the temperature decreases regularly by a small amount at each sweep, while measurements are performed on the plaquette energy $E = 1 - \langle \text{Tr } U_p \rangle$. Once a sufficiently low temperature is reached, the process is reversed to return to the initial temperature. The resulting curve for the energy is displayed on figure 3. One observes two hysteresis loops, which reveal the existence of two transitions and give a rough determination of their position.

A precise determination of the critical temperatures and of the order of the transitions requires a more careful study. The existence of a hysteresis loop is not characteristic of a first order transition under present circumstances. In the latter case, it is due to the existence of metastable phases in which the system remains until a sufficient free energy fluctuation induces a jump to the stable phase. Observables such as the energy E undergo a discontinuity. For a second order transition, the hysteresis may be due to a very long relaxation time as compared to the computing time. The energy at equilibrium remains continuous as a function of temperature, and the hysteresis loop should shrink with slower

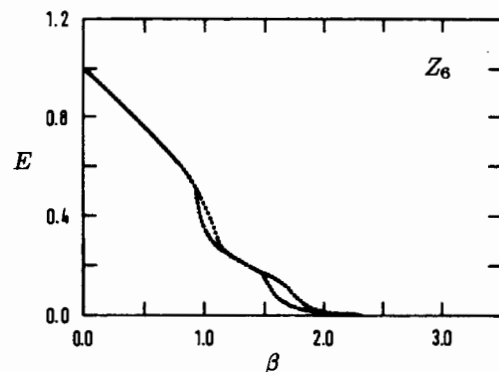


Fig. 3 The plaquette energy in the Z_6 lattice gauge model as the temperature decreases and increases (according to M. Creutz, L. Jacobs and C. Rebbi, *Phys. Rev. D* **20**, 1915 (1979)).

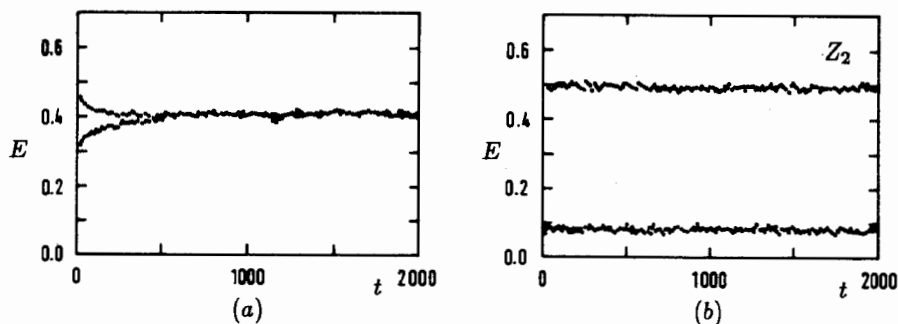


Fig. 4 Evolution of systems at the critical temperature, starting from an ordered or a disordered configuration. (a) Z_6 gauge model (second order). (b) Z_2 gauge model (first order). Same source as figure 3.

changes in the temperature. However, this is not such a good criterion, since the most important factor is the finite size of the system. Curves are rounded and the two characteristic shapes of observables for first and second order transitions become hardly distinguishable. In any case, there are no transitions in finite systems, and therefore the tests to be presented below should be clear and unambiguous enough for the answer to remain valid in the thermodynamic limit.

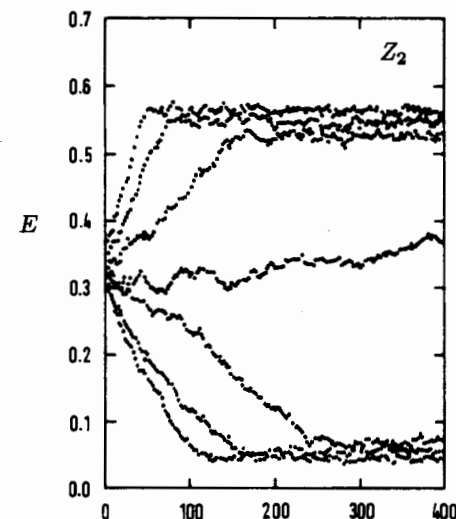


Fig. 5 Time evolution of a Z_2 gauge system initially containing two pure phases as β increases (from top to bottom) from 0.41 to 0.47. Same source as figure 3.

As an example of such a test, one prepares the system in an ordered and a disordered state and let each of these configurations evolve at (or very near) the critical temperature. In the case of a second order transition, both states reach the same equilibrium (figure 4a) while they remain in different phases in the case of a first order transition (figure 4b). Nevertheless, there still remains the possibility that the relaxation time is very long. This can be studied using a sample with an interface separating two pure phases and evolving at various temperatures near the critical one. Figure 5 displays such an experiment for the four-dimensional Z_2 gauge system, with temperatures $\beta = 0.41, 0.42, \dots, 0.47$. The curves are sufficiently characteristic to ascertain that we have here a first order transition with a critical temperature slightly above 0.44 (duality yields the exact value 0.4407).

In the case of a second order transition, one wants to get a handle on critical exponents and amplitudes, or more generally on all universal quantities. Let us estimate the difficulty of such problems for the Ising model. We measure for instance the

susceptibility

$$\chi = \left\langle \left(\sum_i \sigma_i \right)^2 \right\rangle - \left\langle \sum_i \sigma_i \right\rangle^2 \quad (23)$$

or the specific heat

$$C = \left\langle \left(\sum_{(i,j)} \sigma_i \sigma_j \right)^2 \right\rangle - \left\langle \sum_{(i,j)} \sigma_i \sigma_j \right\rangle^2 \quad (24)$$

A technical detail should be mentioned here. As stressed above, these quantities cannot be obtained by numerical differentiation of measured magnetization or energy, due to the statistical uncertainties. This method would not distinguish the signal from the noise. The correct technique is to measure magnetization and energy in various *statistically independent* samples and to estimate their variance as indicated in equations (23) and (24). We insist on the statistical independence, since consecutive configurations of the generating Markov chain are correlated and thus induce an underestimation.

The effect of finite size is illustrated on figure 6 for the three-dimensional Ising system. The critical temperature is determined as the limit of the location of the maximum. The fits of critical exponents are much more delicate and are very sensitive to the determination of T_c . Therefore other methods are needed. A first attempt using additional information from finite size scaling allows the fitting of the measurements by more constrained formulae.

8.2.2 Finite size effects

Scaling laws have been studied in detail in previous chapters. We summarize them here for our present purpose. We consider a system with a finite size L^d depending on only one length L (results can be generalized, e.g. to bars or slices). We have to take into account both the shift in the “critical” temperature $T_c(L)$ and in various exponents ($\gamma(L)$, ...). It is assumed that the susceptibility behaves as

$$\chi_L \underset{t_L \rightarrow 0}{\sim} A(L) t_L^{-\gamma(L)} \quad \text{with } t_L = \frac{T - T_c(L)}{T_c(\infty)} \quad (25)$$

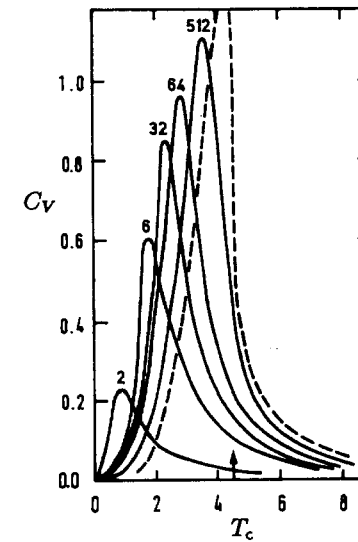


Fig. 6 The three-dimensional Ising model specific heat on finite lattices of various sizes. The dashed curve is the expected thermodynamic limit (from K. Binder, *Physica* **62** 508 (1972)).

For a finite system, all quantities are of course regular and it is *stricto sensu* incorrect to use the term critical temperature. Nevertheless, we define it here as the location of the maximum of χ_L . Hence the corresponding exponent $\gamma(L)$ vanishes. In other cases, such as slices, there exists a nontrivial exponent $\gamma(L)$. For instance, when in the slice $d - 1$ dimensions are large and can be considered in practice infinite as compared to the others, $\gamma(L)$ interpolates between the value $\gamma = \gamma(\infty)$ for the d -dimensional bulk system and $\gamma^* = \gamma(1)$ of the $(d - 1)$ -dimensional one.

As the temperature approaches $T_c \equiv T_c(\infty)$, the correlation length ξ increases first according to its expected power law $t_\infty^{-\nu}$ until it reaches the order of magnitude L , where finite size effects make their appearance. Hence, the displacement of the critical temperature is expected to be such that $L \approx \xi$. The behaviour of the temperature shift is thus estimated to be

$$\frac{T_c(L) - T_c}{T_c} \sim L^{-\theta} \quad \theta = \frac{1}{\nu} \quad (26)$$

It is natural to expect that the only relevant variable governing the system is, besides the temperature, $L/\xi \sim Lt^\nu$. This leads to the scaling behaviour (Fisher 1969)

$$\chi_L(T) \sim L^\omega X(L^\theta t_L) \quad (27)$$

This hypothesis is confirmed below the upper critical dimension (chapter 4 in volume 1) from a general analysis. It can be proved to hold for some solvable models and is observed in low dimensional precise simulations. As $x \rightarrow \infty$, one should recover the critical behaviour obtained in the thermodynamic limit $L \rightarrow \infty$. This leads to the following two conditions on the form factor $X(x)$

$$X(x) \underset{x \rightarrow \infty}{\sim} A_\infty x^{-\gamma} \quad (28)$$

and

$$\omega = \theta\gamma = \frac{\gamma}{\nu} = 2 - \eta \quad (29)$$

Conversely, for fixed L , the behaviour of $X(x)$ near $x = 0$ is governed by equation (25), and one finds

$$X(x) \underset{x \rightarrow 0}{\sim} A(L) L^{(\gamma(L)-\gamma)/\nu} x^{-\gamma(L)} \quad (30)$$

In the case of a finite system, $\gamma(L)$ vanishes and $X(0)$ is finite. The height of the peak is therefore predicted to behave as

$$\chi_{L,\max} \sim L^{-\gamma/\nu} \quad (31)$$

Relations (26) and (31) are crucial in obtaining a reasonable accuracy in the determination of critical exponents using data such as those displayed on figure 6.

Logarithmic divergences $\chi \sim \ln t$ are not correctly taken into account in the previous scaling laws. Generalize the above derivation to this case. Formula (27) should be subtracted at some fixed temperature T_0 , by writing for instance

$$\chi_L(T) \sim L^\omega [X(L^\theta t_L) - X(L^\theta t_0)] + \chi_L(T_0) \quad (32)$$

As a result, one finds that the height of the peak behaves logarithmically.

The method also allows information about interface energies to be extracted. Assume for instance antiperiodic boundary conditions in a given direction. At low temperature, an interface will appear (a surface crossing frustrated links) with an "area"

$S(L) = L^{d-1}$, of order $1/L$ with respect to the volume $V(L) = L^d$ of the system. The free energy admits an expansion

$$\mathcal{F}(T, L) = V(L)\mathcal{F}_\infty(T) + S(L)\mathcal{F}_s(T) + \mathcal{O}(1/L^2) \quad (33)$$

where $\mathcal{F}_\infty(T)$ is the bulk free energy per unit volume in the thermodynamic limit and $\mathcal{F}_s(T)$ is the interface free energy per unit area. The singularities of these functions differ, but are related if the scaling law (27) is valid. Indeed, writing one more term in the large x expansion (28) of the form factor pertaining to the specific heat, with γ replaced by α and ω by α/ν

$$X(x) = X_\infty x^{-\alpha} + Y_\infty x^{-\phi} + \dots \quad (34)$$

and inserting into formula (33), one derives the critical exponent α_s for the interface free energy

$$\alpha_s = \phi = \alpha + \nu = 2 - \nu(d-1) \quad (35)$$

Write the inequalities to be satisfied by the exponents in order that the previous expansions make sense. In particular, one finds $\theta \geq 1$. The limiting logarithmic case $\theta = 1$ can be also studied, allowing an estimate of the corresponding amplitude $A_s(\infty)$. A value $\theta < 1$ would correspond to long-range forces, and the splitting between volume and surface terms is no longer valid.

In practice, when one fits measurements, the most sensitive parameter is the critical temperature, and an error in its determination induces important variations of the critical exponents, which are therefore rather imprecise. The scaling laws have introduced additional constraints which improve the situation. However, a correct determination of the exponents following the previous method implies very precise measurements on samples as large as possible, and hence expansive simulations. An alternative, ingenious and very promising technique has been devised to treat efficiently the critical regime.

8.2.3 Monte Carlo renormalization group

Chapter 4 (volume 1) has been devoted to elementary aspects of the renormalization group in real space. The block spin technique seems well adapted to simulation, and we discuss now its practical

implementation. We reproduce once more the essential formulae needed in the discussion of Ising like systems.

Let us choose a set of operators $\{X_\alpha\}$, which, for practical reasons, should be as small as possible, leading to fast numerical computations, and for theoretical reasons, as large as possible to describe a general action, given the constraints of symmetry and short-range interactions. The action is expressed as a combination of the X_α 's with a set $\{g_\alpha\}$ of coupling constants. It reads

$$S(\{\sigma\}; g_\alpha) = \sum_\alpha g_\alpha X_\alpha(\{\sigma\}) \quad (36)$$

One also defines block variables $\mu_i(\{\sigma\})$, using for instance the majority rule

$$\mu_i = \text{Sign} \left(\sum_{j \in \text{block } i} \sigma_j \right)$$

which is easily implemented on computers. If a sum in a block vanishes, μ_i is chosen at random. The system $\{\mu_i\}$ is hence derived numerically from a given configuration $\{\sigma_i\}$, and both systems describe the same physics in two different scales, with a ratio λ given by size of the blocks. Recall, from chapter 4, that the two corresponding actions are related through

$$\exp S(\{\mu\}; g_\alpha^{(1)}) = \sum_{\{\sigma\}} \left(\prod_i \delta(\mu_i - \mu_i(\{\sigma\})) \right) \exp S(\{\sigma\}; g_\alpha^{(0)}) \quad (37)$$

This relation is here only approximate since the set of operators X_α is by necessity finite. The fixed point g_α^* of the transformation $g^{(0)} \rightarrow g^{(1)}$ corresponds to the critical point, near which the linearized version of the transformation reads

$$g_\alpha^{(1)} - g_\alpha^* = \sum_\beta T_{\alpha\beta} (g_\beta^{(0)} - g_\beta^*) \quad (38)$$

The *Monte Carlo Renormalization Group* method (Swendsen, Wilson) gives numerical access to the matrix T rather than to the coupling constants of the decimated action. This yields directly the critical exponents, without the need for a precise *a priori* knowledge of the critical temperature. Referring to the discussion in chapter 4 and to figure 4.2, the approach to the critical point depends on the magnitude of the eigenvalues of T with

respect to unity. In the basis diagonalizing T , relevant operators are those corresponding to eigenvalues of modulus greater than unity. The associated coupling constants will flow away from the fixed point. In particular, the temperature is associated to the largest eigenvalue $y_{\max}$ in the even sector, and thus the reduced temperature satisfies

$$t^{(1)} \sim t^{(0)} y_{\max} \quad (39)$$

Recalling that the correlation lengths in both configurations differ by a factor λ

$$\xi^{(1)} = \frac{\xi^{(0)}}{\lambda} \quad (40)$$

we derive, from the scaling behaviour $\xi \sim t^{-\nu}$ and from equation (39), the critical exponent ν in the form

$$\nu = \frac{\ln \lambda}{\ln y_{\max}} \quad (41)$$

To obtain other exponents such as γ , one needs the introduction of an external field (i.e. more generally odd operators X_α).

In practice, one generates configurations $\{\sigma^{(0)}\}$ very close to the critical temperature. Then the decimated configurations $\{\sigma_i^{(k)} = \mu_i(\{\sigma^{(k-1)}\})\}$ are determined and one measures the corresponding values $X_\alpha^{(k)}$ of the operators X_α . Performing statistical averages over independent configurations $\{\sigma^{(0)}\}$, one obtains the quantities

$$\frac{\partial}{\partial g_\alpha^{(p)}} \langle X_\beta^{(q)} \rangle = \langle X_\alpha^{(p)} X_\beta^{(q)} \rangle - \langle X_\alpha^{(p)} \rangle \langle X_\beta^{(q)} \rangle \quad (42)$$

The determination of T immediately follows, since equation (38) can be rewritten as

$$\frac{\partial}{\partial g_\alpha^{(k)}} \langle X_\beta^{(k+1)} \rangle = \sum_\gamma T_{\alpha\gamma} \frac{\partial}{\partial g_\gamma^{(k+1)}} \langle X_\beta^{(k+1)} \rangle \quad (43)$$

We emphasize again that an important aspect of this method is that critical exponents are obtained independently from a precise estimate of the critical temperature. The numerical study is to be completed by examining the stability of the results with respect to the number of decimations. One has here a control on the shift between the initial and the critical temperature. If the results

are not sufficiently stable, the choice of the operator set $\{X_\alpha\}$ is inadequate and should be modified.

To illustrate the method, we display some of the data obtained by Swendsen for the two-dimensional Ising model. The set of operators are chosen to be

Even operators	Odd operators
1- Nearest neighbours	1- Magnetic field
2- Next nearest neighbour (11)	2- Three spins on a plaquette
3- Four spins coupled on a plaquette	3- Three aligned spins
4- Third neighbour (20)	4- Three spins (00),(10),(21)
5- Fourth neighbour (21)	
6- Four spins along a rectangle	
7- Fifth neighbour (22)	

The symbols such as (10), (11), ..., indicate the relative coordinates of the interacting neighbours. Starting from a 64×64 lattice with only the usual Ising coupling fixed at its (known) critical value, tables Ia, Ib display the leading and subleading critical eigenvalue exponents in the absence of an external field (only even operators are considered). The exact values for these exponents are respectively $2 - \nu = 1$ and -1 . The number of couplings introduced increases along the vertical direction. Clearly, a single coupling is insufficient, while the introduction of more than two terms brings little improvement in table Ia. Table Ib pertaining to the second largest eigenvalue requires of course more operators. Note that this example is particularly favorable, although one expects that the method will work with only a few adequately chosen couplings. The number of iterations (by 2×2 blocks) varies horizontally. The instability of the flow is visible on right side of the table and is due here to finite size effects, while, in a more general case, it could also be due to a discrepancy with the critical temperature. Table Ic displays the same analysis for the magnetic exponent (i.e. the dimension of the magnetic field with an exact value $2 - \beta = 1.875$), using odd interactions, and implies similar comments. The study is also completed by using other lattice sizes (the results, not given here, confirm the validity of the method). A further consistency check uses the same data, but considers two consecutive iterations as a single one with scaling factor λ^2 . If there is a discrepancy, this means that neglected interactions are important (note that unfortunately agreement does not mean that the method is really correct).

Table I. Monte Carlo Renormalization Group determination of exponents for a two-dimensional 64×64 Ising system, according to the work of Swendsen quoted in the notes.

(a)	1	2	3	4
1	0.912(2)	0.963(4)	0.957(2)	0.940(7)
2	0.967(3)	0.999(4)	0.998(2)	0.993(6)
3	0.968(3)	1.001(4)	0.999(2)	0.992(6)
4	0.969(4)	1.002(5)	0.999(2)	0.988(5)
5	0.969(4)	1.001(5)	0.997(2)	0.990(5)
6	0.969(3)	1.001(5)	0.997(2)	0.988(5)
7	0.969(5)	1.000(5)	0.997(2)	0.984(4)
(b)	1	2	3	4
2	-2.19(3)	-2.20(4)	-2.20(3)	-2.32(8)
3	-2.17(3)	-2.34(6)	-2.56(6)	-2.56(22)
4	-1.11(3)	-1.14(3)	-1.16(3)	-1.35(20)
5	-0.94(3)	-0.94(3)	-0.91(2)	-0.80(8)
6	-1.04(3)	-1.04(3)	-0.98(2)	-0.84(9)
7	-1.08(21)	-1.03(4)	-0.99(20)	-0.76(20)
(c)	1	2	3	4
1	1.8810(1)	1.8757(2)	1.8731(4)	1.8706(5)
2	1.8804(1)	1.8758(2)	1.8740(4)	1.8735(7)
3	1.8806(1)	1.8758(2)	1.8740(4)	1.8732(8)
4	1.8808(1)	1.8759(2)	1.8741(4)	1.8737(9)

Errors on the last digit are indicated in parentheses. (a) Leading thermal exponent. (b) Subleading thermal exponent. (c) Magnetic exponent.

- i) Finite size effects are not included in the previous discussion. A more complex procedure allows their elimination. This involves two independent simulations with two lattices differing by a factor λ in size. Coupling constants are adjusted in such a way that the $(n+1)$ th decimation of the larger system and the n th decimation of the smaller one yield identical mean values for the operators X_α . As the physical sizes have to agree, finite size effects are identical and disappear in the final result.
- ii) Several improvements have been proposed. Let us describe for instance a version (Gupta and Cordery) in which the configurations $\{\sigma^{(0)}\}$ are generated with a weight

$$\exp \left[S \left(\left\{ \sigma^{(0)} \right\}; g_\alpha^{(0)} \right) - S \left(\left\{ \sigma_i^{(1)} \right\}; g_\alpha^g \right) \right]$$

with

$$\sigma_i^{(1)} = \mu_i \left(\left\{ \sigma^{(0)} \right\} \right) \quad (44)$$

depending on the block variables. The coupling $g_\alpha^{(1)}$ of the decimated action has been approximated by a guess g_α^g . The benefit is that the introduction of both site and block couplings suppresses the long time correlations due to the diverging correlation length. If the guess is perfect, the block spins are totally uncorrelated

$$\langle X_\alpha^{(1)} \rangle = 0 \quad \langle X_\alpha^{(1)} X_\beta^{(1)} \rangle = N_\alpha \delta_{\alpha\beta} \quad (45)$$

If the guess is not perfect, one has to first order

$$\langle X_\alpha^{(1)} \rangle = \langle X_\alpha^{(1)} X_\beta^{(1)} \rangle \left(g_\beta^{(1)} - g_\beta^g \right) \quad (46)$$

so that it is possible to determine the renormalized coupling with no truncation error as

$$g_\alpha^{(1)} = g_\alpha^g + \frac{\langle X_\alpha^{(1)} \rangle}{N_\alpha} \quad (47)$$

The procedure can be iterated. If the irrelevant eigenvalues of T are small, two or three iterations are sufficient to reach the fixed point action, assumed to be of short range. The determination of the exponents proceeds exactly as above.

8.2.4 Dynamics and the Langevin equation

We have already mentioned the importance of the underlying dynamics generated by the Markov chain used in the Monte Carlo method. However, we have said very little on this subject, which is not easy to analyze theoretically. Another suggestion has been made of a method to generate configurations with better control of the dynamics. This *stochastic quantization* method uses the classical equations of motion with an additional perturbation generated by white noise. The technique has interesting applications for the determination of correlation functions and in fermionic simulations. It was originally proposed as a substitute for the gauge fixing procedure (Parisi, Wu Yong Shi).

Consider a generic field theory described by an action $S[\varphi]$. The classical evolution of fields with a random, time dependent driving term $\eta(t)$ is described by the Langevin equation

$$\dot{\varphi}_i(t) = \frac{\delta S}{\delta \varphi_i(t)} + \eta_i(t) \quad (48)$$

The first term is driving the system towards classical equilibrium. The excitation is chosen to be a random Gaussian variable

$$\begin{aligned} \langle \eta_i(t) \rangle &= 0 \\ \langle \eta_i(t) \eta_j(t') \rangle &= \delta_{ij} \delta(t - t') \\ &\dots \end{aligned} \quad (49)$$

The notation $\langle \cdot \rangle$ denotes here the stochastic average with respect to the η 's, while the usual field theoretic expectation values will be denoted by an overbar. The Fokker-Planck equation satisfied by the probability distribution $P(\varphi, t)$ for the field configurations can be easily derived from (48) and reads

$$\frac{\partial}{\partial t} P(\varphi, t) = \frac{\delta}{\delta \varphi_i} \left[- \frac{\delta S}{\delta \varphi_i} P(\varphi, t) + \frac{\delta P(\varphi, t)}{\delta \varphi_i} \right] \quad (50)$$

In the stationary limit $\dot{P} = 0$, one recovers the usual Boltzmann weight $\exp S$. Hence any observable A satisfies

$$\langle A \rangle \xrightarrow[t \rightarrow \infty]{} \bar{A} \quad (51)$$

If the action is invariant under a symmetry group, such as local invariance in lattice gauge systems, the result is valid for the invariant part of the operators A . Indeed, "longitudinal" components (i.e. corresponding to gauge transforms) have a vanishing driving term in equation (48). As a result, they undergo a random walk, and this yields divergent terms proportional to t (chapter 1). This divergence of course does not affect the measurement of invariant quantities. These properties were the reason for the original proposal and also apply in the context of continuous field theory.

Simulations of the Langevin equation are easy to do. Using random draughts of the noise $\eta_i(t)$, a configuration $\{\varphi\}$ is deterministically followed in time. Mean values of observables are then computed from a time average

$$\bar{A} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt A \quad (52)$$

However, as in all numerical computations, time should be discretized. Denoting the elementary time step by ε , a discrete

version of the Langevin equation reads

$$\varphi_i(\tau+1) = \varphi_i(\tau) + \varepsilon \frac{\delta S}{\delta \varphi_i} + \sqrt{\varepsilon} \eta_{i,\tau} \quad (53)$$

where we have used a reduced integral time variable $\tau = t/\varepsilon$ and rescaled η by a factor $\sqrt{\varepsilon}$, in such a way that the $\eta_{i,\tau}$ are independent normal Gaussian variables

$$\langle \eta_{i,\tau} \eta_{i',\tau'} \rangle = \delta_{ii'} \delta_{\tau\tau'} \quad (54)$$

The computer running time is proportional to ε^{-1} . It is thus interesting to use ε 's as large as possible. We have therefore to discuss the corresponding discretization effects. In particular, the limiting probability distribution of the fields will be shifted from the Boltzmann weight $\exp S$ to $\exp S'$ related to a different effective action S' . Deriving the analog of equation (50) for the discretized case and expanding up to first order in ε , one obtains a continuous Fokker-Planck equation with an effective action given by

$$S' = S + \frac{1}{4}\varepsilon \sum_i \left[2 \frac{\delta^2 S}{\delta \varphi_i^2} + \left(\frac{\delta S}{\delta \varphi_i} \right)^2 \right] + \mathcal{O}(\varepsilon^2) \quad (55)$$

It is possible to redefine the fields so that the quadratic term in the coefficient of ε disappears

$$\begin{aligned} \varphi'_i &= \varphi_i - \frac{\varepsilon}{4} \frac{\delta S}{\delta \varphi_i} + \dots \\ S'' &= S + \frac{\varepsilon}{4} \sum_i \frac{\delta^2 S}{\delta \varphi_i^2} + \dots \end{aligned} \quad (56)$$

In the $\lambda\varphi^4$ theory, the effect is interpreted (to this order) as a shift in the bare square mass $m^2 \rightarrow m^2 + \frac{1}{4}\lambda\varepsilon$. Irrelevant terms also affect the bare coupling constant.

- i) The difficulty in implementing this algorithm for a pure $SU(n)$ lattice gauge theory lies in the nontrivial structure of the manifold on which the fields take their values. The stochastic equation of motion which replaces (48) has to apply to an element in the tangent space, i.e. the Lie algebra. Let us denote by t_α the generators of the Lie algebra, normalized according to $[t_\alpha, t_\beta] = if_{\alpha\beta\gamma} t_\gamma$ and $\text{Tr} t_\alpha t_\beta = \frac{1}{2} \delta_{\alpha\beta}$. If ∂_α stands for the derivative with respect to the group parameter corresponding

to t_α , the simplest discretized Langevin equation reads

$$\begin{aligned} U_{ij}(\tau+1) &= e^{-i\lambda_{ij} \cdot t} U_{ij}(\tau) \\ \lambda_\alpha &= \varepsilon \partial_\alpha S + \sqrt{\varepsilon} \eta_\alpha \end{aligned} \quad (57)$$

For the Wilson action $S = (-\beta/2n) \sum \text{Tr}(U_p + U_p^\dagger)$, this leads to

$$i\lambda_{ij} \cdot t = \frac{\varepsilon\beta}{4n} \sum_{kl} (U_{ijkl} - U_{ijkl}^\dagger) - \frac{\varepsilon\beta}{4n^2} \sum_{kl} \text{Tr}(U_{ijkl} - U_{ijkl}^\dagger) + i\sqrt{\varepsilon} H_{ij}(\tau) \quad (58)$$

where H is a traceless Hermitian random noise matrix satisfying

$$\langle (H_{ij}(\tau))_{ab} (H_{kl}(\tau'))_{cd} \rangle = (\delta_{ac} \delta_{bd} - \delta_{ab} \delta_{cd}/n) \delta_{(ij),(kl)} \delta_{\tau,\tau'} \quad (59)$$

After some manipulation, the discretization of the Langevin equation leads to a Fokker-Planck equation, with an effective action of the form

$$S' = (1 + \frac{1}{12}\varepsilon n) S + \frac{1}{4}\varepsilon \sum_\alpha \{ 2\partial_\alpha^2 S - (\partial_\alpha S)^2 \} + \mathcal{O}(\varepsilon^2) \quad (60)$$

This formula can again be interpreted as a rescaling of the fields together with a shift of the coupling, according to

$$\begin{aligned} U &\rightarrow e^{-\frac{1}{4}i\varepsilon \partial S \cdot t} U \\ \beta &\rightarrow \beta \left(1 - \varepsilon \frac{5n^2 - 6}{12n} \right) \end{aligned} \quad (61)$$

- ii) Is it possible to design a more sophisticated discrete Langevin equation leading to errors of order higher than ε ?

8.3 Simulating fermions

Grassmannian integrals have been defined in an algebraic way and cannot be computed as usual integrals. The concept of sampling the integration space becomes meaningless, and the existence of "Grassmannian chips" remains a joke. Fermions cannot be simulated using the previous Monte Carlo techniques, and it is necessary to perform the integration over Grassmannian variables in closed form using the properties described in chapter 2 (volume 1). The corresponding numerical simulations therefore have to deal with the estimation of determinants and inverses of huge matrices in the best of all cases where one has reduced the fermionic action to a quadratic one.

Quantum Chromodynamics is one of the domains where fermionic simulations are used extensively. At present, the corresponding programs are extremely heavy and costfull and reach the capacity limits of existing computers. Parts of the following discussion might soon become obsolete.

Using the language and notation of lattice gauge theories, we consider a model with bosonic gauge fields U and fermionic fields $\bar{q}_i, q_i$ interacting through the action

$$S = S_g(U) + \sum_{ij} \bar{q}_i D_{ij}(U) q_j \quad (62)$$

with only quadratic fermionic terms. More complicated systems such as the Gross-Neveu model (see chapter 10) can be handled within this scheme at the price of the introduction of additional bosonic fields. The occurrence of flavor, spinor, ... indices is immaterial for the forthcoming analysis. Their introduction would only complicate the expressions. We omit them and leave it to the reader to restore them for specific applications.

Performing the Grassmannian integrals, one finds that the gauge fields evolve according to an effective action

$$S_{\text{eff}}(U) = S_g(U) + \ln \det D(U) \quad (63)$$

The additional difficulty in the Monte Carlo sampling for the gauge configurations is the computation of the determinant term. Note that it is nonlocal and introduces long-range interactions between gauge fields. Once the configurations have been generated, another difficulty arises in the measurement of the observables involving Fermi fields. We have

$$\langle f(U) \rangle = Z^{-1} \int DU \exp S_{\text{eff}}(U) f(U) \quad (64a)$$

$$\langle \bar{q}_i q_j f(U) \rangle = Z^{-1} \int DU \exp S_{\text{eff}}(U) f(U) (D(U)^{-1})_{ij} \quad (64b)$$

$$\begin{aligned} \langle \bar{q}_i q_j \bar{q}_k q_l f(U) \rangle &= Z^{-1} \int DU \exp S_{\text{eff}}(U) f(U) \\ &\times \left[(D(U)^{-1})_{ij} (D(U)^{-1})_{kl} - (D(U)^{-1})_{il} (D(U)^{-1})_{kj} \right] \quad (64c) \end{aligned}$$

and the problem is to estimate the inverse matrix $D(U)^{-1}$, depending on the configuration $\{U\}$ and occurring in the average over fermionic observables. In a first step, we present an approx-



Fig. 7 The two terms entering in the determination of meson masses.

imation which greatly simplifies the simulation and amounts to neglecting the virtual quark pairs.

8.3.1 The quenched approximation

The most drastic assumption is that the term $\ln \det D(U)$ in (63) has little influence in the sampling of gauge configurations, at least in a qualitative, and perhaps even semiquantitative sense (this is called the “quenched approximation”). The practical interest of this assumption is obvious, since the evaluation of the determinant is very long. Another interest is that the effective action becomes independent on the details of the fermionic model (choice for the discretization of fermion fields, number of flavours, bare masses of quarks, etc.), and the same configurations can thus be used while varying the parameters and details of the fermionic part of the model (still necessary to compute averages such as those in equations (64b)–(64c)), allowing an accurate comparison without statistical errors.

An interpretation of the quenched approximation is that it amounts to a decoupling of virtual pairs for large internal quark masses. This is easily understood if one splits the quark propagator into kinetic and mass terms $D(U) = \Delta(U) + m$. Expanding in $1/m$, the effective action reads (up to an unessential additive constant)

$$S_{\text{eff}}(U) = S_g(U) - \sum_k \frac{1}{k} \left(\frac{-1}{m} \right)^k \text{Tr} \Delta^k(U) \quad (65)$$

so that at face value the determinant becomes negligible when m tends to infinity. This naive argument neglects of course renormalization effects and may become totally wrong in the presence of anomalies. The latter corresponds precisely to cases where the combination of a vanishing coefficient (an inverse power of m) and a divergent expression conspire to yield a finite result in

one (or several) term(s) in the above expansion. We set aside this possibility. Nevertheless the quenched approximation becomes questionable whenever the creation or annihilation of dynamical quark pairs is expected to play an important role. For instance, let us consider the meson propagator $\langle (\bar{q}_i q_i)(\bar{q}_j q_j) \rangle$. The two terms entering in the r.h.s. of equation (64c) are graphically represented on figure 7. The diagram (b) generates fermion loops by iteration, so that its contribution is strongly related to the terms neglected in the quenched approximation. Due to selection rules on angular momentum or internal quantum numbers, the contribution of this graph vanishes in most cases of interest, such as when studying the ρ or π meson propagators. It remains however important for states such as the η meson which is a singlet for the flavour group (this case is also related to the above mentioned anomaly). Such states have relatively high masses, and one can therefore expect that the approximation is reasonable in the spectroscopy of light mesons (π , ρ , K , K^* , ...) and light baryons (N , Δ , Λ , Σ , ...).

Formula (65) yields a procedure to compute systematic corrections in $1/m$ to the quenched approximation. However, the resulting *hopping parameter* expansion seems to be difficult to handle and has not yet provided accurate improvements.

Fermionic simulations using the quenched approximation still have to face the determination of the inverse matrix $G = D(U)^{-1}$. In practice, this task takes an important part of the total computer time.

Two remarks can be used to simplify this matrix inversion. First, it is generally sufficient to compute only a fixed column of the matrix G . Indeed, in many applications, one is interested in determining meson and baryon masses from the decay laws of correlations $\langle (\bar{q}_i \gamma q_i)(\bar{q}_j \gamma q_j) \rangle$ and $\langle (\bar{q}_i \bar{q}_i)(q_j q_j) \rangle$, working at fixed i and varying j . Secondly, the matrix D has a nearest neighbour structure and multiplication by this matrix is thus not expansive. These facts can be used in iterative procedures. The Gauss-Seidel method, valid for a positive definite matrix D , uses the sequence

$$G^{(n+1)} = G^{(n)} + \varepsilon[1 - DG^{(n)}] \quad (66)$$

These equations decouple for each column of G . One expects to reach the fixed point $G = D^{-1}$ of this sequence if ε is sufficiently small.

Study the convergence properties of the sequence, using a basis diagonalizing D . Show in particular that, for lattice gauge simulations performed at rather high coupling β , convergence is ensured if $\varepsilon < Cst m$, that is, in practice, for sufficiently massive quarks. The number of iteration steps required to reach a given precision grows as $1/m$ for small quark mass m .

The technique can be improved in several ways. For instance, one can use a second order relaxation procedure

$$G^{(n+1)} = G^{(n)} + \{\varepsilon + \rho\varepsilon^2[1 - D]\}(1 - DG) \quad (67)$$

In practice, the values of ρ and ε are empirically adjusted to yield the quickest convergence. The Runge-Kutta value $\rho = \frac{1}{2}$ is theoretically the best one to approach the expected result, but slightly larger values of this parameter yield sequences which overshoot for some time the exact result and provide a better estimate when one uses only a small number of iterations.

Another improvement is to rederive from the sequence $\{G^{(n)} - G^{(n-1)}\}$ an estimate of the errors and thus the best value to be chosen for ε at the next iteration step. This leads to the so-called conjugate gradient method.

8.3.2 Dynamical fermions

Let us now abandon the crude approximation of the previous subsection and analyze the available practical techniques to handle the full nonlocal action (62).

Luckily, in the Monte Carlo updating algorithm, it is not necessary to compute at every stage the complete action (62). Only its variation δS_{eff} is effectively required when varying a field $U \rightarrow U + \delta U$. We have

$$\delta S_{\text{eff}}(U) = \delta S_g(U) + \text{Tr} \left[D(U)^{-1} \frac{\delta D}{\delta U} \right] + \mathcal{O}(\delta U^2) \quad (68)$$

We are therefore facing again the problem of obtaining an estimate for $G = D(U)^{-1}$. This computation occurs now very frequently, not only for the measurements of observables as it did in the quenched approximation. We need it at each step of the updating procedure. The most straightforward procedures are the ones described above. These are however very costly. Alternatively,

the pseudofermion method introduces auxiliary complex *bosonic* fields $\{\varphi_i\}$ and uses the identity

$$(D^{-1})_{ij} = \langle \varphi_i^\dagger \varphi_j \rangle = \frac{\int \prod_k (d\varphi_k d\varphi_k^\dagger) \varphi_i^\dagger \varphi_j \exp(-\varphi^\dagger D \varphi)}{\int \prod_k (d\varphi_k d\varphi_k^\dagger) \exp(-\varphi^\dagger D \varphi)} \quad (69)$$

This expression is computed statistically using a standard Monte Carlo method for the pseudofermions φ .

A difficulty arises from the fact that the matrix D may be non-positive definite, so that the integrals in (69) are meaningless. This feature is corrected by using a pseudofermionic action $\varphi^\dagger D^\dagger D \varphi$ which involves now a next to nearest neighbour interaction. The mean value $\langle \varphi_i^\dagger \varphi_j \rangle$ provides the inverse matrix $(D^\dagger D)^{-1}$ from which one extracts the desired quantity $D^{-1} = (D^\dagger D)^{-1} D^\dagger$ by a quick multiplication with a “short-range” matrix.

We deal now with two Monte Carlo simulations, one for the pseudofermions inside the other one for gauge fields. In principle, δS_{eff} has to be computed every time with an updated gauge field and the pseudofermion algorithm could be very slow. At present, one generally computes δS_{eff} only once at the end of a full sweep. This approximation is expected to be accurate enough provided the shift δU remains small.

Extend the method of stochastic quantization to take into account the fermions and generalize formula (58) for the updating of gauge fields to include the effect of the fermion determinant.

8.3.3 Hadron mass calculation in lattice gauge theory

We conclude this chapter with a quick overview of the present attempts to predict the mass spectrum of light “elementary” particles from numerical simulations of lattice quantum chromodynamics. Let us first recall some more details about the model. Strong interactions are described in terms of a $SU(3)$ “colour” gauge group coupling several families of quarks. We restrict ourselves to the light, nonstrange hadrons and thus consider only the lowest mass quarks. These obey a $SU(2)$ isotopic symmetry, with respect to which they form a doublet of “up” and “down” states. Observed particles are bound states which are colour singlets due

Table II. Experimental characteristics of some light particles. We list isospin I , G -parity when relevant, angular momentum J , parity P and charge conjugation C . The quark content is indicated for the mesonic states.

Particle	$I^G(J^P)C$	Mass (MeV)	expected in
π	$1^-(0^-)+$	139	$\bar{q}\hat{\gamma}_5\tau q$
ρ	$1^+(1^-)-$	770	$\bar{q}\gamma_\mu\tau q$
ω, φ	$0^-(1^-)-$	783, 1020	$\bar{q}\gamma_\mu q$
η, η'	$0^+(0^-)+$	549, 958	$\bar{q}\hat{\gamma}_5 q$
A_1	$1^-(1^+)+$	1270	$\bar{q}\hat{\gamma}_5\gamma_\mu\tau q$
B	$1^+(1^+)-$	1235	$\bar{q}[\gamma_\mu, \gamma_\nu]\tau q$
δ	$1^-(0^+)$	980	$\bar{q}\tau q$
S	$0^+(0^+)+$	975	$\bar{q}q$
N	$\frac{1}{2}\left(\frac{1}{2}^+\right)$	939	
Δ	$\frac{3}{2}\left(\frac{3}{2}^+\right)$	1232	

to confinement. The corresponding operators in the quark language are most likely nonlocal. However, as they have a local component involving only quark fields at the same location, one studies for simplicity the correlation lengths of the corresponding operators, $\bar{q}_i\hat{\gamma}_5\tau q_i$ for mesons, $q_i q_i q_i$ for baryons. Spinorial and isospin indices are arranged to yield a state of definite spin and isospin, taking into account the limitations due to the finite lattice rotation group. For instance, the isospin 1 pseudoscalar π meson is obtained from the operator $\bar{q}_i\hat{\gamma}_5\tau q_i$, τ being the three 2×2 Pauli matrices acting on isospin quark indices. We display in table II the characteristics of the most important light particles. A more complete study has also to include heavy quarks, and in particular the strange quark which appears in rather light strange particles such as the pseudoscalar isotopic doublet K (495 MeV).

In the spectroscopy of light hadrons, chirality plays a very important role although it is only an approximate symmetry. In the limit of massless quarks, the Goldstone theorem predicts a vanishing mass for the pion. Its low value (139 MeV) as compared to the order of magnitude 1 GeV for other particles, as well as the experimental verification of current algebra predictions show the importance of spontaneously broken chiral invariance. Monte Carlo simulations are presently carried out using concurrently

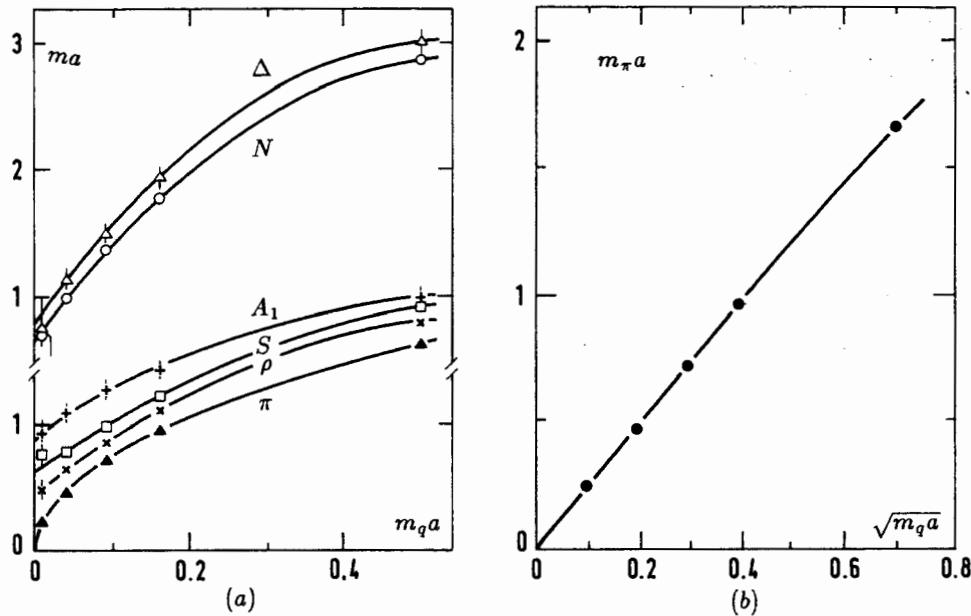


Fig. 8 (a) Measured masses of light states in the staggered fermion model as a function of the quark mass. (b) Pion mass as a function of the square root of the quark mass, showing the prediction of chiral invariance (from D. Barkai, K.J.M. Moriarty and C. Rebbi, *Phys. Lett.* **156B**, 385 (1985)).

either the staggered fermionic action (6.157), or Wilson's action (6.150). The first one has, as was discussed, a discrete chiral invariance, and it is therefore expected that the pion mass vanishes with the quark mass. Figure 8 shows this behaviour.

On the other hand, using Wilson's action, chiral invariance is explicitly broken by the lattice regularization but one expects to recover it in the continuous limit. With r the parameter occurring in this action, this will however not occur for the naive value $1/2dr$ of the hopping parameter $K = 1/[2(ma + dr)]$ as m goes to zero, due to renormalization effects on the bare quark mass. These effects are not under control in the present case for lack of chiral symmetry. The critical value is displaced to K_c , and chiral invariance will hopefully be restored as the hopping parameter approaches K_c from below. This is observed as expected in the simulations displayed on figure 9, where the critical point corresponds to the vanishing of the pion mass.

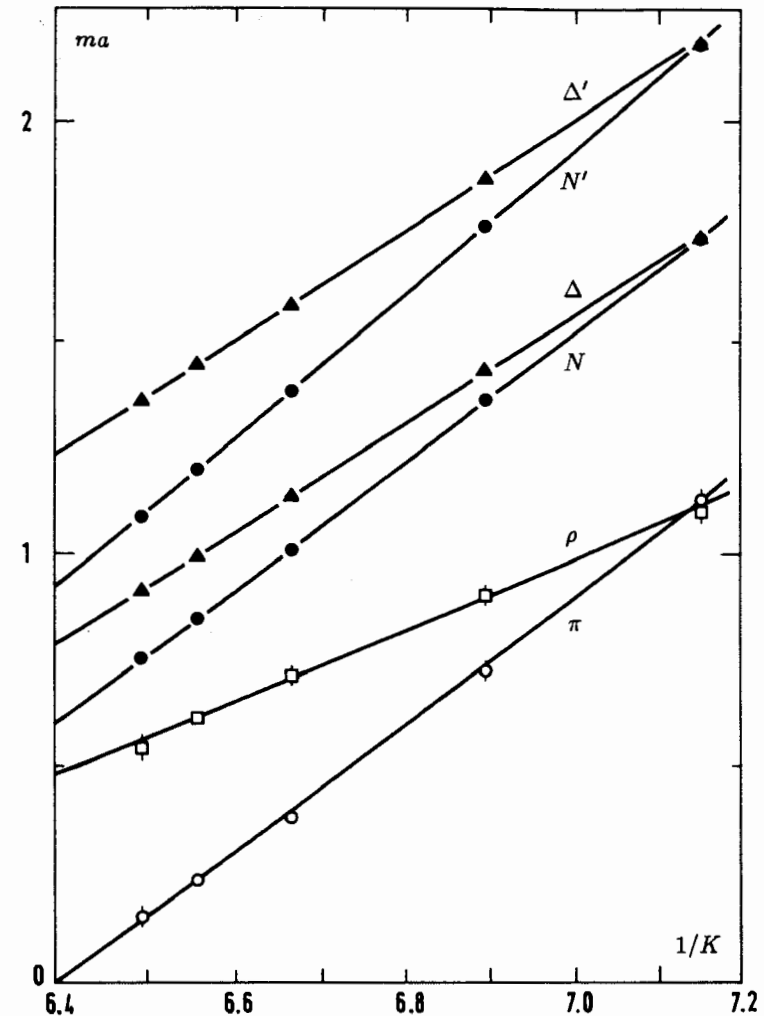


Fig. 9 Hadron masses versus the inverse hopping parameter with the Wilson's fermionic action (from S. Itoh, Y. Iwasaki and T. Yoshie, *Phys. Lett.* **B167**, 443 (1986)).

As the Monte Carlo techniques are in any case extremely expansive, simulations are done extensively only for a fixed gauge coupling constant β_J . In the choice of this parameter, one is squeezed between two opposite constraints. The first one is to have a coupling as low as possible to have a faster convergence rate and to avoid the use of large lattices (the spacing a decreases as β_J

increases due to asymptotic freedom) which are also very onerous. The second constraint is to have a sufficiently high coupling in the scaling window in order to allow a contact with the continuous limit using asymptotic freedom formulae. Unfortunately, an extensive numerical study of the scaling properties in the presence of fermions has not yet been performed. Let us recall that, in the scaling region, the masses remain constant in the lattice mass scale Λ_L given by (6.31). In the case of a $SU(3)$ colour gauge group, the formulae (6.22), valid for pure gauge theory, have to be modified as

$$\begin{cases} b_0 = \frac{1}{(4\pi)^2} (11 - \frac{2}{3}N_f) \\ b_1 = \frac{1}{(4\pi)^4} (102 - \frac{38}{3}N_f) \end{cases} \quad (70)$$

where N_f is the number of flavours, two in the present discussion. Note that this number cannot exceed 16 to ensure asymptotic freedom ($b_0 > 0$). Given the present lack of numerical results about scaling properties, one assumes that the scaling window is the same as in the pure gauge case. String tension data (see figure 16 of chapter 6) yield a lower bound $\beta_J \geq 5.7$. Up to now, simulations have generally been done either at $\beta_J = 5.7$ or 6. These choices are perhaps too optimistic.

It is important to check that the various lattice formulations leading to the same continuous model yield compatible results. Figure 10 displays the ratio m_N/m_ρ as a function of m_π/m_ρ , from a compilation of several simulations. This figure illustrates the state of the art up to 1987, displaying in particular the error bars. With such errors, all data are compatible. However, the simulations done at $\beta_J = 6$ and in the quenched approximation are systematically below the dashed line, while those performed at 5.7 are above this line. This is an evidence for an insufficient scaling for these gauge coupling values.

In conclusion, the small sample of the very numerous results obtained by many groups that we have reported show a remarkable qualitative agreement with the physical spectrum. It is to be expected that more elaborate simulations incorporating for instance dynamical quarks will improve the present evidence that quantum chromodynamics is indeed the best candidate for a theory of strong interactions.

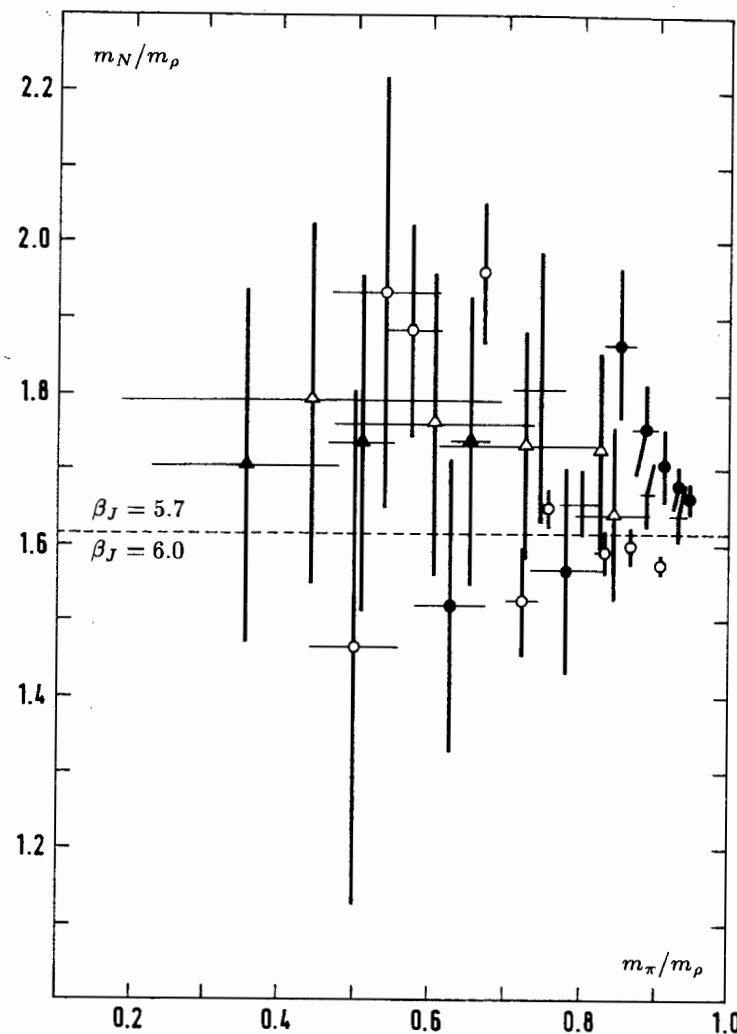


Fig. 10 The relation between the pion and nucleon masses, expressed in units of the ρ mass, according to several sources (from Bowler *et al*, *Phys. Lett.* **162B**, 354 (1985)).

Notes

The first proposition for a practical Monte Carlo simulation was done by N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953). A

general review on the Monte Carlo method is given by K. Binder, *Monte Carlo Investigations of Phase Transitions and Critical Phenomena* in the Domb and Green series, vol. 5B, Academic Press, New York (1976). Applications to lattice gauge theory are presented by M. Creutz, L. Jacobs and C. Rebbi, *Phys. Reports* **95** 201 (1983), and additional material can be found in *Advances in Lattice Gauge Theory*, eds D.W. Duke and F.J. Owens, World Scientific, Singapore (1985). See also in the series "Topics in current physics", *Monte Carlo Methods in Statistical Physics*, vol. 7 (1979) and *Applications of the Monte Carlo Method*, vol. 36 (1984), ed. K. Binder, Springer Verlag, Berlin. General computer algorithms are described in the book by D. Knuth, *The Art of Computer Programming*, Addison-Wesley, Reading (1973), in particular volume 2 discusses random generators.

For complementary material on the critical slowing down, see L. Van Hove, *Phys. Rev.* **93** 268 (1954) and the review by P.C. Hohenberg and B.I. Halperin, *Rev. Mod. Phys.* **49** 435 (1977).

The Monte Carlo Renormalization Group method is presented by K. Wilson in *Recent Developments in Gauge Theories*, Cargèse summer school 1979, G. 't Hooft *et al* eds., Plenum, New York (1980). See also the works by S.K. Ma, *Phys. Rev. Lett.* **37** 461 (1976), L.P. Kadanoff, *Rev. Mod. Phys.* **49** 267 (1977). The article by R.H. Swendsen in *Real Space Renormalization*, Topics in current physics, **30**, p. 57, T.W. Burkhardt and J.M.J. van Leeuwen eds, Springer, Berlin (1982) inspired the discussion of section 2.3.

The use of the Langevin equation in Monte Carlo simulations was suggested by G. Parisi and Wu Yong Shi, *Sci. Sin.* **24** 483 (1981), G. Parisi, *Nucl. Phys.* **B180** [FS2] 378 (1981) and **B205** [FS5] 337 (1982).

9

CONFORMAL INVARIANCE

It was noted by Polyakov and others in the early seventies that critical models implement a global conformal invariance which goes beyond pure scale invariance. The latter affects relative distances by a constant (i.e. space independent) factor, while other conformal transformations involve a space dependent factor. This invariance property enables one to fix not only the form of two-point but also three-point functions at criticality, when they are nonvanishing. However the conformal group is in general a finite dimensional Lie group, so that the resulting constraints are limited in number. In two dimensions, a new phenomenon arises, well known in the theory of analytic functions, namely there exists a plethora of local conformal transformations. As a result, it was tempting to investigate the possible consequences of local scale invariance in two dimensions. This is what was brilliantly undertaken by Belavin, Polyakov and Zamolodchikov in 1983, launching a new wave of applications in statistical physics. As the subject is still in its development, the present chapter will not be as elementary as previous ones, nor will it presumably remain up to date, especially as it is closely related to string field theory, a promising new approach to the quantum description of extended objects, which attempts to embrace all known interactions including gauge theories and gravity.

The ideas of local scale invariance, and possibly of a local renormalization group, are not yet fully appreciated. They prompt us to look at several aspects which were not in the forefront in the study of critical phenomena. One of them is the relevance of finite size effects at criticality where, due to the lack of an intrinsic length scale, there is a maximal sensitivity to geometrical effects like boundaries. A second novel aspect is the meaning of the values obtained for critical exponents, critical correlations, etc. It will appear that for many two-dimensional systems

general review on the Monte Carlo method is given by K. Binder, *Monte Carlo Investigations of Phase Transitions and Critical Phenomena* in the Domb and Green series, vol. 5B, Academic Press, New York (1976). Applications to lattice gauge theory are presented by M. Creutz, L. Jacobs and C. Rebbi, *Phys. Reports* **95** 201 (1983), and additional material can be found in *Advances in Lattice Gauge Theory*, eds D.W. Duke and F.J. Owens, World Scientific, Singapore (1985). See also in the series "Topics in current physics", *Monte Carlo Methods in Statistical Physics*, vol. 7 (1979) and *Applications of the Monte Carlo Method*, vol. 36 (1984), ed. K. Binder, Springer Verlag, Berlin. General computer algorithms are described in the book by D. Knuth, *The Art of Computer Programming*, Addison-Wesley, Reading (1973), in particular volume 2 discusses random generators.

For complementary material on the critical slowing down, see L. Van Hove, *Phys. Rev.* **93** 268 (1954) and the review by P.C. Hohenberg and B.I. Halperin, *Rev. Mod. Phys.* **49** 435 (1977).

The Monte Carlo Renormalization Group method is presented by K. Wilson in *Recent Developments in Gauge Theories*, Cargèse summer school 1979, G. 't Hooft *et al* eds., Plenum, New York (1980). See also the works by S.K. Ma, *Phys. Rev. Lett.* **37** 461 (1976), L.P. Kadanoff, *Rev. Mod. Phys.* **49** 267 (1977). The article by R.H. Swendsen in *Real Space Renormalization*, Topics in current physics, **30**, p. 57, T.W. Burkhardt and J.M.J. van Leeuwen eds, Springer, Berlin (1982) inspired the discussion of section 2.3.

The use of the Langevin equation in Monte Carlo simulations was suggested by G. Parisi and Wu Yong Shi, *Sci. Sin.* **24** 483 (1981), G. Parisi, *Nucl. Phys.* **B180** [FS2] 378 (1981) and **B205** [FS5] 337 (1982).

9

CONFORMAL INVARIANCE

It was noted by Polyakov and others in the early seventies that critical models implement a global conformal invariance which goes beyond pure scale invariance. The latter affects relative distances by a constant (i.e. space independent) factor, while other conformal transformations involve a space dependent factor. This invariance property enables one to fix not only the form of two-point but also three-point functions at criticality, when they are nonvanishing. However the conformal group is in general a finite dimensional Lie group, so that the resulting constraints are limited in number. In two dimensions, a new phenomenon arises, well known in the theory of analytic functions, namely there exists a plethora of local conformal transformations. As a result, it was tempting to investigate the possible consequences of local scale invariance in two dimensions. This is what was brilliantly undertaken by Belavin, Polyakov and Zamolodchikov in 1983, launching a new wave of applications in statistical physics. As the subject is still in its development, the present chapter will not be as elementary as previous ones, nor will it presumably remain up to date, especially as it is closely related to string field theory, a promising new approach to the quantum description of extended objects, which attempts to embrace all known interactions including gauge theories and gravity.

The ideas of local scale invariance, and possibly of a local renormalization group, are not yet fully appreciated. They prompt us to look at several aspects which were not in the forefront in the study of critical phenomena. One of them is the relevance of finite size effects at criticality where, due to the lack of an intrinsic length scale, there is a maximal sensitivity to geometrical effects like boundaries. A second novel aspect is the meaning of the values obtained for critical exponents, critical correlations, etc. It will appear that for many two-dimensional systems

they possess an unexpected group theoretical (and even number theoretical) interpretation, being related to the representation theory of some infinite Lie algebras. A third question, partly answered, but not treated here, is the relation with integrable models. Finally the deepest questions perhaps are to look for suitable generalisations both in the direction of studying the effects of moving off criticality, and of trying to gain a deeper understanding of critical phenomena in higher dimensions. The present success in the two-dimensional case can partly be traced to the use of analytic functions which enables us to reduce two-dimensional problems to one-dimensional ones.

9.1 Energy-momentum tensor – Virasoro algebra

9.1.1 Conformal invariance

Since finite size effects play a major role at criticality, we must first ascertain what we mean by a critical system enclosed in a finite domain. This can be phrased in various ways, either by dilating the box, or by saying that the system has no intrinsic length scale except for those dictated by the geometry, or finally by studying the short-distance behaviour of correlations and making sure that it agrees with the infinite volume one at criticality. This is one of the crucial aspects of what is to follow, namely the local characterization of critical systems, and it will be fully elucidated in two dimensions.

Let us first give an example to demonstrate the phenomenon. Consider the correlation function of Ising spins in the two-dimensional case at criticality. The dimension of the spin field is equal to $\frac{1}{8}$, as we know from chapter 2. In the infinite plane

$$\langle \sigma(\mathbf{r}_1) \sigma(\mathbf{r}_2) \rangle = 1/|\mathbf{r}_{12}|^{1/4} \quad (1)$$

It was also mentioned in chapter 2 that in a strip of width L ($0 \leq x < L$), with periodic boundary conditions, the corresponding correlation reads

$$\begin{aligned} & \langle \sigma(\mathbf{r}_1) \sigma(\mathbf{r}_2) \rangle \\ &= \left(\frac{\pi}{L} \right)^{\frac{1}{4}} \left[\sinh^2 \frac{\pi}{L} y_{12} \cos^2 \frac{\pi}{L} x_{12} + \cosh^2 \frac{\pi}{L} y_{12} \sin^2 \frac{\pi}{L} x_{12} \right]^{-\frac{1}{8}} \end{aligned} \quad (2)$$

When $L \rightarrow \infty$, or $\mathbf{r}_{12} \rightarrow 0$, this expression reduces to (1). Rather than changing the geometry (varying L) it is better to characterize intrinsically the critical model by its short-distance scale invariant behaviour.

A conformal transformation of an Euclidean d -dimensional space (or more generally of a Riemannian manifold) is a point to point transformation which preserves angles but not necessarily distances. These transformations form a group which includes the Euclidean group as a subgroup (translations, d parameters, rotations $\frac{1}{2}d(d-1)$ parameters), dilatations (one parameter), and special conformal transformations (d parameters) obtained by composing an inversion, a translation and a second inversion. This describes the connected part of the noncompact conformal group, $SO(d+1, 1)$ with $\frac{1}{2}(d+1)(d+2)$ parameters. To be precise, the Euclidean space is completed by a point at infinity (giving it the topology of a compact sphere) to make the above transformations well-defined everywhere.

To understand globally the conformal group, it is convenient to map the completed Euclidean space $\mathcal{R}^d$, with metric $\mathbf{r}^2 = r_1^2 + \dots + r_d^2$, onto the unit sphere S_d of the Euclidean space $\mathcal{R}^{d+1}$, $x_-^2 + \mathbf{x}^2 = 1$ using a stereographic (conformal) projection as shown on figure 1. Correspondingly

$$\begin{aligned} \mathbf{x} &= \frac{2\mathbf{r}}{1+r^2} & \mathbf{r} &= \frac{\mathbf{x}}{1-x_-} \\ x_- &= \frac{r^2-1}{r^2+1} & r^2 &= \frac{1+x_-}{1-x_-} \end{aligned} \quad (3)$$

Using homogeneous coordinates, S_d may be identified with the section $x_+ = 1$ of the light cone $\mathcal{C}$ of Minkowski space $\mathcal{R}^{d+1,1}$ with metric $x_+^2 - x_-^2 - \mathbf{x}^2$. Rays on $\mathcal{C}$ are in one-to-one correspondence with points of $\mathcal{R}^d \cup \{\infty\}$. Conformal transformations are interpreted as homogeneous Lorentz transformations in $\mathcal{R}^{d+1,1}$. The little group of a lightlike vector can be identified with the Euclidean group in d -dimensions (the point singled out may be thought of as the point at infinity). For instance if the lightlike vector has components $(x_+, x_-, \mathbf{x}) = (1, 1, 0)$ a translation $\Lambda_{\mathbf{a}}$ is represented as

$$\begin{aligned} \Lambda_{\mathbf{a}}(1, 1, 0) &= (1, 1, 0) \\ \Lambda_{\mathbf{a}}(1, -1, 0) &= (1, -1, 0) + \mathbf{a}^2(1, 1, 0) + (0, 0, \mathbf{a}) \\ \Lambda_{\mathbf{a}}(0, 0, \mathbf{v}) &= \mathbf{a} \cdot \mathbf{v}(1, 1, 0) + (0, 0, \mathbf{v}) \end{aligned} \quad (4)$$

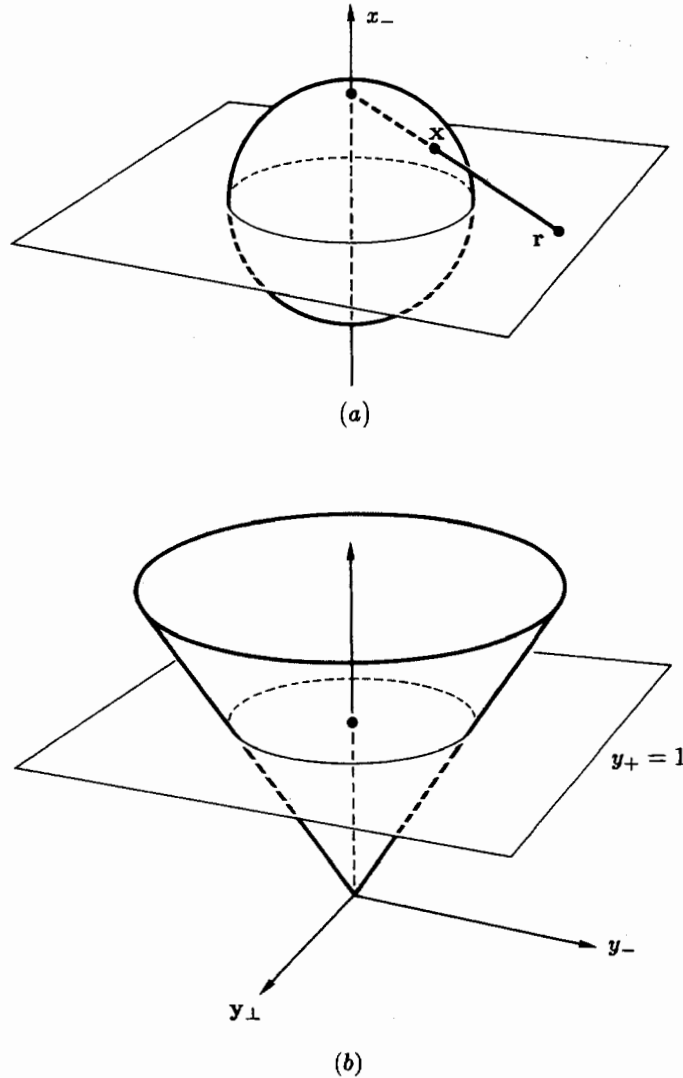


Fig. 1 (a) Stereographic projection of S^d onto $\mathcal{R}^d$. (b) The isotropic cone (light cone) in Minkowski space $\mathcal{R}^{d+1,1}$.

An inversion belonging to a disconnected component of $O(d+1, 1)$ corresponds to $x_- \rightarrow -x_-$, and a dilatation through λ is represented by an hyperbolic rotation in the x_+, x_- plane through

$$\begin{aligned}\Lambda_\lambda(1, 1, 0) &= \lambda(1, 1, 0) \\ \Lambda_\lambda(1, -1, 0) &= \lambda^{-1}(1, -1, 0) \\ \Lambda_\lambda(0, 0, \mathbf{v}) &= (0, 0, \mathbf{v})\end{aligned}\quad (5)$$

Returning to d -dimensional space a special conformal transformation takes the form

$$\mathbf{r} \rightarrow \mathbf{r}' = \frac{\mathbf{r} + \mathbf{r}^2 \mathbf{b}}{1 + 2\mathbf{b} \cdot \mathbf{r} + \mathbf{b}^2 \mathbf{r}^2} \quad (6)$$

In infinitesimal form

$$\mathbf{r} \rightarrow \mathbf{r}' = \mathbf{r} + \delta \mathbf{r} \quad \delta \mathbf{r} = \mathbf{r}^2 \left(\delta \mathbf{b} - \frac{2\mathbf{r}(\mathbf{r} \cdot \delta \mathbf{b})}{\mathbf{r}^2} \right) \quad (7)$$

Euclidean displacements conserve the metric and have therefore a unit Jacobian. This is not the case for dilatations, nor for special transformations, in which case the Jacobian is

$$\frac{D(\mathbf{r}')}{D(\mathbf{r})} = \frac{1}{(1 + 2\mathbf{b} \cdot \mathbf{r} + \mathbf{b}^2 \mathbf{r}^2)^d} \quad (8)$$

Assume that critical models are not only scale invariant, but also conformal invariant. This hypothesis will receive support as we proceed. Consider for simplicity scalar observables, with corresponding dimension denoted Δ . These assumptions mean that under conformal transformations, the correlation functions satisfy

$$\langle A_1(\mathbf{r}_1) \dots A_n(\mathbf{r}_n) \rangle = \prod_{i=1}^n \left| \frac{D(\mathbf{r}'_i)}{D(\mathbf{r}_i)} \right|^{\Delta_i/d} \langle A_1(\mathbf{r}'_1) \dots A_n(\mathbf{r}'_n) \rangle \quad (9)$$

For a two-point function, using translational, rotational and dilatation invariance, we get, with a normalizing constant C_{12}

$$\langle A_1(\mathbf{r}_1) A_2(\mathbf{r}_2) \rangle = \frac{C_{12}}{r_{12}^{\Delta_1 + \Delta_2}} \quad (10)$$

Under a special transformation, we have

$$\mathbf{r}'_{12} = \frac{\mathbf{r}_{12}^2}{(1 + 2\mathbf{b} \cdot \mathbf{r}_1 + \mathbf{b}^2 \mathbf{r}_1^2)(1 + 2\mathbf{b} \cdot \mathbf{r}_2 + \mathbf{b}^2 \mathbf{r}_2^2)} \quad (11)$$

Consequently equations (9) and (10) are only compatible if $\Delta_1 = \Delta_2$

$$\begin{aligned} \Delta_1 \neq \Delta_2 \quad \langle A_1(\mathbf{r}_1) A_2(\mathbf{r}_2) \rangle &= 0 \\ \Delta_1 = \Delta_2 = \Delta \quad \langle A_1(\mathbf{r}_1) A_2(\mathbf{r}_2) \rangle &= \frac{C_{12}}{r_{12}^{2\Delta}} \end{aligned} \quad (12)$$

We can define an "orthonormal basis" of observables by choosing the appropriate combinations which diagonalize the symmetric real matrix of coefficients C_{ab} . A single rescaling of the fields will then replace C_{ab} by δ_{ab} . Similar reasoning enables one to obtain the three-point functions

$$\langle A_1(\mathbf{r}_1) A_2(\mathbf{r}_2) A_3(\mathbf{r}_3) \rangle = \frac{C_{123}}{r_{12}^{\Delta_1+\Delta_2-\Delta_3} r_{23}^{\Delta_2+\Delta_3-\Delta_1} r_{31}^{\Delta_3+\Delta_1-\Delta_2}} \quad (13)$$

Check that equation (13) is the unique solution to condition (9).

Beyond three-point functions, conformal invariance alone is not sufficient to determine the structure of correlations. This follows from the fact that some combinations such as $r_{13}r_{24}/r_{23}r_{14}$ (where r_{ij} stands for $|\mathbf{r}_i - \mathbf{r}_j|$) are invariant under conformal transformations. More information is required, which will be provided in the two-dimensional case from an assumption of invariance under local conformal transformations. For the time being, let us translate condition (9) in infinitesimal form by writing

$$\begin{aligned} \mathbf{r}' &= \mathbf{r} + \delta\mathbf{r} \\ A'(\mathbf{r}) &= \left(\frac{D(\mathbf{r}')}{D(\mathbf{r})} \right)^{\Delta/d} A(\mathbf{r}') = A(\mathbf{r}) + \delta A(\mathbf{r}) \\ \delta A(\mathbf{r}) &= \left[\delta\mathbf{r} \cdot \nabla + \frac{\Delta}{d} (\nabla \cdot \delta\mathbf{r}) \right] A(\mathbf{r}) \end{aligned} \quad (14)$$

The requirement that $\mathbf{r} \rightarrow \mathbf{r} + \delta\mathbf{r}$ be a conformal map in infinitesimal form means that $\delta\mathbf{r}$ should be at most quadratic in $\mathbf{r}$, i.e. of the form

$$\begin{aligned} \delta r_i &= \delta a_i + [\delta\lambda\delta_{ij} + \delta\varepsilon_{ij} + \delta b_i r_j - 2r_i\delta b_j] r_j \\ \nabla \cdot \delta\mathbf{r} &= d(\delta\lambda - 2\mathbf{r} \cdot \delta\mathbf{b}) \end{aligned} \quad (15)$$

where δa_i , $\delta\lambda$, $\delta\varepsilon_{ij} = -\delta\varepsilon_{ji}$ and δb_i are respectively the infinitesimal parameters for translations, dilatation, rotations and special

transformations. Explicitly

$$\begin{aligned} \delta A(\mathbf{r}) &= \left[\delta\mathbf{a} \cdot \nabla + \delta\lambda\mathbf{r} \cdot \nabla + \frac{1}{2}\delta\varepsilon_{ij}(r_i\nabla_j - r_j\nabla_i) \right. \\ &\quad \left. + r^2\delta\mathbf{b} \cdot \nabla - 2(\mathbf{r} \cdot \delta\mathbf{b})(\mathbf{r} \cdot \nabla) + \Delta_A(\delta\lambda - 2\mathbf{r} \cdot \delta\mathbf{b}) \right] A(\mathbf{r}) \end{aligned} \quad (16)$$

Therefore, if $\delta A_p(\mathbf{r}_p)$ stands for the corresponding expression pertaining to the (scalar) field A_p , equation (9) states that

$$\sum_{p=1}^n \langle A_1(\mathbf{r}_1) \dots \delta A_p(\mathbf{r}_p) \dots A_n(\mathbf{r}_n) \rangle = 0 \quad (17)$$

Generalize the above expressions in the case of tensorial fields.

9.1.2 Energy-momentum tensor

Critical theories in infinite Euclidean space are obtained from a limiting procedure starting from a sum (or integral) over Boltzmann weights. Local observables generally are expressed in terms of fundamental fields occurring in these weights. Let us denote the latter collectively by A . Correlation functions are averages over the Boltzmann weights with a partition function normalization. If we forget about the limiting procedures, we may think of performing in the numerator a dummy change of variable of the type $A(\mathbf{r}) \rightarrow A(\mathbf{r}) + \delta A(\mathbf{r})$.

$$\delta A(\mathbf{r}) = \left[\delta\mathbf{r} \cdot \nabla + \frac{\Delta}{d} \nabla \cdot \delta\mathbf{r} \right] A(\mathbf{r}) \quad (18)$$

where now $\delta\mathbf{r}(\mathbf{r})$ is an arbitrary infinitesimal vector field, i.e. not necessarily of the form (15). We further assume that the observables A_p transform also as (18), in which case they are likely to belong to a restricted class (with corresponding dimensions Δ_i). At any rate we denote by δA_p the corresponding variation of the observables A_p . Then the total variation of the correlation functions will not only contain a contribution analogous to (17), but will also include a term arising from the noninvariance of the measure (including the Boltzmann weight) under the change of variable (18). This term will be of first order in $\delta\mathbf{r}$ and should obviously vanish when $\delta\mathbf{r}$ is a constant (corresponding to a translation). It therefore only depends on derivatives of $\delta\mathbf{r}$. Since we are working with a local theory we write the simplest possible

form, as a postulate defining the (local) energy-momentum tensor density

$$\sum_{p=1}^n \langle A_1(\mathbf{r}_1) \dots \delta A_p(\mathbf{r}_p) \dots A_n(\mathbf{r}_n) \rangle + \int d^d \mathbf{r} \langle A_1(\mathbf{r}_1) \dots A_n(\mathbf{r}_n) T_{ij}(\mathbf{r}) \rangle \partial_i \delta r_j(\mathbf{r}) = 0 \quad (19)$$

This is analogous to the corresponding definition in Minkowskian local field theory, where this construction provides a Noether current for the conserved quantities, energy and momentum. This explains the origin of the name. For a given Lagrangian, one obtains a natural "classical" candidate for T_{ij} , but there might be difficulties in the way, which have their origin in the behaviour of the remaining integration measure. Therefore it is at this point more honest to assume (19) as a postulate, referred to as the basic Ward identity. In a more general setting, the energy-momentum tensor appears as the response to an infinitesimal change of metric, which insures that it is symmetric. We refrain from proceeding along these lines, although most two-dimensional applications suggest that there is much to learn from quantum field theory in curved space.

Returning to equation (19), let us look at the constraints implied by assuming invariance under Euclidean transformations and global dilatations. We substitute the corresponding expressions for δA_p and $\delta \mathbf{r}$. Then we have

- i) Invariance under translations, $\delta \mathbf{r}$ constant, equation (19) reduces to (17).
- ii) Invariance under rotations, $\partial_i \delta r_j = \delta \varepsilon_{ji} = -\delta \varepsilon_{ij}$ constant. For equation (19) to reduce to (17), it is required that T_{ij} be symmetric

$$T_{ij}(\mathbf{r}) = T_{ji}(\mathbf{r}) \quad (20)$$

- iii) Invariance under dilatations, $\delta_i \delta r_j = \delta_{ij} \delta \lambda$. To recover (17) the energy-momentum must be traceless

$$\sum_i T_{ii}(\mathbf{r}) = 0 \quad (21)$$

If one now considers an infinitesimal special transformation, with $\partial_i \delta r_j = 2(r_i \delta b_j - r_j \delta b_i) - \delta_{ij}(\delta \mathbf{b} \cdot \mathbf{r})$, the second term in (19) vanishes identically if we take (20) and (21) into account. In other words in a local theory, where the Ward identity (19) holds, Euclidean invariance and scale invariance alone insure at the formal level invariance under the full conformal group. This is the important original observation.

If an infinitesimal coordinate transformation is interpreted in an active sense, then the infinitesimal square length between neighbouring points is changed according to $(g_{ij}^{(0)} = \delta_{ij})$

$$ds^2 = d\mathbf{r}^2 \rightarrow d(\mathbf{r} + \delta \mathbf{r})^2 = (g_{ij}^{(0)} + \delta g_{ij}) d\mathbf{r}_i d\mathbf{r}_j \quad (22)$$

$$\delta g_{ij} = \partial_i \delta r_j + \partial_j \delta r_i$$

If we substitute $\frac{1}{2} T_{ij} \delta g_{ij}$ for $T_{ij} \partial_i \delta r_j$ we find that $\int d^d \mathbf{r} T_{ij}(\mathbf{r}) \delta g_{ij}(\mathbf{r})$ appears as the response of the functional measure (including the Boltzmann weight) to a change in the metric. Since here everything is understood as relative to the flat metric $g_{ij}^{(0)}$, we did not bother to distinguish between covariant and contravariant indices.

Let us assume that for a general coordinate variation the variations δA_p are local, of the form (18). Furthermore consider the case where $\delta \mathbf{r}$ is arbitrary and of compact support. Then after a partial integration, we can rewrite equation (19), by identifying the coefficient of $\delta \mathbf{r}$, as

$$\begin{aligned} \partial_i \langle T_{ij}(\mathbf{r}) A_1(\mathbf{r}_1) \dots A_n(\mathbf{r}_n) \rangle &= \\ &= \sum_{p=1}^n \left\{ \delta(\mathbf{r} - \mathbf{r}_p) \partial_j^{(p)} - \frac{\Delta_p}{d} (\partial_j \delta(\mathbf{r} - \mathbf{r}_p)) \right\} \langle A_1(\mathbf{r}_1) \dots A_n(\mathbf{r}_n) \rangle \end{aligned} \quad (23)$$

This is what is meant by saying that the energy-momentum is conserved. Indeed the right-hand side vanishes for $\mathbf{r}$ in any domain omitting a neighbourhood of the points $\mathbf{r}_1, \dots, \mathbf{r}_n$, so does therefore the left-hand side with the insertion of the divergence of T .

It is not possible for the time being to go much further in the general case. We turn therefore to the special features occurring in two dimensions.

9.1.3 Two-dimensional conformal transformations

In two dimensions, the global conformal group is the familiar Lorentz group $SO(3,1)$. It may be realized by acting on the complex plane (completed by a point at infinity) through linear rational transformations (or Möbius transformations) exhibiting the isomorphism of $SO(3,1)$ with $SL(2, \mathbb{C})/\{\pm 1\}$. Instead of using two real independent coordinates x and y , it is convenient to introduce the two independent complex combinations $z = x + iy$, $\bar{z} = x - iy$. Using this notation, the flat metric becomes antidiagonal, and care must be exercised with tensors. We write

$$\begin{aligned} ds^2 &= g_{zz} dz^2 + (g_{z\bar{z}} + g_{\bar{z}z}) dz d\bar{z} + g_{\bar{z}\bar{z}} d\bar{z}^2 = dz d\bar{z} \\ dx \wedge dy &= d\bar{z} \wedge dz / 2i \quad \partial_x^2 + \partial_y^2 = 4\partial_z \partial_{\bar{z}} \\ \partial_z &= \frac{1}{2}(\partial_x - i\partial_y) \quad \partial_{\bar{z}} = \frac{1}{2}(\partial_x + i\partial_y) \end{aligned} \quad (24)$$

For short we sometimes abbreviate ∂_z as ∂ and $\partial_{\bar{z}}$ as $\bar{\partial}$. It follows that

$$\begin{aligned} g_{zz} = g_{\bar{z}\bar{z}} &= 0 & g_{z\bar{z}} = g_{\bar{z}z} &= \frac{1}{2} \\ g^{zz} = g^{\bar{z}\bar{z}} &= 0 & g^{z\bar{z}} = g^{\bar{z}z} &= 2 \end{aligned} \quad (25)$$

A symmetric traceless tensor such as T is obtained in the $z, \bar{z}$ basis by forming the scalar

$$T_{ij} \xi^i \xi^j = \frac{1}{2} (T_{11} - iT_{12}) \xi^2 + \frac{1}{2} (T_{11} + iT_{12}) \bar{\xi}^2$$

where $\xi = \xi^1 + i\xi^2$, $\bar{\xi} = \xi^1 - i\xi^2$. Thus

$$T_{zz} = \frac{1}{2} (T_{11} - iT_{12}) \quad T_{\bar{z}\bar{z}} = \frac{1}{2} (T_{11} + iT_{12}) \quad T_{z\bar{z}} = T_{\bar{z}z} = 0 \quad (26)$$

With a, b, c, d complex satisfying $ad - bc = 1$, the global conformal transformations read

$$z, \bar{z} \rightarrow z', \bar{z}' \quad z' = f(z) = \frac{az + b}{cz + d} \quad \bar{z}' = \overline{f(z)} = \frac{\bar{a}\bar{z} + \bar{b}}{\bar{c}\bar{z} + \bar{d}} \quad (27)$$

$$\frac{D(z', \bar{z}')}{D(z, \bar{z})} = f'(z) \overline{f'(z)} = \frac{1}{(cz + d)^2 (\bar{c}\bar{z} + \bar{d})^2} \quad (28)$$

We can also envision other global conformal transformations which change the orientation and correspond to a disconnected part of the group. They are generated by complex conjugation and will not play a crucial role, since we will mostly be concerned with infinitesimal transformations.

Tensor fields transform with integral powers of f' and $\bar{f}'$. But here we encounter a specificity of two dimensions, in that it makes perfect sense to think of a generalization of the notion of tensor with a fractional number of z or $\bar{z}$ indices. More precisely we define a field with conformal weights $(h, \bar{h})$ as one which transforms according to

$$A(z, \bar{z}) \rightarrow A'(z, \bar{z}) = f'(z)^h \overline{f'(z)}^{\bar{h}} A(f(z), \overline{f(z)}) \quad (29)$$

where $\bar{h}$ is not meant as the complex conjugate of h , but the notation $\bar{h}$ is suggestive of the fact that it relates to the behaviour associated to the $\bar{z}$ variable. The quantity $h - \bar{h}$ plays the role of angular momentum or spin, indicating the behaviour under rotations $z \rightarrow e^{i\varphi} z$ with A' picking a factor $e^{i(h-\bar{h})\varphi}$. For most (but not all) of what follows, we shall consider integral spin fields ($h - \bar{h}$ an integer). Indeed these are the only observable ones. However it is natural to envision also cases involving Fermi fields with half integral spins, and further generalizations have been investigated. The scaling dimension Δ is the sum

$$\Delta = h + \bar{h} \quad (30)$$

and for a real scalar field we have

$$h = \bar{h} = \frac{1}{2}\Delta \quad (31)$$

We can interpret equation (29) by saying that, in infinitesimal form

$$(A + \delta A) dz^h d\bar{z}^{\bar{h}} = A(z + \delta z, \bar{z} + \delta \bar{z}) [d(z + \delta z)]^h [d(\bar{z} + \delta \bar{z})]^{\bar{h}} \quad (32)$$

Therefore, with

$$\delta z = \varepsilon(z) \quad \delta A = \delta_z A + \delta_{\bar{z}} A \quad (33)$$

we have

$$\begin{aligned} \delta_z A(z, \bar{z}) &= \{\varepsilon(z) \partial_z + h \varepsilon'(z)\} A(z, \bar{z}) \\ \delta_{\bar{z}} A(z, \bar{z}) &= \{\overline{\varepsilon(z)} \partial_{\bar{z}} + \bar{h} \overline{\varepsilon'(z)}\} A(z, \bar{z}) \end{aligned} \quad (34)$$

The novel feature arising in two dimensions is the existence of an infinite class of conformal transformations generated by an arbitrary analytic function $f(z)$ instead of a rational linear function as was used until now (we forget here the possibility that f might be an analytic function of $\bar{z}$). That these transformations

are conformal follows immediately from the fact that the metric changes into

$$dzd\bar{z} \rightarrow dfd\bar{f} = f'(z)\overline{f'(z)}dzd\bar{z} \quad (35)$$

so that angles are preserved, but lengths are locally multiplied by a dilatation factor $|f'(z)|$. The interpretation of such transformations is not straightforward. The point is that in general

- i) $f(z)$ might not be defined throughout the complex plane
- ii) and even when this is the case (admitting isolated poles) it will not be a one-to-one map.

The only one-to-one analytic maps of the (completed) complex plane are the linear fractional transformations, i.e. those corresponding to the global group. There are two possible interpretations, both of them interesting. In the first one, we use the assumption of local invariance to obtain quantities pertaining to the transformed domain. Shortly we shall look at such a map from the punctured plane onto a periodic strip. In the second interpretation, we study the effect of infinitesimal local conformal transformations such as (33) with $\varepsilon(z)$ considered of first order but otherwise an arbitrary function of z . It is good to keep in mind that analytic functions are rigid, i.e. $\varepsilon(z)$ cannot be made to be smooth and vanishing outside a bounded domain while keeping it analytic.

Under an arbitrary "infinitesimal" holomorphic change of coordinates

$$\begin{aligned} \delta z &= \delta \varepsilon g(z) \\ \delta \bar{z} &= \delta \varepsilon \bar{g}(\bar{z}) \end{aligned} \quad (36)$$

fields which transform as

$$\delta A(z, \bar{z}) = \delta \varepsilon [g\partial_z + h(\partial_z g) + \bar{g}\partial_{\bar{z}} + \bar{h}(\partial_{\bar{z}} \bar{g})] A(z, \bar{z}) \quad (37)$$

are called primary ones. Here the emphasis is on the fact that g is arbitrary, and not restricted to be a second degree polynomial in z (in which case A could be called quasiprimary). We shall deal almost exclusively with primary fields, the most notable exception being, surprisingly, the energy-momentum tensor. For a general $g(z, \bar{z})$, we have

$$T_{ij}\partial_i\delta\mathbf{r}_j = 2\delta\varepsilon [T_{zz}\partial_z g + T_{\bar{z}\bar{z}}\partial_{\bar{z}} \bar{g}] \quad (38)$$

This quantity vanishes for g ($\bar{g}$) holomorphic (antiholomorphic). Let us insert this expression into equation (19). Of course the second term vanishes in those (finite) regions of the $z, \bar{z}$ plane where g and $\bar{g}$ have the above analyticity property, suggesting a covariance property of correlations of primary fields under general analytic transformations. Let us make this statement more precise. It is convenient to rescale the energy-momentum tensor in order to insert a factor $1/2\pi$ into the integral defining T . Choosing a pair of functions $g, \bar{g}$, vanishing sufficiently fast at infinity to allow integrations by parts, holomorphic for g and antiholomorphic for $\bar{g}$ in the vicinity of the points $z_1, \dots, z_n$, so that we can use equation (37), and setting $g_p \equiv g(z_p, \bar{z}_p)$, $\partial_p \equiv \partial_{z_p}$, we obtain

$$\begin{aligned} \sum_{p=1}^n [g_p \partial_p + h_p(\partial_p g_p) + \bar{g}_p \bar{\partial}_p + \bar{h}_p(\bar{\partial}_p \bar{g}_p)] \langle A_1 \dots A_n \rangle \\ = \int \frac{d\bar{z} \wedge dz}{2i\pi} [g(z, \bar{z}) \partial_{\bar{z}} \langle T_{zz}(z, \bar{z}) A_1 \dots A_n \rangle \\ + \bar{g}(z, \bar{z}) \partial_z \langle T_{\bar{z}\bar{z}}(z, \bar{z}) A_1 \dots A_n \rangle] \end{aligned} \quad (39)$$

Comparing both sides, it follows that $\partial_{\bar{z}} \langle T_{zz}(z, \bar{z}) A_1 \dots A_n \rangle$ vanishes except at the points $z_1 \dots z_n$. A similar conclusion holds for $\partial_z \langle T_{\bar{z}\bar{z}}(z, \bar{z}) A_1 \dots A_n \rangle$. Indeed this equation shows that the first of these functions is a meromorphic one in z with single and double poles located at $z_1 \dots z_n$, and the corresponding conjugate property for the second, so that we are entitled to write the Ward identity in the useful form

$$\begin{aligned} \langle T_{zz}(z) A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle &= \\ \sum_{p=1}^n \left[\frac{h_p}{(z - z_p)^2} + \frac{1}{(z - z_p)} \partial_{z_p} \right] \langle A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle \\ \langle T_{\bar{z}\bar{z}}(\bar{z}) A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle &= \\ \sum_{p=1}^n \left[\frac{\bar{h}_p}{(\bar{z} - \bar{z}_p)^2} + \frac{1}{(\bar{z} - \bar{z}_p)} \partial_{\bar{z}_p} \right] \langle A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle \end{aligned} \quad (40)$$

Given the parallel treatment of the z and $\bar{z}$ variables, we shall in the following not duplicate the equations, giving only the z part. However, the reader should be aware that some fields, such as the A 's above, might have both a z and $\bar{z}$ dependence, while it follows from equation (40) that T_{zz} depends only on z at least in the sense

that all correlation with T_{zz} insertions are meromorphic functions of z . Henceforth we shall denote it $T(z)$, similarly for $T_{\bar{z}\bar{z}}$ denoted $\bar{T}(\bar{z})$.

If now $g(z)$ is analytic in a domain enclosing the points $z_1, \dots, z_n$ and if $\mathcal{C}$ is a curve in this domain encircling these points once in the positive sense, the integral form of (40) reads

$$\sum_{p=1}^n [g(z_p) \partial_{z_p} + h_p g'(z_p)] \langle A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle = \int_{\mathcal{C}} \frac{dz}{2i\pi} g(z) \langle T(z) A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle \quad (41)$$

and a similar conjugate equation is obtained with $\bar{T}$. The generators of global conformal transformations correspond to functions $g(z)$ which are second degree polynomials in z , in which case the left-hand side of the above equation must vanish. As a result, a correlation function with a $T(z)$ insertion vanishes like z^{-4} for large z

$$\langle T(z) A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle \underset{z \rightarrow \infty}{\sim} z^{-4} \Gamma(z_1, \bar{z}_1; \dots; z_n, \bar{z}_n) \quad (42)$$

with a conjugate relation for $\bar{T}$. The canonical dimension of T is 2 (and d in d dimensions) since its integral must be dimensionless (as is the action). A naive idea would have been that the left-hand side of (42) vanishes like z^{-2} at infinity. The present statement is stronger, implying global conformal invariance. A second degree polynomial depends on three complex parameters, which is of course the (complex) dimension of the group $SL(2, \mathbb{C})$. Combining (42) with (40), we find by an expansion around $z = \infty$, that global conformal transformations imply

$$\begin{aligned} \sum_{p=1}^n \partial_{z_p} \langle A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle &= 0 \\ &\text{translations} \\ \sum_{p=1}^n (z_p \partial_{z_p} + h_p) \langle A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle &= 0 \\ &\text{(complex) dilatations} \\ \sum_{p=1}^n (z_p^2 \partial_{z_p} + 2h_p z_p) \langle A_1(z_1, \bar{z}_1) \dots A_n(z_n, \bar{z}_n) \rangle &= 0 \end{aligned}$$

In the same way that we do not write explicitly the conjugate equations, we shall sometimes omit to indicate the $\bar{z}$ variables in the z -equations. In two dimensions, rotations and real dilatations are conjugate transformations. In the complex formalism they combine into complex dilatations as indicated in equation (43).

Relations (41) express a generalized form of covariance under infinitesimal local conformal transformations. To proceed further, we need some properties of correlations involving the energy-momentum tensor, and the first task is to study its behaviour under the above mentioned transformations.

9.1.4 Central charge

From translational invariance in the plane, the average value $\langle T(z) \rangle$ is a constant. It is implicit in equation (42) that this constant has been chosen equal to zero

$$\langle T(z) \rangle = \langle \bar{T}(\bar{z}) \rangle = 0 \quad (44)$$

Similarly translational and global scale invariance imply that the corresponding two-point correlation be a homogeneous function of z_{12} . From equation (42) we should therefore have

$$\begin{aligned} \langle T(z_1) T(z_2) \rangle &= \frac{\frac{1}{2}c}{(z_{12})^4} & \langle \bar{T}(\bar{z}_1) \bar{T}(\bar{z}_2) \rangle &= \frac{\frac{1}{2}c}{(\bar{z}_{12})^4} \\ \langle T(z_1) \bar{T}(\bar{z}_2) \rangle &= 0 \end{aligned} \quad (45)$$

from which we learn that T has conformal weights $(2,0)$ ($\bar{T}$ has weights $(0,2)$) as expected. The choice of the factor $\frac{1}{2}$ in the normalization of the real constant c common to T and $\bar{T}$ is conventional. It is designed to produce the value $c = 1$ for the scalar massless field (see below). Since the scale of T is already fixed by relations such as equation (40), we cannot dispose of it. If the theory admits an underlying Hamiltonian quantum mechanical interpretation, with a positive definite Hilbert space of states and energies bounded from below, then c has to be positive, as we shall see later. But this is by no means necessary for a consistent statistical interpretation, and we shall find interesting cases with $c < 0$.

The dimensionless constant c is called the central charge for reasons which will appear below. One of the main conclusions of this chapter will be that knowledge of c is (almost) sufficient

any rate, it gives essential information from which much can be derived.

The nonvanishing of c corresponds to an interesting anomaly. Such anomalies occur when a classical symmetry cannot be implemented quantum mechanically as a consequence of renormalization effects. In the context of path integrals, this can also be interpreted by saying that the complete functional measure cannot be made invariant. Such instances arise frequently when dealing with local symmetries. One can also say that the renormalization group equations express an anomalous behaviour under global dilatations off the fixed points.

We just stated that $\langle T(z) \rangle$ vanishes in the infinite plane. If we had invariance under infinitesimal local conformal transformations, we would expect that this mean value would remain equal to zero. But except for those transformations generated by a second degree polynomial in z , we know that one cannot, in a strict sense, speak of an invariance, since the transformations are not one-to-one mappings of the (completed) complex plane onto itself. As a result, we do not expect $\langle \delta T(z) \rangle$ to vanish. Indeed combining equations (41) and (45), we have

$$\begin{aligned} z &\rightarrow z' = z + \delta\epsilon g(z) \\ T &\rightarrow T' = T + \delta T \end{aligned} \quad (46a)$$

$$\langle \delta T(z) \rangle = \delta\epsilon \int_c \frac{dz'}{2i\pi} g(z') \langle T(z') T(z) \rangle = \frac{1}{2} c \delta\epsilon \int_c \frac{dz'}{2i\pi} \frac{g(z')}{(z' - z)^4}$$

and as a result

$$\langle \delta T(z) \rangle = \frac{1}{12} c g'''(z) \delta\epsilon \quad (46b)$$

with a similar expression for $\bar{T}$. When g is a second degree polynomial, g''' vanishes, and we have the expected invariance. Equation (46) is a first expression of the anomaly. Given that $\langle T(z) \rangle = 0$, and that the weights of T are (2,0), it is natural to postulate that the minimal transformation law incorporating the above anomaly is

$$\delta T(z) = \delta\epsilon \left[(2g'(z) + g(z)\partial_z) T(z) + \frac{1}{12} c g'''(z) \right] \quad (47)$$

with a similar formula for $\bar{T}$. This behaviour should be contrasted with the one of primary fields, given by equation (33). Equation (47) teaches us that T is not a primary field. It exhibits an inhomogeneous behaviour, with an extra contribution proportional

to the identity. This suggests that the anomaly may be derived from the corresponding behaviour of the free energy, or partition function.

Let us transcribe here the expressions (12) and (13) for arbitrary two- and three-point functions, omitting the $\bar{z}$ variables,

$$\begin{aligned} \langle A_h(z_1) A_h(z_2) \rangle &= \frac{\alpha_{12}}{(z_{12})^{2h}} \\ \langle A_{h_1}(z_1) A_{h_2}(z_2) A_{h_3}(z_3) \rangle &= \frac{\alpha_{123}}{z_{12}^{h_1+h_2-h_3} z_{23}^{h_2+h_3-h_1} z_{31}^{h_3+h_1-h_2}} \end{aligned} \quad (48)$$

Consequently

$$\langle T(z_1) T(z_2) T(z_3) \rangle = \frac{c}{(z_{12} z_{23} z_{31})^2} \quad (49)$$

Check that this is in agreement with

$$\langle \delta T(z_1) T(z_2) \rangle + \langle T(z_1) \delta T(z_2) \rangle = \delta\epsilon \int_c \frac{dz_3}{2i\pi} g(z_3) \langle T(z_1) T(z_2) T(z_3) \rangle$$

It is interesting to know the generalization of equation (47) to a finite conformal transformation mapping the plane onto a possibly different domain. We now have $z' = f(z)$. To $T(z)$ will correspond $T'(z')$. With the active interpretation, the corresponding formula reads

$$T(z) dz^2 = T'(z') dz'^2 + \frac{1}{12} c \{z', z\} dz^2 \quad (50)$$

where the symbol $\{z', z\}$ stands for the Schwarzian derivative of the map $z \rightarrow z' = f(z)$, namely

$$\{z', z\} = \frac{f'''}{f'} - \frac{3}{2} \left(\frac{f''}{f'} \right)^2 \quad (51)$$

To first order in $\delta\epsilon$, this reduces to the third derivative of $\delta\epsilon g(z)$ for $f(z) = z + \delta\epsilon g(z)$.

It is not obvious how, from the composition of infinitesimal maps, one recovers the Schwarzian derivative. We shall present a proof using free field theory in section 2.1. Let us digress a little to give a simple interpretation of the Schwarzian derivative. It is related to the global conformal group very much as the ordinary derivative is related to translations. As will shortly be clear, if we use the

notation

$$[z', z] = \{z', z\} dz^2 \quad (52a)$$

for the corresponding quadratic differential (recall that z' and z are analytically related), the compatibility of equation (50) requires that

$$[z_1, z_2] + [z_2, z_1] = 0 \quad (52b)$$

$$[z_1, z_2] + [z_2, z_3] + [z_3, z_1] = 0 \quad (52c)$$

$$[z_1, z_2] = \left[\frac{az_1 + b}{cz_1 + d}, z_2 \right] = \left[z_1, \frac{az_2 + b}{cz_2 + d} \right] \quad (52d)$$

In equation (52c), it is of course assumed that $z_1 = f_{12}(z_2)$, $z_2 = f_{23}(z_3)$ and $z_3 = f_{31}(z_1)$ with $f_{12} \circ f_{23} \circ f_{31}$ the identity transformation. Finally, if z_1 and z_2 are related by a linear fractional transformation, the corresponding quadratic differential vanishes, as a consequence of (52b) and (52d)

$$az_1z_2 + bz_1 + cz_2 + d = 0 \Leftrightarrow [z_1, z_2] = 0 \quad (52e)$$

The anharmonic or cross ratio of four complex numbers is invariant under Möbius transformations. To measure how a given transformation $y = f(x)$ departs locally from a fractional linear transformation, let us compare the anharmonic ratios of four points x_i and of their images. Setting $x_{ij} = x_i - x_j$, we compute the quantity

$$Q(x, y) = \frac{x_{13}x_{42}}{x_{12}x_{43}} - \frac{y_{13}y_{42}}{y_{12}y_{43}}$$

for values x_i close to a point x , of the form $x_i = x + t\epsilon_i$. Correspondingly we expand y_i to third order in t . A short calculation shows that the leading term in Q is of order t^2 , of the form

$$Q(x, y) = \frac{\epsilon_{13}\epsilon_{42}}{6\epsilon_{12}\epsilon_{43}} \left\{ \frac{f'''(x)}{f'(x)} - \frac{3}{2} \left(\frac{f''(x)}{f'(x)} \right)^2 \right\} \epsilon_{14}\epsilon_{23}t^2 + \dots$$

The properties (52b)–(52e) follow immediately. Let us finally observe that

$$\{f(x), x\} = -2(f')^{1/2} \frac{d^2}{dx^2} (f')^{-1/2} = \frac{d^2}{dx^2} \ln f' - \frac{1}{2} \left(\frac{d}{dx} \ln f' \right)^2 \quad (53)$$

We can now turn to an interpretation of equation (50) and hence of the anomaly. For that purpose, let us map conformally

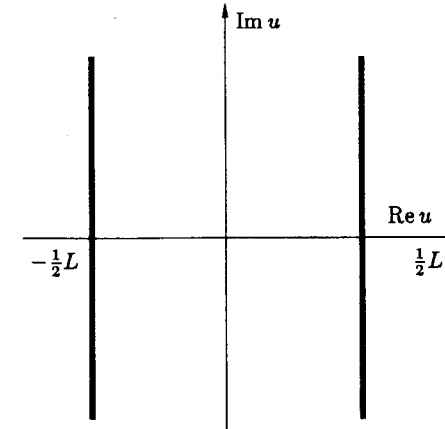


Fig. 2 A periodic strip of width L in the u -plane

the complex z -plane, with the point at the origin deleted, onto a periodic strip in the u -plane, of width L (figure 2). The corresponding map is exponential

$$z = \exp(2i\pi u/L) \quad (54)$$

Therefore

$$T_{\text{strip}}(u) = -(2\pi/L)^2 \left[T_{\text{plane}}(z) z^2 - \frac{1}{24} c \right] \quad (55)$$

and since $\langle T_{\text{plane}}(z) \rangle = 0$, we find

$$\langle T_{\text{strip}}(u) \rangle = \frac{1}{24} c (2\pi/L)^2 \quad (56)$$

This is the basis of an understanding of the anomaly as a Casimir effect, i.e. a shift in free energy due to the finite geometry. This effect was originally derived, then measured, in quantum electrodynamics where it gives rise to forces between neutral conductors. The very fact that it is measurable, although very tiny, is evidence of the long-range character of electromagnetic interactions. Similarly we expect in the present case an L -dependent shift in the free energy due to the confining boundary conditions. For the purpose of normalization, we assume that the free energy per unit area at criticality, in the infinite plane, is equal to zero. Including the extra factor $1/2\pi$, introduced above, the variation of the total free energy is given by a formula analogous

to (19), namely

$$\delta \ln Z = - \int_{\mathcal{D}} \frac{d^2 u}{2\pi} \langle T_{\mu\nu}(u, \bar{u}) \rangle \partial_\mu \delta r_\nu \quad (57)$$

where the domain of integration $\mathcal{D}$ is the infinite strip. Let us choose in particular $\delta r_1 = \delta \varepsilon u_1$, $\delta r_2 = 0$, meaning an horizontal dilatation of the strip. Of course, this is not a conformal transformation (it is rather a quasi conformal one, transforming circles into ellipses, with a fixed ratio of axes). But the definition of $T_{\mu\nu}$ does not imply conformal transformations either. The only nonvanishing term in $\partial_\mu \delta r_\nu$ is $\partial_1 \delta r_1 = \delta \varepsilon$, and $T_{11} = T(u) + \bar{T}(\bar{u})$. We face the difficulty that for an infinite strip $\ln Z$ is infinite. From translational invariance, we expect that there is a well-defined free energy per unit longitudinal length, in other words that

$$F(L) = \lim_{M \rightarrow \infty} \frac{1}{M} \ln Z(L, M) \quad (58)$$

exists, provided that an extra boundary condition at a distance M is imposed in the longitudinal direction. Assuming that the limit is independent of the longitudinal boundary condition, we conclude from equations (56)–(58), that

$$\delta F(L) = \delta \varepsilon L \frac{dF(L)}{dL} = -\frac{1}{12} c \frac{2\pi}{L} \delta \varepsilon$$

Hence

$$F(L) = \frac{1}{6} c \frac{\pi}{L} \quad (59a)$$

We have used the fact that the free energy per unit area in the infinite plane, F_0 , vanishes. Otherwise it would contribute an extra term $F_0 L$ to the r.h.s. of (59a). An equivalent statement if one uses a transfer matrix method, is that, at criticality, the logarithm of the largest eigenvalues λ_0 behaves as

$$\ln \lambda_0 = F_0 L + \frac{1}{6} c \frac{\pi}{L} \quad (59b)$$

as L tends to infinity. We shall discuss in section 2 the validity of equation (59) with respect to the assumption of a smooth limit starting from a finite domain. Clearly if some correlations grow with distance, as can occur in nonunitary models, the result is expected to be invalid. For the time being, assuming that all correlations decrease with distance, we see that a relation such

as (59) enables one to have access to the value of the central charge, as has been exemplified in numerical studies. For instance for polymers ($O(n)$ model $n \rightarrow 0$) or percolation (Q -state Potts model with $Q \rightarrow 1$), one finds as expected $c = 0$, while we shall see that the Ising model corresponds to $c = \frac{1}{2}$, the three-state Potts model to $c = \frac{4}{5}$ and the free Bose field (Gaussian model) to $c = 1$, to quote just a few of these values.

9.1.5 Virasoro algebra

The above mapping on a strip suggests the use of an operator formalism with a transfer matrix T . Since it corresponds to the use of polar coordinates in the plane, this formalism deserves the name of radial quantization. Time evolution amounts to dilatations and the eigenvalues of the corresponding generator are the conformal weights. To simplify notations, we absorb the previous factor $2i\pi/L$ in the definition of the variable u , so that

$$z = e^u = \rho e^{i\theta} \quad \ln \rho = \text{Re } u \quad \theta = \text{Im } u$$

$$T_{\text{strip}}(u) = T_{\text{plane}}(z) z^2 - \frac{1}{24} c \quad (60)$$

According to (60), the propagation axis is now along the real u -axis. Mean values in the strip are identified with traces of ordered products of operators (the ordering being along the “time” axis $\text{Re } u$). To an observable $A(u)$ this procedure associates an operator $\hat{A}(u)$. The transfer matrix is the analog of the exponential of a Hamiltonian, and relates operators at different values of $\text{Re } u$ according to

$$\hat{A}(\text{Re } u, \text{Im } u) = T^{-\text{Re } u} \hat{A}(0, \text{Im } u) T^{\text{Re } u} \quad (61)$$

Correlations are written for a decreasing sequence of values of $\text{Re } u$ as

$$\begin{aligned} & \langle A_1(u_1) \cdots A_n(u_n) \rangle_{\text{strip}} \\ &= \lim_{M \rightarrow \infty} \left[\text{Tr} \left(T^{\frac{1}{2}M - \text{Re } u_1} \hat{A}_1(0, \text{Im } u_1) T^{\text{Re } (u_1 - u_2)} \hat{A}_2(0, \text{Im } u_2) \cdots \right. \right. \\ & \quad \left. \left. \cdots \hat{A}_n(0, \text{Im } u_n) T^{\text{Re } u_n + \frac{1}{2}M} \right) \right] / \text{Tr } T^M \\ &= \lim_{M \rightarrow \infty} \frac{\text{Tr } T^M \hat{A}_1(u_1) \cdots \hat{A}_n(u_n)}{\text{Tr } T^M} \end{aligned} \quad (62)$$

If required, necessary subtractions are understood to convert these expressions to connected ones. In a quantum mechanical context, this is the Gell-Mann-Low formula. Let us assume for simplicity the model to be unitary. In Euclidean space, this means $T = \exp -H$, where H is Hermitian bounded from below. Let its unique ground state be denoted $|0\rangle$, with $\langle 0|$ the adjoint bra. For a unitary model, the limit $M \rightarrow \infty$ projects on the ground state if we assume a finite gap (in energy) to excited states. As the situation under consideration involves a strip with finite width, these assumptions look realistic. Therefore we find

$$\langle A_1(u_1) \dots A_n(u_n) \rangle = \langle 0 | \hat{A}_1(u_1) \dots \hat{A}_n(u_n) | 0 \rangle \quad (63)$$

$\text{Re } u_1 > \text{Re } u_2 > \dots > \text{Re } u_n$

For a real observable A , the corresponding operator is Hermitian

$$\begin{aligned} A^\dagger(0, \text{Im}u) &= \hat{A}(0, \text{Im}u) \\ A^\dagger(u) &= \hat{A}(-\bar{u}) \end{aligned} \quad (64)$$

and the correlations satisfy the reflection property

$$\langle 0 | \hat{A}_1(u_1) \dots \hat{A}_n(u_n) | 0 \rangle^* = \langle 0 | \hat{A}_n(-\bar{u}_n) \dots \hat{A}_1(-\bar{u}_1) | 0 \rangle \quad (65)$$

The second argument $\bar{u}$ has not been written explicitly, but is understood. With the help of the exponential map, we can view time translations in the u -plane as dilatations in the (punctured) z -plane, provided we keep in mind the correspondence between operators

$$\hat{A}_{\text{strip}}(u, \bar{u}) = \hat{A}_{\text{plane}}(z, \bar{z}) z^h \bar{z}^{\bar{h}} \quad (66)$$

As a result

$$\langle A_1(z_1) \dots A_n(z_n) \rangle = \langle 0 | \hat{A}_1(z_1) \dots \hat{A}_n(z_n) | 0 \rangle \quad (67)$$

$|z_1| > |z_2| > \dots > |z_n|$

Under a local conformal map $z \rightarrow z' = z + \delta\epsilon g(z)$, with $g(z)$ analytic, a primary field behaves as

$$\delta \hat{A}(z) = \delta\epsilon \left[h \frac{dg(z)}{dz} + g(z) \frac{d}{dz} \right] \hat{A}(z) \quad (68)$$

and equation (39) implies the existence of an operator $\hat{T}(z)$ such that for $|z_1| > |z_2| > \dots > |z_n|$,

$$\begin{aligned} \sum_{p=1}^n \langle 0 | \hat{A}_1(z_1) \dots \delta \hat{A}_p(z_p) \dots \hat{A}_n(z_n) | 0 \rangle \\ = \delta\epsilon \left\{ \int_C \frac{dz}{2i\pi} g(z) \langle 0 | \hat{T}(z) \hat{A}_1(z_1) \dots \hat{A}_n(z_n) | 0 \rangle \right. \\ \left. - \int_{C'} \frac{dz}{2i\pi} g(z) \langle 0 | \hat{A}_1(z_1) \dots \hat{A}_n(z_n) \hat{T}(z) | 0 \rangle \right\} \end{aligned}$$

where C encircles all points z_i while C' leaves all these points outside. Introducing intermediate contours, we see that this is equivalent to the operator statement

$$\delta \hat{A}(z) = \delta\epsilon \left\{ \oint_{|z'|>|z|} \frac{dz'}{2i\pi} g(z') \hat{T}(z') \hat{A}(z) - \oint_{|z|>|z'|} \frac{dz'}{2i\pi} g(z') \hat{A}(z) \hat{T}(z') \right\}$$

As a particular case

$$\begin{aligned} \delta \hat{T}(z) &= \delta\epsilon \left\{ \left[2g'(z) + g(z) \frac{d}{dz} \right] \hat{T}(z) + \frac{1}{12} c g'''(z) \right\} \\ &= \delta\epsilon \left\{ \oint_{|z'|>|z|} \frac{dz'}{2i\pi} g(z') \hat{T}(z') \hat{T}(z) - \oint_{|z|>|z'|} \frac{dz'}{2i\pi} g(z') \hat{T}(z) \hat{T}(z') \right\} \end{aligned} \quad (70)$$

Expand now $\hat{T}$ as well as $\hat{\bar{T}}$ in Laurent series

$$\hat{T}(z) = \sum_{n=-\infty}^{+\infty} \frac{L_n}{z^{n+2}} \quad \hat{\bar{T}}(\bar{z}) = \sum_{n=-\infty}^{+\infty} \frac{\bar{L}_n}{\bar{z}^{n+2}} \quad (71)$$

The extra factor z^{-2} has its origin in relation (60). We substitute (71) into (70), with $g(z)$ also expanded in a Laurent series with arbitrary coefficients. Identifying the factor of each of those on both sides, we obtain a set of commutation relations defining two infinite Lie algebras, called Virasoro algebras

$$\begin{aligned} [L_j, L_k] &= (j-k)L_{j+k} + \frac{1}{12} c j(j^2-1) \delta_{j+k,0} \\ [\bar{L}_j, \bar{L}_k] &= (j-k)\bar{L}_{j+k} + \frac{1}{12} c j(j^2-1) \delta_{j+k,0} \\ [L_j, \bar{L}_k] &= 0 \end{aligned} \quad (72)$$

These algebras were originally introduced by Virasoro in the study of the dual model of particle physics, an ancestor of string theory, as a means of enforcing what we now interpret as reparametrization invariance. Viewed in the z -plane, the origin plays no special role, and $\hat{T}$ and $\bar{\hat{T}}$ could be expanded around any other point obtained as the image of the origin through a Möbius transformation.

The formulas (72) can be understood as follows. Suppose first that $c = 0$, and look at one of the two equivalent algebras. One observes that the operators can be realized as first order differential operators of the form

$$\ell_m = -z^{1+m} \frac{\partial}{\partial z} \quad [\ell_j, \ell_k] = (j-k)\ell_{j+k} \quad (73)$$

If we restrict z to be of unit modulus, the ℓ_m 's act on functions defined on the circle, and are generators for the diffeomorphisms of the unit circle, meaning one-to-one infinitely differentiable maps (here connected to the identity). The extra term in (72) proportional to c commutes with all elements in the algebra. It therefore indicates an extension of the Lie algebra (73) by a one-dimensional algebra. Since the latter commutes with all other elements, one speaks of a central extension, and hence of a central charge. The situation is the analog at the Lie algebra level of the one which prevails in quantum mechanics when one looks for representations of symmetries, up to a phase. For instance, half integral spins exhibit a central extension of $SO(3)$ by Z_2 , in the form of SU_2 . A simple calculation shows that, up to a change of basis, the form indicated in equation (72) is the only central extension of the algebra (73).

The operators L_0 , $L_{\pm 1}$, or $\bar{L}_0$, $\bar{L}_{\pm 1}$, or ℓ_0 , $\ell_{\pm 1}$, generate in each case a closed three-dimensional algebra, which is nothing but the complex Lie algebra of the global conformal group $SL(2, \mathbb{C})$. Note that correspondingly, the coefficient of the anomaly vanishes. Furthermore, L_0 , the generator of dilatations, is diagonal in the adjoint representation $[L_0, L_{-p}] = pL_{-p}$, with integral eigenvalues. It serves to define a grading of the enveloping algebra generated by the L 's, i.e. a splitting in homogeneous subspaces with a definite eigenvalue for L_0 . Said otherwise, L_{-p} acting on an eigenstate of L_0 increases the eigenvalue by p .

The vacuum state is invariant under global conformal transformations. As a result, it is annihilated by $L_0, L_{\pm 1}$. But since $T(z)|0\rangle$ remains regular as $z \rightarrow 0$ (as well as $\langle 0|T(z)$ as z goes to infinity) it follows that one has the stronger properties

$$\begin{aligned} p \geq -1 & \quad L_p |0\rangle = 0 \\ p \leq 1 & \quad \langle 0|L_p = 0 \end{aligned} \quad (74)$$

with similar equations for $\bar{L}_p$.

- i) Show that in order to obtain (74), it is sufficient to require that beyond $L_0, L_{\pm 1}$, L_2 annihilates the state $|0\rangle$ to the right, or L_{-2} annihilates $\langle 0|$ to the left.
- ii) As a consequence of (74), $\langle 0|L_p|0\rangle = 0$, for any p . Rederive from the commutation relation (72) alone that, for $|z_1| > |z_2|$ say

$$\langle 0|\hat{T}(z_1)\hat{T}(z_2)|0\rangle = \frac{\frac{1}{2}c}{z_{12}^4}$$

In a unitary model, a Hermitian observable satisfies the relation

$$\hat{A}(z, \bar{z})^\dagger = \bar{z}^{-2h} z^{-2\bar{h}} \hat{A}(\bar{z}^{-1}, z^{-1}) \quad (75)$$

When applied to T or $\bar{T}$, this entails that a necessary condition for unitarity of a representation is that

$$L_p^\dagger = L_{-p} \quad , \quad \bar{L}_p^\dagger = \bar{L}_{-p} \quad (76)$$

Equation (40) implies that, as $z_1 \rightarrow z_2$, $\hat{T}(z_1)\hat{A}(z_2)|0\rangle$ behaves as $h_A z_{12}^{-2} \hat{A}(z_2)|0\rangle$. Let us define the state $|h, \bar{h}\rangle$ through

$$\begin{aligned} \lim_{z \rightarrow 0} \hat{A}(z, \bar{z})|0\rangle &= |h, \bar{h}\rangle \\ \langle h, \bar{h}| &= \lim_{z \rightarrow 0} \langle 0|\hat{A}^\dagger(z, \bar{z}) = \lim_{z \rightarrow \infty} \langle 0|\hat{A}(z, \bar{z}) z^{2h} \bar{z}^{2\bar{h}} \end{aligned} \quad (77a)$$

It then follows that

$$\begin{aligned} p > 0 & \quad L_0 |h, \bar{h}\rangle = h |h, \bar{h}\rangle & \bar{L}_0 |h, \bar{h}\rangle = \bar{h} |h, \bar{h}\rangle \\ & \quad L_p |h, \bar{h}\rangle = \bar{L}_p |h, \bar{h}\rangle = 0 \end{aligned} \quad (77b)$$

and similar conjugate relations

$$\begin{aligned} p < 0 & \quad \langle h, \bar{h}|L_0 = \langle h, \bar{h}|h & \langle h, \bar{h}|\bar{L}_0 = \langle h, \bar{h}|\bar{h} \\ & \quad \langle h, \bar{h}|L_p = \langle h, \bar{h}|\bar{L}_p = 0 \end{aligned}$$

In a unitary representation the generator of (real) dilatations, $L_0 + \bar{L}_0$, and the one of rotations, $L_0 - \bar{L}_0$, are both Hermitian.

We conclude that the space of states in a critical model carries a representation of the product of algebras $\{L_p\} \times \{\bar{L}_p\}$, which we denote $V \times \bar{V}$. Such a representation will be in general reducible, containing states such as $|h, \bar{h}\rangle$ and their "descendants" i.e. states generated by applying a product of L_p and $\bar{L}_p$'s.

Let us focus our attention on one of these algebras, V say. By isomorphism the reasoning applies to $\bar{V}$ as well. Let us consider a representation with a so-called highest weight vector, i.e. a state $|h\rangle$ fulfilling the requirements (77). We henceforth omit the label $\bar{h}$ pertaining to $\bar{V}$. The name highest weight vector derives from the terminology used in the representation theory of Lie algebras. In the present context, it is rather a ground state in a given sector. Consider the so-called *Verma module* generated by $|h\rangle$, i.e. the infinite-dimensional vector space spanned by linear combinations of monomials in the L 's applied to $|h\rangle$. For our purpose, we may think of finite linear combinations, so that we need not worry about convergence properties. This space splits into homogeneous subspaces characterized by the eigenvalues $h+n$ of L_0 , where n is a non-negative integer, which will be called the level. The quantity $h+n$ is sometimes also called the weight of the corresponding states. Here it is nothing but a conformal weight of a nonprimary field which creates this state when applied to the vacuum state $|0\rangle$. When considering the states of level n , we may use the commutation relations (72) as well as the properties of the state $|h\rangle$ annihilated by $L_p, p > 0$, to show that they all reduce to combinations of some basic ones, involving only $L_p, p < 0$, and ordered conveniently, for instance states of the form

$$|\{p\}, h\rangle \equiv L_{-p_1} L_{-p_2} \dots L_{-p_k} |h\rangle \quad (78a)$$

$$0 < p_1 \leq p_2 \leq \dots \leq p_k, \quad k \leq n, \quad p_1 + p_2 + \dots + p_k = n$$

Similarly in the conjugate space

$$\langle h, \{p\}| \equiv \langle h| L_{p_k} \dots L_{p_1} \quad (78b)$$

with the same constraints on the indices. Such infinite-dimensional spaces carry obviously a representation of the algebra.

The Virasoro algebra is closely akin to the Kac-Moody algebras defined as infinite-dimensional extensions of semisimple Lie alge-

bras (appendix C) which occur in the study of two-dimensional current algebras as well as in infinite-dimensional integrable systems. In both cases, we deal with graded Lie algebras, and the techniques for handling them are closely related.

Applying an operator $L_{p'}, p' > 0$ to a state such as $|\{p\}, h\rangle$ decreases its level by p' . Therefore if a product of such operators has a degree equal to n , it will produce a multiple of $|h\rangle$ when applied to $|\{p\}, h\rangle$. We write

$$L_{p'_q} \dots L_{p'_1} L_{-p_1} \dots L_{-p_k} |h\rangle = (\{p'\}, \{p\})_{c,h} |h\rangle \quad (79a)$$

$$0 < p_1 \leq p_2 \leq \dots \leq p_k \quad 0 < p'_1 \leq p'_2 \leq \dots \leq p'_q$$

$$\Sigma p'_i = \Sigma p_j = n$$

The coefficients $(\{p'\}, \{p\})_{c,h}$ for fixed c, h and level n can be computed solely from the structure of the Virasoro algebra and the properties of the state $|h\rangle$. They form a matrix indexed by partitions of the integer n , called the contragredient form. In the unitary case, this contragredient form is the matrix of scalar products at level n ,

$$(\{p'\}, \{p\})_{c,h} = \frac{\langle h, \{p'\} | \{p\}, h \rangle}{\langle h | h \rangle} \quad (79b)$$

where we recall that

$$L_0 |\{p\}, h\rangle = (h + \Sigma p_i) |\{p\}, h\rangle \quad (80)$$

$$\langle h, \{p\} | L_0 = \langle h, \{p\} | (h + \Sigma p_i)$$

To achieve unitarity, we require that at each level n , the contragredient form is positive semidefinite. If it is definite positive, the Verma module itself yields a unitary irreducible representation. The remaining option, when the contragredient form is only semidefinite positive, is that there exist "null states" generating invariant subspaces.

Since for any $p > 0$

$$L_p L_{-p} |h\rangle = [L_p, L_{-p}] |h\rangle = [2ph + \frac{1}{12}cp(p^2 - 1)] |h\rangle \quad (81)$$

a necessary condition for unitarity is $h \geq 0$ and $c \geq 0$. We therefore find only positive weights in unitary models. But negative ones, which correspond to correlations growing with increasing distance, are *a priori* not excluded in a nonunitary situation. Thus we have

$$\text{unitarity} \implies c \geq 0, \quad h \geq 0 \quad (82)$$

In any representation the number of linearly independent states at level n cannot exceed $p(n)$, the number of partitions of the integer n . If one agrees to define $p(0) = p(1) = 1$, Euler's generating function for partitions is

$$\sum_0^\infty p(n)q^n = \frac{1}{\prod_1^\infty (1 - q^n)} \equiv \frac{1}{P(q)} \quad (83)$$

where for convergence $|q| < 1$. Dedekind's function $\eta(\tau) = q^{1/24}P(q)$, where $q = \exp 2i\pi\tau$, will play a prominent role in the following discussion.

A means of studying the contragredient form is to compute its determinant at each level n (at level 0 it is 1 by definition). These determinants, first obtained by Kac

$$\det_n(c, h) = \det(\{p'\}, \{p\})_{c,h} \quad (84)$$

are obviously polynomials in c and h . They enjoy the property that $\det_n(c, h)$ vanishes if and only if at some level $n' \leq n$ there exists a linear combination

$$|s\rangle = \sum \alpha_{\{p\}} L_{-p_1} \cdots L_{-p_k} |h\rangle \quad \sum p_i = n' \leq n$$

which is annihilated by all $L_p, p > 0$. If such a vector exists at level $n' \leq n$, it generates at level n a subspace of vectors annihilated by the $L_p, p > 0$. Conversely if the determinant vanishes, some of its columns satisfy linear relations. The corresponding vectors of the form $\sum \alpha_{\{p\}} L_{-p_1} \cdots L_{-p_k} |h\rangle, \sum p_i = n$ are annihilated by all products of $L_p, \sum p_i = n$. A recursive reasoning shows that there must exist a singular vector of degree $n' \leq n$. If such singular (or "null") vectors exist with $n' > 0$, in order that the representation of V be reduced, the invariant subspace that they generate must be eliminated using an equivalence relation. This means that the field that creates these singular vectors when applied to the vacuum state $|0\rangle$ must be set equal to zero. Since the Kac determinants are polynomials in h , this will enable us to single out for fixed c special values for the conformal weights.

These determinants can at first be computed by merely using the commutation relations. Beyond the trivial cases

$$\det_0(c, h) = 1 \quad \det_1(c, h) = 2h \quad (85)$$

where the vanishing of $\det_1$ for $h = 0$ reflects the fact that $|0\rangle$ itself is a null vector, let us look at the instructive case $n = 2$ with

$$(\{p'\}, \{p\})_{c,h} = \begin{pmatrix} 4h + \frac{1}{2}c & 6h \\ 6h & 4h(2h + 1) \end{pmatrix} \quad (86a)$$

where the lines or columns are indexed respectively by the partitions $\{2\}$ and $\{1, 1\}$. Therefore

$$\det_2(c, h) = 2h [16h^2 + 2(c - 5)h + c] \quad (86b)$$

We note that $\det_2$ contains $\det_1$ as a factor. The new information is the existence of two new zeroes, occurring for

$$h_\pm = \frac{5 - c \pm \sqrt{(1 - c)(25 - c)}}{16} \quad (87)$$

These zeroes are complex conjugate for c real between 1 and 25, and real otherwise. We exclude the former case from consideration and look at the corresponding singular vectors of level two, i.e. with weight $h_\pm + 2$. We can check directly that

$$|s_\pm\rangle = \left(\frac{3}{2(2h_\pm + 1)} L_{-1}^2 - L_{-2} \right) |h\rangle \quad (88a)$$

satisfy

$$L_0 |s_\pm\rangle = (h_\pm + 2) |s_\pm\rangle \quad L_p |s_\pm\rangle = 0 \quad p > 0 \quad (88b)$$

To verify (88b), we observe that L_1 and L_2 generate, through commutation, all the $L_p, p > 0$. Let us introduce the following useful parametrization of the central charge, for any real or complex, finite or infinite m

$$c = 1 - \frac{6}{m(m + 1)} \quad (89)$$

The above values $h_\pm$ take the form

$$\begin{aligned} h_+ &= \frac{m + 3}{4m} = \frac{[2(m + 1) - m]^2 - 1}{4m(m + 1)} \\ h_- &= \frac{m - 2}{4(m + 1)} = \frac{[m + 1 - 2m]^2 - 1}{4m(m + 1)} \end{aligned} \quad (90)$$

All this is stated in operator language. We can relate it also to short-distance expansions of a primary field A (of weight h) with a product of fields $T(z)$. For short, we omit again the argument

$\bar{z}$ of A . Equations (40) can be reinterpreted by saying that, when inserted in correlation functions, one has the short-distance expansion

$$\begin{aligned} T(z)A(u) &= \frac{A^{(0)}(u)}{(z-u)^2} + \frac{A^{(-1)}(u)}{(z-u)} + A^{(-2)}(u) + (z-u)A^{(-3)} + \dots \\ A^{(0)}(u) &= hA(u) \\ A^{(-1)}(u) &= \frac{\partial A(u)}{\partial u} \\ &\dots \end{aligned} \quad (91a)$$

as well as

$$T(z)T(u) = \frac{\frac{1}{2}c}{(z-u)^4} + \frac{2T(u)}{(z-u)^2} + \frac{1}{(z-u)} \frac{\partial T(u)}{\partial u} + \dots \quad (91b)$$

We have therefore a correspondence $A(0) \leftrightarrow |h\rangle$, $A^{(0)}(0) \leftrightarrow L_0|h\rangle$, $A^{(-p)}_{(0)} \leftrightarrow L_{-p}|h\rangle$. One can repeat with the derived fields $A^{(-p)}$ the above short-distance expansion defining $A^{(-p_1, -p_2)}$ in correspondence with $L_{-p_1}L_{-p_2}|h\rangle$, and so on.

The primary field A or $A^{(0)} = hA(z)$ has known transformation properties. Similarly for $A^{(-1)}(u) = \partial_u A(u)$. It is instructive to pursue the idea, in order to obtain an alternative interpretation of singular vectors in terms of the corresponding fields. Recall that if A is primary

$$\begin{aligned} A(z) &= A'(\xi) \left(\frac{d\xi}{dz} \right)^h \\ T(z) &= T'(\xi) \left(\frac{d\xi}{dz} \right)^2 + \frac{1}{12}c \{ \xi, z \} \end{aligned} \quad (92)$$

It is obvious that $A^{(-1)}$ does not transform as a primary field. To find the behaviour of $A^{(-2)}$, we compare the short-distance expansions for $T(z')A(z)$ and $T'(\xi')A'(\xi)$, using the transformation laws (92). Thus

$$\begin{aligned} hA(z) + (z' - z)\partial A(z) + (z' - z)^2 A^{(-2)}(z) + \dots \\ = hA'(\xi) \left(\frac{z' - z}{\xi' - \xi} \right)^2 \left(\frac{d\xi}{dz} \right)^h \left(\frac{d\xi'}{dz'} \right)^2 \\ + (z' - z) \frac{z' - z}{\xi' - \xi} \left(\frac{d\xi}{dz} \right)^h \left(\frac{d\xi'}{dz'} \right)^2 \partial_\xi A'(\xi) \end{aligned}$$

$$\begin{aligned} + (z' - z)A^{(-2)}(\xi) \left(\frac{d\xi}{dz} \right)^h \left(\frac{d\xi'}{dz'} \right)^2 \\ + \frac{1}{12}c(z' - z) \{ \xi', z' \} A'(\xi) \left(\frac{d\xi}{dz} \right)^h + \dots \end{aligned}$$

After some rather laborious calculations, we find

$$\begin{aligned} A^{(-2)}(z) &= A^{(-2)}(\xi) \left(\frac{d\xi}{dz} \right)^{h+2} \\ &+ hA'(\xi) \left(\frac{d\xi}{dz} \right)^h \left[\frac{2}{3} \frac{d^3\xi/dz^3}{d\xi/dz} - \frac{1}{4} \left(\frac{d^2\xi/dz^2}{d\xi/dz} \right)^2 \right] \\ &+ \frac{3}{2} \frac{dA'(\xi)}{d\xi} \left(\frac{d\xi}{dz} \right)^{h+1} \frac{d^2\xi/dz^2}{d\xi/dz} \\ &+ \frac{1}{12}cA'(\xi) \left(\frac{d\xi}{dz} \right)^h \left[\frac{d^3\xi/dz^3}{d\xi/dz} - \frac{3}{2} \left(\frac{d^2\xi/dz^2}{d\xi/dz} \right)^2 \right] \end{aligned} \quad (93a)$$

We let the courageous reader find the corresponding formula for $A^{(-p_1, -p_2, \dots)}$! To say the least, the derived fields have rather complicated transformation laws. But let us compare (93a) with the behaviour of $\partial_z^2 A(z)$ corresponding to $L_{-1}^2|h\rangle$ we have

$$\begin{aligned} \partial_z^2 A(z) &= \partial_\xi^2 A'(\xi) \left(\frac{d\xi}{dz} \right)^{h+2} + \partial_\xi A'(\xi)(2h+1) \left(\frac{d\xi}{dz} \right)^h \frac{d^2\xi}{dz^2} \\ &+ hA'(\xi) \left(\frac{d\xi}{dz} \right)^h \left[\frac{d^3\xi/dz^3}{d\xi/dz} + (h-1) \left(\frac{d^2\xi/dz^2}{d\xi/dz} \right)^2 \right] \end{aligned} \quad (93b)$$

If we now form the combination

$$\chi_\pm(z) = \frac{3}{2(2h_\pm + 1)} \partial_z^2 A(z) - A^{(-2)}(z) \quad (94a)$$

for either value $h_\pm$ given by (87) or (90), it is straightforward to verify that

$$\chi_\pm(z) = \chi'_\pm(\xi) \left(\frac{d\xi}{dz} \right)^{h_\pm+2} \quad (94b)$$

In other words, singular vectors are associated to combinations of derived fields, arising from the short-distance expansion with the energy-momentum tensor, which transform as primary fields with an incremented weight. This is a useful equivalence. It shows that if we call χ as above such a combination, requiring χ to vanish is a consistent (covariant) condition. This amounts to factoring out

invariant subspaces leaving irreducible representations. Finally, such conditions lead to partial differential equations satisfied by correlation functions involving the original field. Of course, its weight h has to be selected as a zero of a Kac determinant.

Let us consider a field A_1 with $h_1 \equiv h_{\pm}$. Requiring $\chi_{\pm} = 0$, leads to second order partial differential equations of the form

$$\frac{3}{2(2h_1 + 1)} \frac{\partial^2}{\partial z_1^2} \langle A_1(z_1) \dots A_n(z_n) \rangle \\ = \sum_{p=2}^n \left[\frac{h_p}{(z_1 - z_p)^2} + \frac{1}{(z_1 - z_p)} \partial_{z_p} \right] \langle A_1(z_1) \dots A_n(z_n) \rangle \quad (95)$$

obtained from equation (40) by using the short-distance expansion of T with A_1 . Higher order equations will correspond to singular vectors of higher level. We see that the theory offers the possibility of being able to discuss interesting circumstances involving both the selection of specific values of the conformal weights (or critical exponents) and the determination of the corresponding correlation functions. This appears an amazing achievement. It is therefore quite useful that an explicit expression is known for the Kac determinant at any level.

Show that at level 3

$$\det_3(c, h) = 48h^2 [16h^2 + 2(c-5)h + c] [3h^2 + (c-7)h + c + 2]$$

and obtain the corresponding singular vectors. Show that in general, $\det_{n-1}$ is a factor of $\det_n$.

9.1.6 The Kac determinant

The following remarkable expression was found by Kac

$$\det_n(c, h) = \text{cst} \times \prod_{\substack{r,s=1 \\ 1 \leq r \leq s \leq n}}^n (h - h_{r,s})^{p(n-rs)} \quad (96)$$

where the central charge is parametrized as in (89), the zeroes $h_{r,s}$ are indexed by two positive integers, r and s , and read

$$h_{r,s} = \frac{[r(m+1) - sm]^2 - 1}{4m(m+1)} \quad (97)$$

A given zero occurs first at level $n' = \text{Inf}(rs)$ and propagates with a multiplicity $p(n - n')$ at level n (the infimum is over the set of pairs r, s giving the same value for $h_{r,s}$). One can check (96) for small values of n by direct computation. The parametrization $c = 1 - 6/m(m+1)$ is singular for $c = 1$ or 25 . These cases require special consideration.

The proof given by Feigin and Fuchs of the fundamental result (96), (97) is involved but interesting, so we shall reproduce it below. As the values obtained for the weights solve in essence the question of finding two-dimensional critical exponents, it rightly deserves to be presented in some detail, but can, however, be skipped without harm, should the reader wish to do so.

It is not difficult to obtain the degree of $\det_n(c, h)$ as a polynomial in h . One readily convinces oneself that the term of highest degree arises from the product of diagonal terms in the contragredient form, each of which is the coefficient of $|h\rangle$ in

$$L_n^{\alpha_n} \dots L_1^{\alpha_1} L_{-1}^{\alpha_{-1}} \dots L_{-n}^{\alpha_{-n}} |h\rangle \quad \sum_1^n k\alpha_k = n \quad \alpha_k \geq 0$$

of degree $\sum_1^n \alpha_k$ in h . We therefore have the following formula for the degree ρ_n

$$\rho_n = \sum_{\alpha_k \geq 0, \sum_1^n k\alpha_k = n} \left(\sum_{k=1}^n \alpha_k \right) \quad (98)$$

where the first sum is over partitions of n , with α_k the number of times a particular integer $1 \leq k \leq n$ occurs in such a partition. We set $\rho_0 \equiv 1$ and write a generating function, for $|q| < 1$, as

$$\sum_0^\infty \rho_n q^n = t \frac{\partial}{\partial t} \sum_{\alpha_1, \alpha_2, \dots=0}^\infty t^{\sum_1^\infty \alpha_k} q^{\sum_1^\infty k\alpha_k} \Big|_{t=1} \\ = t \frac{\partial}{\partial t} \frac{1}{\prod_1^\infty (1 - tq^k)} \Big|_{t=1} = \frac{1}{P(q)} \sum_{r=1}^\infty \frac{q^r}{1 - q^r} \\ = \sum_{n'=0}^\infty p(n') q^{n'} \sum_{r,s=1}^\infty q^{rs} = \sum_{n=0}^\infty q^n \left[\sum_{\substack{r,s \geq 1 \\ rs \leq n}} p(n - rs) \right] \quad (99)$$

which is the expected result.

As a side remark, Coste has obtained the numerical, c -independent coefficient, omitted in (96). This prefactor is, with ρ_n as in equation (98)

$$2^{\rho_n} \prod_{\text{partitions of } n} \alpha_1! \cdots \alpha_n! 1^{\alpha_1} \cdots n^{\alpha_n}$$

Using their notations, let us now turn to the proof of Feigin and Fuchs. It involves four stages.

- i) Introduce the set $F_{\lambda, \mu}$ of Laurent series (in the sequel we think of $z = e^u$, allowing for arbitrary powers in z)

$$z^\mu f(z) = \sum_{-\infty}^{+\infty} f_k z^{k+\mu} \quad (100)$$

of weight $-\lambda$, i.e. transforming as $f(z) \mapsto f'(\xi)$, with

$$z^\mu f(z) = \xi^\mu f'(\xi) \left(\frac{d\xi}{dz} \right)^{-\lambda} \quad (101)$$

in an analytic map $z \mapsto \xi$. Of course, the prime on $f'(\xi)$ does not mean a derivative. Taking a basis $f_{(k)}(z) = z^k$, we have for an infinitesimal transformation $\xi = z - \delta \varepsilon z^{1-j}$ for j integer

$$\begin{aligned} f_{(k)} &\rightarrow f'_{(k)} = f_{(k)} + \delta f_{(k)} & \delta f_{(k)} &= \delta \varepsilon (\ell_j f_{(k)}) \\ \ell_j f_{(k)} &= [\mu + k + \lambda(j-1)] f_{(k-j)} & [\ell_j, \ell_k] &= (j-k) \ell_{j+k} \end{aligned} \quad (102)$$

showing that the set $\{\ell_j\}$ gives a representation of the Lie algebra of diffeomorphisms of the unit circle, generalizing (73). Provided that $\lambda + \lambda' + 1 = \mu + \mu' + p + 1 = 0$, with p an integer, the integral

$$\int \frac{dz}{2i\pi z^{1+p}} g_{\lambda'\mu'}(z) f_{\lambda\mu}(z) \quad (103a)$$

gives an invariant bilinear form on $F_{\lambda'\mu'} \otimes F_{\lambda\mu}$. Similarly when $\lambda + \lambda' + 1 = \mu - \mu' + 2\lambda' + p + 1 = 0$, the contragredient form

$$\int \frac{dz}{2i\pi z^{1+p}} g_{\lambda'\mu'}(z^{-1}) f_{\lambda\mu}(z) \quad (103b)$$

is also invariant. Consider now the n th exterior product $\bigwedge^n F_{\lambda\mu}$. Let us show that the determinant

$$\varphi(z) = \begin{vmatrix} f_0 & \frac{df_0}{dz} & \cdots & \frac{d^{n-1}}{dz^{n-1}} f_0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{n-1} & \frac{df_{n-1}}{dz} & \cdots & \frac{d^{n-1}}{dz^{n-1}} f_{n-1} \end{vmatrix}$$

defines a map $\bigwedge^n F_{\lambda\mu} \rightarrow F_{n\lambda - \frac{1}{2}n(n-1), n\mu}$. For that purpose we associate to z the n -component vector $f_0(z), \dots, f_{n-1}(z)$. Consider now n points $z_0, \dots, z_{n-1}$. Under a map $z \mapsto \xi$, by definition

$$\det f'_i(\xi_j) = \prod_{k=1}^n \left(\frac{z_k}{\xi_k} \right)^\mu \left(\frac{dz_k}{d\xi_k} \right)^{-\lambda} \det f_i(z_j)$$

If $z_i = z + \varepsilon_i$, where all ε_i tend to zero and are of the same order, we have

$$\det f_i(z + \varepsilon_j) \sim \frac{1}{\prod_0^{n-1} k!} \prod_{0 \leq s < r \leq n-1} (\varepsilon_r - \varepsilon_s) \varphi(z)$$

Similarly $\xi_i = \xi + \varepsilon_i$ $d\xi/dz + \dots$, so that, comparing the leading terms,

$$\varphi'(\xi) = \left(\frac{z}{\xi} \right)^{n\mu} \left(\frac{dz}{d\xi} \right)^{\frac{1}{2}n(n-1) - n\lambda} \varphi(z)$$

which is the required property.

The constant function belongs to $F_{0,0}$. As a result, using (103), if

$$\lambda = (n-2)(n+1)/2n \quad \mu = -1 - p/n \quad (104a)$$

we conclude that whenever n functions f_j belong to $F_{\lambda, \mu}$ the integral over the unit circle

$$\gamma = \oint \frac{dz}{2i\pi z^{p+n}} \det \left[\frac{d^k f_j}{dz^k} \right]_{0 \leq k, j \leq n-1} \quad (104b)$$

is an invariant quantity. The definition (103a) with $p = 0$ enables one to assert the existence of an element of $\bigwedge^n F_{-1-\lambda, -1-\mu} \equiv \bigwedge^n F_{-(n-1)(n+2)/2n, p/n}$, the dual of $\bigwedge^n F_{\lambda, \mu}$ corresponding to the invariant γ . All what we have to do is to substitute in (104b) the expansion

$$\frac{d^k}{dz^k} f_j(z) = \sum_{q=-\infty}^{+\infty} q(q-1) \cdots (q-k+1) z^{q-k} \oint \frac{dz'}{2i\pi z'^{q+1}} f_j(z')$$

Consequently with P denoting a permutation on n symbols, $(-1)^P$ the signature of P , we find

$$\begin{aligned} \gamma &= \oint \frac{dz}{2i\pi z^{p+n}} \sum_P (-1)^P \sum_{q_0, \dots, q_{n-1} = -\infty}^{\infty} \int \frac{1}{(2i\pi)^n} \frac{dz_0 \cdots dz_{n-1}}{z_0^{1+q_0} \cdots z_{n-1}^{1+q_{n-1}}} \\ &\quad \times f_{P_0}(z_0) \cdots f_{P_{n-1}}(z_{n-1}) z^{\sum_{i=0}^{n-1} q_i - \frac{1}{2}n(n-1)} \end{aligned}$$

$$\begin{aligned}
& \times q_1 q_2 (q_2 - 1) \cdots q_{n-1} (q_{n-1} - 1) \cdots (q_{n-1} - n + 2) \\
& = \sum_{q_0, \dots, q_{n-1} = -\infty}^{+\infty} \delta_{\sum_{i=0}^{n-1} q_i, \frac{1}{2}n(n-1)+n-1+p} \oint \prod_{r=0}^{n-1} \left(\frac{dz_r f_r(z_r)}{2i\pi z_r^{1+q_r}} \right) \\
& \times \det \begin{vmatrix} 1 & q_0 & \cdots & q_0^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & q_{n-1} & \cdots & q_{n-1}^{n-1} \end{vmatrix}
\end{aligned}$$

The last determinant follows from Vandermonde's formula. Let us change all q_i into $-q_i$, and recall that the basis vectors z^j were denoted $f_{(j)}$. We then reach the conclusion that

$$\gamma = \frac{1}{n!} \sum_{\substack{q_0, \dots, q_{n-1} = -\infty \\ \sum q_i = -\frac{1}{2}n(n+1)+1-p}}^{+\infty} \prod_{0 \leq r < s \leq n-1} (q_r - q_s) \det (f_{(q_i)}, f_{(j)})_0$$

where $(\ , \)_0$ is the bilinear form (103a) for $p = 0$. We read from this expansion that the corresponding invariant element $\varphi_{n,p}$, belonging to $\wedge^n F_{-(n-1)(n+2)/2n, p/n}$ is given by

$$\varphi_{n,p} = \sum_{\substack{q_0, \dots, q_{n-1} = -\infty \\ \sum q_i = -\frac{1}{2}n(n+1)+1-p}}^{+\infty} \left[\prod_{0 \leq r < s \leq n-1} (q_r - q_s) \right] f_{(q_0)} \wedge \cdots \wedge f_{(q_{n-1})} \quad (105)$$

ii) In the second stage, we associate to $F_{\lambda, \mu}$ a fermionic Fock space and a corresponding representation of the Virasoro algebra. To each integer i , positive or negative, we associate a pair $a_i, a_i^\dagger$ of operators satisfying the usual canonical anticommutation relations

$$\{a_i, a_j^\dagger\} = \delta_{ij} \quad \{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \quad (106)$$

The reference "vacuum" state $|\Omega\rangle$ is taken by convention to satisfy

$$\begin{aligned}
i \geq 0 & \quad a_i |\Omega\rangle = 0 \\
i < 0 & \quad a_i^\dagger |\Omega\rangle = 0
\end{aligned} \quad (107)$$

meaning that all "particle" states such that $i < 0$ are occupied, and all those with $i \geq 0$ empty. Corresponding to the choice (107) we have a natural Wick ordering $:$, where the operators a_i ($i \geq 0$) or $a_i^\dagger$ ($i < 0$) are on the right of the a_i ($i < 0$) or $a_i^\dagger$ ($i \geq 0$), taking into account the sign of the permutation required to bring a given monomial in this order. With this construction,

we can associate two quadratic operators, a charge Q and a Hamiltonian H , defined as follows. First

$$Q = \sum_{-\infty}^{+\infty} : a_k^\dagger a_k : = \sum_{k \geq 0} a_k^\dagger a_k - \sum_{k < 0} a_k a_k^\dagger \quad (108)$$

annihilates $|\Omega\rangle$ and will commute with the representation of the Virasoro algebra. In a sector of charge $Q = n$, it is useful to introduce a new reference state $|\Omega, n\rangle$ (such that $|\Omega, 0\rangle \equiv |\Omega\rangle$). This state is annihilated by a_i ($i \geq n$) and $a_i^\dagger$ ($i < n$) and corresponds to filling one-particle states up to $n-1$. If the index is thought of as an "energy", then the Hamiltonian H reads

$$H = \sum_{-\infty}^{+\infty} k : a_k^\dagger a_k : = \sum_{k \geq 0} k a_k^\dagger a_k - \sum_{k < 0} k a_k a_k^\dagger \quad (109)$$

Of course, the one-particle state with $k = 0$ does not contribute to H , which is a positive operator and satisfies

$$H |\Omega, n\rangle = \frac{1}{2}n(n-1) |\Omega, n\rangle \quad (110)$$

valid for n positive or negative. Observe that $|\Omega, 0\rangle$ and $|\Omega, 1\rangle$ have the same zero energy. For any integer j different from zero, we define the second quantized form of (102), namely

$$L_j = \sum_k a_{k-j}^\dagger \{\mu + k + \lambda(j-1)\} a_k \quad (111)$$

$$j \neq 0 \quad [L_j, a_k^\dagger] = \{\mu + k + \lambda(j-1)\} a_{k-j}^\dagger \quad [L_j, Q] = 0$$

The Wick ordering is irrelevant here since k and $k-j$ are distinct. In computing commutators, we use

$$\begin{aligned}
[AB, C] &= A\{B, C\} - \{A, C\}B \\
[AB, CD] &= A\{B, C\}D - AC\{B, D\} + \{A, C\}DB - C\{A, D\}B
\end{aligned} \quad (112)$$

The above formulas for Q, H and L_j assume a more compact form if we introduce a pair of complex Fermi fields

$$\psi(z) = \sum_k a_k z^{-k} \quad \bar{\psi}(z) = \sum_k a_k^\dagger z^k \quad (113)$$

Then

$$\begin{aligned}
Q &= \oint \frac{dz}{2i\pi} : \bar{\psi}(z) \psi(z) : \\
H &= \frac{1}{2} \oint \frac{dz}{2i\pi} : \frac{\partial \bar{\psi}}{\partial z} \psi - \bar{\psi} \frac{\partial \psi}{\partial z} :
\end{aligned} \quad (114)$$

$$j \neq 0 \quad L_j = \oint \frac{dz}{2i\pi} z^j \left\{ \frac{(\mu - \lambda)}{z} \bar{\psi} \psi - \lambda \frac{\partial \bar{\psi}}{\partial z} \psi - (1 + \lambda) \bar{\psi} \frac{\partial \psi}{\partial z} \right\}$$

If now j, k and $j + k$ are all distinct from zero

$$j, k, j + k \neq 0 \quad [L_j, L_k] = (j - k)L_{j+k} \quad (115)$$

To obtain an expression of L_0 and of the central charge, we compute for $j \neq 0$ the following commutator

$$[L_j, L_{-j}] = \sum_k \{ \mu + k + j + \lambda(j - 1) \} \{ \mu + k - \lambda(j + 1) \} (a_k^\dagger a_k - a_{k+j}^\dagger a_{k+j})$$

We reach here a delicate point where care must be exercised. In such a singular expression, it is not advisable to perform a translation on indices. In a subspace of fixed charge Q , and acting on a vector with a finite energy, the operator $a_k^\dagger a_k - a_{k+j}^\dagger a_{k+j}$ will vanish for $|k|$ large enough. Therefore, as it stands, the previous expression is well-defined. Assuming j is positive, we observe that

$$a_k^\dagger a_k - a_{k+j}^\dagger a_{k+j} =: a_k^\dagger a_k : -: a_{k+j}^\dagger a_{k+j} : + \begin{cases} 0 & \text{if } k \geq 0 \\ 1 & \text{if } -j \leq k < 0 \\ 0 & \text{if } k < -j \end{cases}$$

Thus

$$[L_j, L_{-j}] = 2j \sum_k (\mu - \lambda + k) : a_k^\dagger a_k : + \sum_{k=0}^{j-1} \{ \mu + k + \lambda(j - 1) \} \{ \mu + k - j - \lambda(j + 1) \}$$

With

$$\sum_{k=0}^{j-1} 1 = j \quad \sum_{k=0}^{j-1} k = \frac{1}{2}j(j-1) \quad \sum_{k=0}^{j-1} k^2 = \frac{1}{6}j(j-1)(2j-1)$$

we define in the sector of charge $Q = n$

$$L_0 = H - \frac{1}{2}n(n-1) + h_n \quad h_n = \frac{1}{2}(n+\mu)(n+\mu-2\lambda-1) \quad (116)$$

such that

$$[L_j, L_{-j}] = 2jL_0 + \frac{1}{12}cj(j^2 - 1) \quad (117)$$

$$c = -2[6\lambda(\lambda + 1) + 1] \quad (118)$$

Obviously equation (117) also holds for negative j . Finally combining the definitions (111) and (116), we can check the

last commutation relations of the Virasoro algebra

$$[L_0, L_j] = -jL_j \quad (119)$$

This is a typical example of a representation, which we denote $\mathcal{F}_{\lambda, \mu, n}$ in the sector with charge n , exhibiting the central charge as a quantum anomaly. Here c is a function of λ only, satisfying

$$c(\lambda) = c(-1 - \lambda) \quad (120)$$

The adjoint of $L_j(\lambda, \mu)$ for real λ and μ is such that

$$L_{-j}^\dagger(\lambda, \mu) = L_j(-1 - \lambda, \mu - 1 - 2\lambda) \quad (121)$$

The correspondence $\lambda \rightarrow \lambda' = -1 - \lambda$, $\mu \rightarrow \mu' = \mu - 1 - 2\lambda$, which leaves both the central charge c and the c -number part of L_0 invariant, is recognized to be the same as occurring in the contragredient form (103b) for $p = 0$. Thus $\mathcal{F}_{\lambda, \mu}$ and $\mathcal{F}_{-1-\lambda, \mu-1-2\lambda}$ are contragredient.

Parenthetically we observe that when $\lambda = 1$, corresponding in the present notation to vector fields (or dually when $\lambda' = -1 - \lambda = -2$ to quadratic differentials) the value (118) of the central charge is $c = -26$. This is the significant value in the quantization of the bosonic string theory, giving the anomaly from the Faddeev-Popov ghost system required to compensate for reparametrization invariance. This anomaly has to be cancelled by the coordinate contribution associated to free fields, each of central charge unity, and leads to a requirement of 26 dimensions for a consistent model (chapter 11).

There was nothing special in our choice of the reference state $|\Omega\rangle$ and the corresponding definition of the charge zero states. The shift operator defines a canonical transformation

$$(a_i, a_i^\dagger) \rightarrow (a_{i-n}, a_{i-n}^\dagger) \quad (122)$$

$$|\Omega, n\rangle \rightarrow |\Omega, 0\rangle$$

which exhibits an isomorphism between the charge n and charge 0 sectors. Of course, its effect on Wick ordering is not trivial. Consequently, the representations $\mathcal{F}_{\lambda, \mu, n}$ and $\mathcal{F}_{\lambda, \mu+n, 0}$, denoted simply $\mathcal{F}_{\lambda, \mu+n}$, are isomorphic. Indeed μ was only appearing in the combination $\mu + n$ in (116), and the central charge is independent both from μ and n . We verify that

$$L_0 |\Omega, n\rangle = h_n |\Omega, n\rangle \quad (123)$$

$$L_j |\Omega, n\rangle = 0 \quad j > 0$$

Therefore $\mathcal{F}_{\lambda, \mu, n}$ admits $|\Omega, n\rangle$ as a vector corresponding to a dominant weight h_n .

As an example, the simplest case of this construction corresponds to $\lambda = \mu = -\frac{1}{2}$. In the charge zero sector, for any j

$$L_j = \frac{1}{2} \oint \frac{dz}{2i\pi} z^j [: \partial_z \bar{\psi} \psi : - : \bar{\psi} \partial_z \psi :] \quad (124)$$

$c = 1$

This sector involves two types of fermions with opposite charge and $c = 1$. One therefore suspects that there exists a "real" fermionic theory (Majorana fermions), with only one type of fermions, which will be studied in a later part of this chapter, corresponding to $c = \frac{1}{2}$. For instance, if $\mathcal{Z}$ denotes the set of integers, and if we define

$$\psi(z) = \sum_{k \in \mathcal{Z} + \frac{1}{2}} a_k z^{-k} \quad \{a_k, a_{k'}\} = \delta_{k+k', 0} \quad (125)$$

the field ψ is double valued, but bilinears are well-defined on the unit circle. We define

$$L_j = \frac{1}{4} \sum_{k \in \mathcal{Z} + \frac{1}{2}} (2k - j) : a_{j-k} a_k : = \frac{1}{4} \oint \frac{dz}{2i\pi} z^j (: \partial_z \psi \psi : - : \psi \partial_z \psi :)$$

$$[L_j, L_k] = (j - k) L_{j+k} + \frac{1}{24} j(j^2 - 1) \delta_{j+k, 0} I \quad (126)$$

In the expression above, the $::$ symbol requires the a_k 's, $k > 0$, to be ordered to the right of those with $k < 0$, keeping track of the permutation sign. The reference state is annihilated by the operators a_k , $k > 0$ and $L_0 = \sum_{k \geq 1/2} k a_{-k} a_k$ is a positive operator.

iii) We now set λ equal to the value

$$\lambda_0 = -\frac{(n-1)(n+2)}{2n} \quad (127)$$

with n an integer larger or equal to 2 in order to construct a function $\varphi_{n,p}$ as in (105) with $\mu = p/n$. The operator

$$\phi' = \sum_{\substack{q_0, \dots, q_{n-1} = -\infty \\ \sum q_i = -\frac{1}{2}n(n+1)+1-p}}^{+\infty} \prod_{0 \leq r < s \leq n-1} (q_r - q_s) a_{q_0}^\dagger \cdots a_{q_{n-1}}^\dagger \quad (128)$$

is such that

$$[Q, \phi'] = n\phi' \quad [H, \phi'] = (1 - p - \frac{1}{2}n(n+1)) \phi' \quad (129)$$

Hence ϕ'^k increases the charge by nk , and by construction satisfies

$$L_j \left(\lambda_0, \mu = \frac{p}{n} \right) \phi'^k |\Omega\rangle = 0 \quad j > 0 \quad (130)$$

which means that it generates, using (122), a singular vector in $\mathcal{F}_{\lambda_0, \mu=p/n, nk} \sim \mathcal{F}_{\lambda_0, \mu=p/n+nk}$. In $\mathcal{F}_{\lambda_0, p/n+nk}$ this state may be obtained by acting on $|\Omega, -nk\rangle$ with the operator ϕ^k (this was the reason for the prime above), where

$$\phi = \sum_{\substack{q_0, \dots, q_{n-1} = -\infty \\ \sum q_i = -\frac{1}{2}n(n+1)-n^2k+1-p}}^{+\infty} \prod_{0 \leq r < s \leq n-1} (q_r - q_s) a_{q_0}^\dagger \cdots a_{q_{n-1}}^\dagger \quad (131a)$$

satisfying

$$[Q, \phi^k] = nk\phi^k \quad [H, \phi^k] = k(1 - p - \frac{1}{2}n(n+1) - n^2k) \phi^k \quad (131b)$$

Set

$$\mu_0 = \frac{p}{n} + nk \quad (132)$$

Since $H|\Omega, -nk\rangle = \frac{1}{2}nk(nk+1)|\Omega, -nk\rangle$, the singular vector in the zero charge sector

$$|s\rangle = \phi^k |\Omega, -nk\rangle \quad (133)$$

obeys

$$\begin{aligned} Q|s\rangle &= 0 & L_0(\lambda_0, \mu_0)|s\rangle &= (h_0 + \ell k)|s\rangle \\ j > 0 & & L_j(\lambda_0, \mu_0)|s\rangle &= 0 \end{aligned} \quad (134)$$

with λ_0 and μ_0 given by (127) and (132), and the remaining symbols ℓ and h_0 stand for

$$\ell = 1 - p - \frac{1}{2}n^2(1+k) \quad h_0 = \frac{1}{2}\mu_0(\mu_0 - 2\lambda_0 - 1) \quad (135)$$

Finally the central charge c is given by (118), with λ_0 substituted for λ . Since

$$L_0(\lambda_0, \mu_0)|\Omega\rangle = h_0|\Omega\rangle \quad (136)$$

the construction of the singular vector makes sense (i.e. $|s\rangle \neq 0$) provided ℓk is positive, i.e.

$$p < 1 - \frac{1}{2}n^2(k+1) \quad (137)$$

The whole procedure involves three arbitrary integers n, p, k and is such that ℓk is always an integer. As a result ℓ is an integer except when n is odd and k even, in which case ℓ is a half integer. To avoid this possibility we assume n even.

iv) We return to the proof of the Kac formula. Consider an abstract Verma module with central charge $c(\lambda_0)$ and dominant weight h_0 , with corresponding vector $|h_0\rangle$. Let $\{L_{-j}\}$ denote an

ordered string of operators L_{-j} . To the vector $\{L_{-j}\}|h_0\rangle$ we can let correspond one of the two kets

$$\begin{aligned} \{L_{-j}\}|h_0\rangle &\xrightarrow{f} \{L_{-j}(\lambda_0, \mu_0)\}|\Omega\rangle \\ \{L_{-j}\}|h_0\rangle &\xrightarrow{g} \{L_{-j}(-1-\lambda_0, \mu_0-1-2\lambda_0)\}|\Omega\rangle \end{aligned} \quad (138)$$

The second ket is dual to the bra $\langle\Omega|\{L_j^\dagger(\lambda_0, \mu_0)\}$ according to equation (121), with h_0 invariant in the corresponding substitution.

Assume that the Verma module possesses a singular vector at a level smaller or equal to r . If the mapping f , which preserves the level sequence, is surjective, $\mathcal{F}_{\lambda_0, \mu_0}$ will also possess such a singular vector. If the application is not surjective, there must exist in the zero charge sector a ket $|\sigma\rangle$ with an excitation energy smaller or equal to r which is orthogonal to the image of f . Hence $\langle\sigma|\{L_{-j}(\lambda_0, \mu_0)\}|\Omega\rangle = 0$ for all strings of L_{-j} such that $\sum j = r$. If one takes complex conjugates, one concludes that $|\sigma\rangle$ is a singular vector for $\mathcal{F}_{-1-\lambda_0, \mu_0-1-2\lambda_0}$. Conversely, if the image under f or g admits a singular vector at a level $\leq r$, one readily sees that this must also be the case for the abstract module.

For λ and μ arbitrary, consider the quantity

$$\begin{aligned} b_{\alpha, \beta}(\lambda, \mu) &= b_{\beta, \alpha}(\lambda, \mu) = b_{-\alpha, -\beta}(-1-\lambda, \mu-1-2\lambda) \\ &= \{(\alpha+1)(\lambda+\frac{1}{2})-\mu\} \{(\beta+1)(\lambda+\frac{1}{2})-\mu\} \end{aligned} \quad (139)$$

The product $b_{\alpha, \beta}(\lambda, \mu)b_{-\alpha, -\beta}(\lambda, \mu)$, symmetric in α, β , is also invariant in the correspondence $\lambda \rightarrow -1-\lambda, \mu \rightarrow \mu-1-2\lambda$. Hence it is only a function of h (given by equation (116) with $n=0$), and c (given by (118)). We find

$$\begin{aligned} B_{\alpha, \beta}(c, h) &= B_{-\alpha, -\beta}(c, h) = \frac{1}{4}b_{\alpha, \beta}(\lambda, \mu)b_{-\alpha, -\beta}(\lambda, \mu) \\ &= \left[h + \frac{1}{24}(\alpha^2 - 1)(c - 13) + \frac{1}{2}(\alpha\beta - 1)\right] \\ &\quad \times \left[h + \frac{1}{24}(\beta^2 - 1)(c - 13) + \frac{1}{2}(\alpha\beta - 1)\right] + \frac{1}{16}(\alpha^2 - \beta^2)^2 \end{aligned} \quad (140)$$

For $\lambda = \lambda_0 = -(n-1)(n+2)/2n$, $\mu = \mu_0 = nk + p/n$, n even, hence $\ell = 1 - p - n^2(k+1)/2$ integer, we just saw that the Kac determinant of degree $k\ell$ has to vanish. One verifies that

$$b_{-k, -\ell}(\lambda_0, \mu_0) = 0 \quad (141)$$

As a consequence, the determinant of level $k\ell$ has to vanish on an infinite number of points of the curve $B_{k, \ell}(c, h) = 0$. The latter is irreducible if $k \neq \ell$. Hence $\det_{k\ell}(c, h)$ is divisible by $B_{k, \ell}(c, k)$ or

its root when $k = \ell$, and $B_{k, k}$ is a square. As a result, $\det_{k\ell}^2(c, h)$ admits $B_{k, \ell}(c, h)$ as a factor. Since a singular vector at level r generates a zero of order $p(s-r)$ at level $s \geq r$, the quantity

$$\prod_{\substack{k, \ell \\ k\ell \leq r}} B_{k, \ell}(c, h)^{p(r-k\ell)}$$

divides $\det_r(c, h)^2$. One checks that both expressions have the same degree in h . As a consequence, up to an h -independent factor,

$$[\det_r(c, h)]^2 = \text{cst} \prod_{\substack{1 \leq k, \ell \\ k\ell \leq r}} B_{k, \ell}(c, h)^{p(r-k\ell)} \quad (142)$$

Finally $B_{k, \ell}$ is symmetric in k and ℓ , hence occurs twice in the product (142) for $k \neq \ell$, while $B_{k, k}$ is a square

$$B_{k, k}(c, h) = \left[h - \frac{(1-c)}{24}(k^2 - 1)\right]^2$$

Furthermore, when we use the parametrization $c = 1 - 6/m(m+1)$,

$$B_{k, \ell} = (h - h_{k, \ell})(h - h_{\ell, k}) \quad (143)$$

with $h_{k, \ell}$ given by equation (97). This allows the square root in (142) to be taken and produces Kac's formula (96).

9.1.7 Unitary and minimal representations

Let us complete the discussion on the conditions on c and h under which a representation is unitary, i.e.

$$L_j^\dagger = L_{-j} \quad (144)$$

For the module (or vector space) generated by a highest weight h to be a Hilbert space with its scalar product giving rise to the contragredient form (79a), it is necessary that (i) either the latter is positive definite (ii) or else if there exists singular vectors, that the invariant subspaces generated by the latter be factored out, leaving a quotient space in which the form is again positive. These conditions are also sufficient.

We have already seen that a necessary condition for unitarity is

$$c \geq 0, \quad h \geq 0$$

In the range $0 < c < 1$, Friedan, Qiu and Shenker have succeeded in showing that the only allowed values required by unitarity are

such that (i) c has the form (89) with m integer, $m \geq 2$ ($m = 2$ corresponds to $c = 0$ and the trivial identity representation) (ii) the weights h are given by the zeroes of the Kac determinant with the two integers r and s ranging over a finite set of values, $1 \leq s \leq r \leq m - 1$. In summary

$$c = 1 - 6/m(m+1) \quad m \text{ integer } \geq 2$$

$$h \equiv h_{r,s} = \frac{[r(m+1) - sm]^2 - 1}{4m(m+1)} \quad 1 \leq s \leq r \leq m - 1 \quad (145)$$

If we use only the restriction $1 \leq s \leq m, 1 \leq r \leq m - 1$, each weight $h_{r,s}$ occurs twice, since $h_{r,s} \equiv h_{m-r, m+1-s}$.

Truly interesting unitary representations of this sort are therefore characterized by an integer $m \geq 3$ and a finite set of possible weights, hence a finite set of possible scaling dimensions for the fundamental or primary observables. In this sense, we call these representations minimal. The corresponding Verma modules possess singular vectors, i.e. invariant subspaces. For the same value of c , it is also possible to consider fields belonging to representations with a weight $h_{r,s}$ with r and/or s outside the above range. Those might enter into the construction of other types of models which will not principally be dealt with here. Nevertheless, it is useful to introduce a table of these inclusions of invariant subspaces, because they enable one to understand the counting of the remaining linearly independent states at a given level, once the invariant subspaces are factored out. Fix a value of c as in (145). Feigin and Fuchs have established the set of inclusions between Verma modules, depicted on figure 3, arising from the Kac formula. On the picture, an arrow points from the larger space to the subspace. The corresponding weights differ by products of integers. Moreover five subspaces such as M, N', N'', P', P'' are such that M contains the sum (not direct sum) $N' \oplus N''$, while the intersection $N' \cap N''$ is equal to the sum $P' \oplus P''$. To obtain the irreducible representation at the top, with r and s in the restricted range given in equation (145), one has to factor out the sum of its first two "descendents". We shall see in the next section applications of these sets of inclusions.

In the range 1 to 25, there is not constraint on the values of c due to unitarity. At $c = 1$, there exists again a (simpler) form

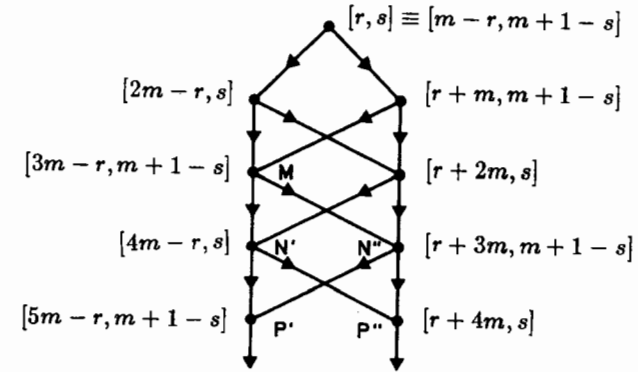


Fig. 3 Inclusions of Verma modules $[r', s']$ corresponding to a central charge $c = 1 - 6/m(m+1)$, $m \geq 3$, and a succession of weights $h_{r',s'} = [(r'(m+1) - s'm)^2 - 1]/4m(m+1)$.

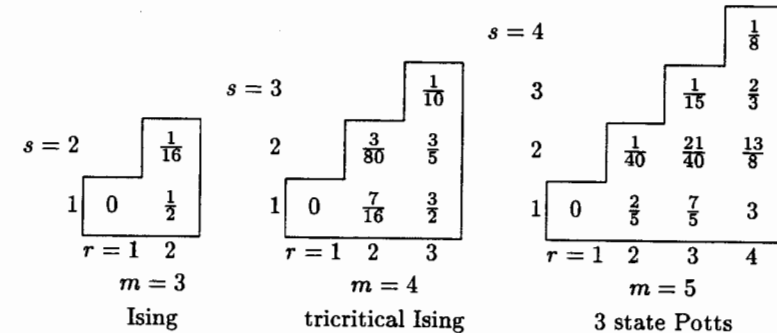


Fig. 4 Grid of conformal weights for the first three instances of unitary minimal representations.

of reducibility (and unitarity) whenever $h = \frac{1}{4}n^2$ with n integer.

As we do not claim here to give an exhaustive presentation, we shall limit ourselves to values of c in the range $c \leq 1$. Such cases are in essence the simplest instances. They include some well-known examples, with at most discrete symmetries for $c < 1$.

For the minimal unitary series, we illustrate the above considerations by displaying in figure 4 the first three examples of the grid

of conformal weights. They are already labelled by some models, with qualifications which will be justified later on. It is fascinating to find that conformal invariance alone predicts the existence of systems with a few rational exponents. What is to be found is the key to the correspondence in some specific instance, as well as the physical meaning of the various observables.

Minimal representations need not be unitary, as shown by the simplest example below. Of course one may admit of certain pathologies, like growing correlations, to be properly interpreted. These minimal cases are obtained if we substitute in the formula $c = 1 - 6/m(m+1)$ not an integer m but a rational fraction. For instance, if $p > p'$ are two positive integers with no common factors (coprimes) such that

$$m = \frac{p'}{p-p'} \quad m+1 = \frac{p}{p-p'} \quad (146a)$$

the corresponding value of c (no longer restricted to be positive) is given by

$$c = 1 - \frac{6(p-p')^2}{pp'} \quad (146b)$$

A conformal grid of minimal representations and their weights, is given by

$$h_{r,s} = h_{p'-r, p-s} = \frac{(rp - sp')^2 - (p-p')^2}{4pp'} \quad (147)$$

$$1 \leq r \leq p' - 1 \quad 1 \leq s \leq p - 1$$

Since for p, p' as above we can find r_0 and s_0 in the indicated range such that $(r_0 p - s_0 p')^2 = 1$, it is clear that, for $p - p' \geq 2$, there will exist negative weights in the table, and therefore the corresponding scalar field $A_{h_0, \bar{h}_0}$ will have correlations increasing with distance. This was excluded in the unitary minimal case. The consistency of these minimal models in the sense that at most a finite set of conformal weights appear for the primary operators will be studied in section 3.

To show that this is not a totally academic possibility, let us present the following example due to Cardy. Let us look at the most elementary case of the above family with a unique nontrivial scalar field, apart for the unit operator (and its descendents like the energy-momentum tensor). Concentrating on the V algebra, we

look for a grid with two entries (taking into account the symmetry indicated in (147)). With duplication this means four entries in $1 \leq r \leq p' - 1$ $1 \leq s \leq p - 1$ with $p > p'$ and coprimes. The only possibility is a unique column ($r = 1, p' = 2$) of four boxes ($p = 5, 1 \leq s \leq 4$). Consequently we have two weights, and the values for h and c are

$$h = 0, -\frac{1}{5} \quad c = -\frac{22}{5} \quad (148)$$

The unique nontrivial scalar field $A_{-\frac{1}{5}, -\frac{1}{5}}$ has a scaling dimension

$$\Delta_A = -\frac{2}{5} \quad (149)$$

Now let us try to guess which scalar interacting field theory (at criticality) this could possibly correspond to. The only candidate seems to be the $i\varphi^3$ theory with imaginary cubic coupling, and coincides with the effective continuous model for the Lee-Yang edge singularity. In this instance no symmetry is broken at criticality (but the correlation length diverges) and the model has no chance to be unitary. Let, as usual, η denote the critical exponent for the two-point correlation. It follows in this two-dimensional case that $\eta = 2\Delta = -\frac{4}{5}$, indeed negative (hence correlations grow!). Furthermore we recall that we defined the exponent σ such that the singular part of the magnetization behaves as

$$m_s \sim h^\sigma \quad (150)$$

where h denotes the deviation from the critical magnetic (external) field. Scaling relations imply that

$$\sigma = \frac{d-2+\eta}{d+2-\eta} \xrightarrow{d=2} \frac{\eta}{4-\eta} = -\frac{1}{6} \quad (151)$$

in excellent agreement with numerical simulations. There is a relation between this singular growing part of the magnetization and the fact that after subtraction we are left at criticality with a connected correlation growing at large distance.

9.1.8 Characters of the Virasoro algebra

For further applications it will be useful to know the so called characters of the representations of the Virasoro algebra. In a representation characterized by c and a highest weight h , these are generating functions for the number of linearly independent states at level n , hence corresponding to an eigenvalue $h+n$ for L_0 . For a fixed $q = \exp 2i\pi\tau$ smaller than one in modulus ($\text{Im}\tau > 0$)

the character appears as the trace in the given representation of the operator q^{L_0} . In other words with $\dim(h+n)$ the dimension of the space at level n we define

$$\chi_{(c,h)}(\tau) = \text{Tr } q^{L_0 - c/24} = \sum_{n=0}^{\infty} \dim(h+n) q^{n+h-c/24} \quad q = \exp 2i\pi\tau \quad (152)$$

The extra factor $q^{-c/24}$ is included in this definition for later convenience. For general values c and h , when the highest weight representation is such that the corresponding Verma module has no singular vector, it is clear that

$$\chi_{(c,h)}(\tau) = \frac{q^{h-c/24}}{P(q)} \quad (153)$$

The factorized form of $P(q)$ shows that the series defining χ converges for any $|q| < 1$. Moreover from equation (83), it is not difficult to show that the leading term in the Hardy-Ramanujan asymptotic expansion for $p(n)$, the number of partitions of n is

$$p(n) \underset{n \rightarrow \infty}{\sim} \frac{\exp \pi \sqrt{2n/3}}{4\sqrt{3}n} \left[1 + O\left(\frac{\ln n}{n^{1/4}}\right) \right] \quad (154)$$

The situation is more interesting for the degenerate representations, when the Verma module contains invariant subspaces. From such diagrams as appear in figure 3, starting from an expression such as given in (153) one must subtract similar terms. Furthermore at any stage such as shown on this figure, the sum of invariant subspaces is not a direct one. As a result starting from r and s in the conformal grid the character formula for minimal representations obtained by Rocha-Caridi reads, with $\chi_{(c,h_{rs})} \equiv \chi_{r,s}$

$$\chi_{r,s}(\tau) = q^{-c/24} \frac{\sum_{k=-\infty}^{+\infty} q^{[(2kpp' + rp - sp')^2 - (p-p')^2]/4pp'} - (s \rightarrow -s)}{P(q)} \quad (155)$$

Here of course, as in (146), p and p' are coprime integers, $p > p'$, r ranges between 1 and $p' - 1$, s between 1 and $p - 1$, with some further restriction such as $rp - sp' > 0$ to avoid double counting. Also $c = 1 - 6(p - p')^2/pp'$. Minimal unitary models are obtained by setting $p' = m$, $p = m + 1$ with $m \geq 2$, integer.

- i) In the minimal unitary case with $m = 2$, the irreducible representation is a trivial one-dimensional one. Setting $\chi \equiv 1$ in (155) leads to Euler's (pentagonal) identity

$$\prod_1^{\infty} (1 - q^n) = \sum_{n=-\infty}^{+\infty} (-1)^n q^{\frac{1}{2}n(3n+1)} \quad (156)$$

- ii) For future use it is possible to present the expression (155) for the characters in a more compact form by introducing the following notation. Set $N = 2pp'$, an even integer. Instead of parametrizing the character by the couple r, s corresponding to $h_{r,s}$ in (147), use an integer modulo N , $\lambda \equiv rp - sp' \pmod{N}$. Observe that if $\omega_0 \equiv r_0p + s_0p' \pmod{N}$, with $r_0p - s_0p' = 1$ (according to the fact that p and p' are coprimes) we have $\omega_0^2 = 1 \pmod{2N}$ and $\omega_0\lambda = rp + sp' \pmod{N}$. Finally trade $P(q)$ for Dedekind's function

$$\eta(\tau) = q^{1/24} P(q) = q^{1/24} \prod_1^{\infty} (1 - q^n) \quad (157)$$

where we continue as above to use the notation $q = \exp 2i\pi\tau$. Then formula (155) can be rewritten as

$$\chi_{\lambda}(\tau) = \frac{\sum_{k=-\infty}^{+\infty} q^{(kN+\lambda)^2/2N} - q^{(kN+\omega_0\lambda)^2/2N}}{\eta(\tau)} \quad (158a)$$

showing that

$$\chi_{\lambda}(\tau) = \chi_{\lambda+N}(\tau) = \chi_{-\lambda}(\tau) = -\chi_{\omega_0\lambda}(\tau) \quad (158b)$$

These characters can be expressed in terms of elliptic functions. In appendix A, we recall some properties of elliptic θ -functions useful in the present context.

Many more nontrivial character formulas do exist, corresponding for instance to some of the irreducible invariant subspaces included in the reducible Verma modules. Also for $c = 1$ and $h = \frac{1}{4}\alpha^2$ (α integer) we have instead a linear chain of inclusions. Thus

$$\chi_{\alpha}(\tau) = \frac{q^{\alpha^2/4} - q^{(\alpha+2)^2/4}}{\eta(\tau)} \quad \alpha \text{ integer.} \quad (159)$$

9.2 Examples

Let us now give some typical and elementary examples of conformal fields. We shall resume the general discussion in section 3.

9.2.1 Gaussian model

The simplest example, and the one which in a some sense is the building block of many of the explicit realizations, is the Gaussian, or massless, neutral free field model. The corresponding action reads, with $\varphi(\mathbf{x})$ real

$$S = \int \frac{d^2x}{2\pi} \frac{1}{2} \sum_{\mu=1,2} (\partial_\mu \varphi)^2 = \int \frac{d^2x}{2\pi} 2(\partial_z \varphi)(\partial_{\bar{z}} \varphi) \quad (160)$$

The normalization agrees with our definition of the energy-momentum tensor. With a sign fixed by equation (57), the latter takes the familiar form

$$T_{\mu\nu} = -\partial_\mu \varphi \partial_\nu \varphi + \frac{1}{2} \delta_{\mu\nu} (\partial\varphi)^2 \quad (161)$$

$$T \equiv T_{zz} = -(\partial_z \varphi)^2 \quad \bar{T} \equiv T_{\bar{z}\bar{z}} = -(\partial_{\bar{z}} \varphi)^2 \quad T_{z\bar{z}} = 0$$

The two-point function for φ involves an arbitrary scale R (infrared cutoff)

$$\langle \varphi(z_1, \bar{z}_1) \varphi(z_2, \bar{z}_2) \rangle = \frac{1}{2} \ln \frac{R^2}{z_{12} \bar{z}_{12}} \quad (162)$$

In the expression (161), a Wick ordering prescription has to be understood to give a definite meaning to correlation functions of T , as we discuss below. The logarithmic behaviour of the φ -correlations shows that it cannot be taken as a bona fide primary field (the tentative assignment of conformal weights is (0,0)). Rather, derivatives or exponentials of φ are candidates for observables. The corresponding correlations will have a more decent scaling behaviour, and should be such that the arbitrary scale R cancels out, as was discussed in chapter 4 of volume 1.

To obtain the central charge, we compute the $\langle TT \rangle$ correlation, omitting tadpole contributions i.e. self-contractions of φ 's at the same point. This results in the expected analytic expression

$$\langle T(z_1) T(z_2) \rangle = 2 \left(\partial_{z_1} \partial_{z_2} \langle \varphi(z_1, \bar{z}_1) \varphi(z_2, \bar{z}_2) \rangle \right)^2 = \frac{1}{2z_{12}^4} \quad (163)$$

with the complex conjugate for $\langle \bar{T} \bar{T} \rangle$. Consequently the central charge is unity

$$c = 1 \quad (164)$$

This was the rationale behind the normalization of the two-point function leading to the definition of the central charge.

Free fields allow us to understand how the anomaly computed for an infinitesimal conformal transformation gets promoted to the Schwarzian derivative for a finite one. The universal structure of the transformation law (50) implies that the calculation performed for $c = 1$ can be generalized by inserting any value for the central charge. The Wick prescription may be understood to stand for a short-distance subtraction procedure, after point splitting the arguments of T

$$T \rightarrow \lim_{z_1, z_2 \rightarrow z} -\partial_{z_1} \varphi \partial_{z_2} \varphi + \partial_{z_1} \partial_{z_2} \left(\frac{1}{2} \ln \frac{R}{z_{12}} \right)$$

to be inserted in correlation functions. In the above formula, only the leading short distance singularity of the correlation function does matter. Now, pretend that under a conformal map $\varphi(z, \bar{z}) = \tilde{\varphi}(u, \bar{u})$. In the new coordinate system with $u = f(z)$

$$T = \lim_{z_1, z_2 \rightarrow z} f'(z_1) f'(z_2) \left[-\partial_{u_1} \tilde{\varphi} \partial_{u_2} \tilde{\varphi} + \partial_{u_1} \partial_{u_2} \frac{1}{2} \ln \frac{\tilde{R}}{u_{12}} \right] + \frac{1}{2} \partial_{z_1} \partial_{z_2} \ln \frac{u_{12}}{z_{12}}$$

We have added and subtracted a term in such a way that the first part is recognized as $\tilde{T}(u) (du/dz)^2$ in the coinciding point limit, while the second one yields the anomaly. Set

$$z_2 = z, \quad z_1 = z + \delta, \quad \frac{u_{12}}{z_{12}} = f' + \frac{1}{2} \delta f'' + \frac{1}{6} \delta^2 f''' + \dots$$

$$\ln \frac{u_{12}}{z_{12}} = \ln f' + \frac{1}{2} \delta \frac{f''}{f'} + \delta^2 \left(\frac{1}{6} \frac{f'''}{f'} - \frac{1}{8} \left(\frac{f''}{f'} \right)^2 \right) + \dots$$

and the anomaly is

$$\frac{1}{2} \left(\frac{\partial^2}{\partial z \partial \delta} - \frac{\partial^2}{\partial \delta^2} \right) \left[\ln f' + \frac{1}{2} \delta \frac{f''}{f'} + \delta^2 \left(\frac{1}{6} \frac{f'''}{f'} - \frac{1}{8} \left(\frac{f''}{f'} \right)^2 \right) + \dots \right] = \frac{1}{12} \left(\frac{f'''}{f'} - \frac{3}{2} \left(\frac{f''}{f'} \right)^2 \right)$$

in agreement with equation (50) for $c = 1$.

an $O(n)$ invariant measure. Let $\gamma, \delta \dots$ denote the indices of the field running from 1 to n , while i, j or coordinates $\mathbf{x}$ label the sites. Let us repeat an argument given previously, in slightly generalized form. The two point function is

$$\langle \varphi_0^\gamma \varphi_{\mathbf{x}}^\delta \rangle = Z^{-1} \int \prod_k d^n \varphi_k \delta(\varphi_k^2 - 1) \varphi_0^\gamma \varphi_{\mathbf{x}}^\delta \exp \left\{ n\beta \sum_{(ij)} \varphi_i \cdot \varphi_j \right\} \quad (113)$$

In any finite volume, the measure is rotationally invariant (in internal space), under a global transformation. We now perform an infinitesimal change of variables in the integral, affecting only φ_i , in the form of an infinitesimal rotation

$$\varphi_i \rightarrow \varphi_i + \varepsilon(A\varphi)_i \quad (114)$$

with A an arbitrary antisymmetric matrix. This should not affect the integral. Hence the coefficient of ε is zero, leading to

$$0 = \delta_{i,0} \langle (A\varphi)_0^\gamma \varphi_{\mathbf{x}}^\delta \rangle + \delta_{i,\mathbf{x}} \langle \varphi_0^\gamma (A\varphi)_{\mathbf{x}}^\delta \rangle + n\beta \sum_{j(i)} \langle \varphi_0^\gamma \varphi_{\mathbf{x}}^\delta (\varphi_j \cdot (A\varphi)_i) \rangle \quad (115)$$

Since the antisymmetric matrix A is arbitrary, we identify the coefficient of $\frac{1}{2}(A^{\alpha\beta} - A^{\beta\alpha})$, and take the trace over indices, with the result that

$$(1 - n^{-1})(\delta_{i,\mathbf{x}} - \delta_{i,0}) \langle \varphi_0 \cdot \varphi_{\mathbf{x}} \rangle = \beta \sum_{j(i)} \left\langle \left[(\varphi_0 \cdot \varphi_j) (\varphi_{\mathbf{x}} \cdot \varphi_i) - (\varphi_0 \cdot \varphi_i) (\varphi_{\mathbf{x}} \cdot \varphi_j) \right] \right\rangle \quad (116)$$

The sum over $j(i)$ runs over neighbours of the point i . In equation (116), we have already assumed an infinite volume limit, but as such it is an exact (Ward) identity. To make use of it as $n \rightarrow \infty$, we imagine evaluating all integrals by a saddle point method. Saddle points are in general degenerate owing to the $O(n)$ invariance, but rotational invariants are insensitive to this degeneracy, and their mean values are then expected to factorize. As a consequence, we may perform the replacement

$$\lim_{n \rightarrow \infty} \langle (\varphi_a \cdot \varphi_b) (\varphi_c \cdot \varphi_d) \rangle \rightarrow \lim_{n \rightarrow \infty} \langle \varphi_a \cdot \varphi_b \rangle \lim_{n \rightarrow \infty} \langle \varphi_c \cdot \varphi_d \rangle \quad (117)$$

In this limit we obtain, up to corrections of order n^{-1} , a quadratic equation for the two point function,

$$\begin{aligned} & (\delta_{i,\mathbf{x}} - \delta_{i,0}) \langle \varphi_0 \cdot \varphi_{\mathbf{x}} \rangle \\ &= \beta \sum_{j(i)} \left\{ \langle \varphi_0 \cdot \varphi_j \rangle \langle \varphi_{\mathbf{x}} \cdot \varphi_i \rangle - \langle \varphi_0 \cdot \varphi_i \rangle \langle \varphi_{\mathbf{x}} \cdot \varphi_j \rangle \right\} \quad (118) \end{aligned}$$

Using a Fourier transform, with

$$\langle \varphi_{\mathbf{x}} \cdot \varphi_{\mathbf{y}} \rangle = \int \frac{d^d q}{(2\pi)^d} e^{i\mathbf{q} \cdot (\mathbf{x} - \mathbf{y})} G(\mathbf{q}) \quad (119)$$

the equation reads

$$G(\mathbf{q}) - G(\mathbf{k}) = 2\beta \sum_1^d (\cos q_\mu - \cos k_\mu) G(\mathbf{q}) G(\mathbf{k}) \quad (120)$$

For $\beta < \beta_c$, $\langle \varphi_0 \cdot \varphi_{\mathbf{x}} \rangle \rightarrow 0$ for $|\mathbf{x}| \rightarrow \infty$, hence $G(\mathbf{q})$ is regular at $\mathbf{q} = 0$, and we conclude that

$$G^{-1}(\mathbf{q}) = \beta \left\{ \xi^{-2} + 2 \sum_1^d (1 - \cos q_\mu) \right\} \quad (121)$$

is the required solution to (120), with ξ^{-2} a constant determined by $\langle \varphi_0^2 \rangle = 1$, which leads back to the saddle point equation (103). However, for $\beta > \beta_c$, a spontaneous magnetization appears, with $\lim_{|\mathbf{x}| \rightarrow \infty} \langle \varphi_0 \cdot \varphi_{\mathbf{x}} \rangle = M^2$, or equivalently a singular part in $G(\mathbf{q})$ of the form $M^2 (2\pi)^d \delta(\mathbf{q})$, where the δ function is assumed periodic on the torus $q_\mu = q'_\mu + 2\pi n_\mu$, n_μ integers. Subtracting M^2 from $\langle \varphi_0 \cdot \varphi_{\mathbf{x}} \rangle$ leads to the representation valid for $d > 2$

$$\beta > \beta_c, d > 2 \quad \langle \varphi_0 \cdot \varphi_{\mathbf{x}} \rangle = M^2 + \frac{1}{\beta} \int \frac{d^d q}{(2\pi)^d} \frac{e^{-i\mathbf{q} \cdot \mathbf{x}}}{2 \sum_1^d (1 - \cos q_\mu)} \quad (122)$$

Now M^2 is determined by the condition $\langle \varphi_0^2 \rangle = 1$, i.e. using equation (102)

$$1 - M^2 = \frac{\beta_c}{\beta} \quad (123)$$

Therefore for $d > 2$, the singularity of M is $[1 - (T/T_c)]^{1/2}$, and the exponent β remains unchanged, $\beta = 1/2$. At criticality

$$\beta = \beta_c \quad \langle \varphi_0 \cdot \varphi_x \rangle = \frac{1}{\beta_c} \int \frac{d^d q}{(2\pi)^d} \frac{e^{-iq \cdot x}}{2 \sum_1^d (1 - \cos q_\mu)} \quad (124)$$

$$\sim_{|x| \rightarrow \infty} \frac{\Gamma(d/2)}{\beta_c 2\pi^{d/2} (d-2)} \frac{1}{|x|^{d-2}}$$

leading to $\eta = 0$ for $d > 2$. In the symmetric phase $\beta < \beta_c$, the susceptibility is isotropic in internal space, and reads

$$\chi = \sum_x \langle \varphi_0 \cdot \varphi_x \rangle = G(\mathbf{q} = 0) = \xi^2 / \beta \quad (125)$$

It diverges as $(\beta_c - \beta)^{2\nu}$ and therefore

$$\gamma = 2\nu = 2/(d-2) \quad (126)$$

The analogous result in dimension two is that χ diverges as $\exp 4\pi\beta$ as $\beta \rightarrow \infty$.

Higher correlations functions in the large n limit easily follow from the factorization property. They vanish for an odd number of fields, while for instance in the symmetric phase

$$\langle \varphi_{x_1}^{\alpha_1} \varphi_{x_2}^{\alpha_2} \varphi_{x_3}^{\alpha_3} \varphi_{x_4}^{\alpha_4} \rangle = \frac{1}{n^2} \sum_{\text{pairings}} \delta^{\alpha_1 \alpha_2} \delta^{\alpha_3 \alpha_4} \langle \varphi_{x_1} \cdot \varphi_{x_2} \rangle \langle \varphi_{x_3} \cdot \varphi_{x_4} \rangle \quad (127)$$

where the sums includes three such terms.

3.3.3 Coupling to an external field

Let us introduce a coupling to an external field $n \sum_i \mathbf{h}_i \cdot \varphi_i$. Applying the saddle point method leads to a generalization of equation (101) involving a function ξ_i^{-2} such that

$$Z = Z_0 \left[\frac{1}{\beta} \left(\frac{2\pi}{n\beta} \right)^{\frac{1}{2}(n-2)} \right]^N \exp \frac{n}{2} \left\{ \beta \sum_i (\xi_i^{-2} + 2d) + \frac{1}{\beta} \sum_{i,j} \mathbf{h}_i \cdot \mathbf{h}_j \left(\frac{1}{\xi_i^{-2} - \Delta} \right)_{ij} - \text{Tr} \ln(\xi^{-2} - \Delta) \right\} \quad (128a)$$

with the stationarity condition

$$\beta = \left(\frac{1}{\xi^{-2} - \Delta} \right)_{ii} + \frac{1}{\beta} \sum_{j,k} \mathbf{h}_j \cdot \mathbf{h}_k \left(\frac{1}{\xi^{-2} - \Delta} \right)_{ji} \left(\frac{1}{\xi^{-2} - \Delta} \right)_{ik} \quad (128b)$$

If $\mathbf{h}$ is uniform, so is ξ . Summing over sites in (128b) yields

$$\beta = \int \frac{d^d p}{(2\pi)^d} \frac{1}{\xi^{-2} + 2 \sum_1^d (1 - \cos p_\mu)} + \frac{1}{\beta} \xi^4 \mathbf{h}^2 \quad (129a)$$

while, up to a β -independent constant F_0 , the free energy per site is

$$F = F_0 + \frac{n}{2} \left\{ -\ln \beta + \beta(\xi^{-2} + 2d) + \frac{\xi^2}{\beta} \mathbf{h}^2 - \int \frac{d^d p}{(2\pi)^d} \ln \left[\xi^{-2} + 2 \sum_1^d (1 - \cos p_\mu) \right] \right\} \quad (129b)$$

For any positive β , equation (129a) has a solution with ξ finite. In other words, no transition occurs in the presence of an external field. We can rewrite the relation between β , ξ and $\mathbf{h}$ as

$$\beta^2 - \beta \xi^2 a(\xi^2) - \mathbf{h}^2 \xi^4 = 0$$

$$a(\xi^2) = \int \frac{d^d p}{(2\pi)^d} \frac{1}{1 + \xi^2 2 \sum_1^d (1 - \cos p_\mu)} \quad (129c)$$

with $a(\xi^2)$ positive, ranging from 1 to 0 as ξ^2 varies from zero to infinity ($\xi^2 a(\xi^2)$ goes to a finite limit). Hence β behaves as ξ^2 for small ξ and as $|\mathbf{h}| \xi^2$ for large ξ . The magnetization per site is given by

$$\mathbf{M} = \frac{1}{nN} \frac{\partial}{\partial \mathbf{h}} \ln Z = \frac{\xi^2}{\beta} \mathbf{h} \quad (130)$$

At β_c , we find from equation (129)

$$\xi^6 \mathbf{h}^2 = \beta_c \int \frac{d^d p}{(2\pi)^d} \frac{1}{2 \sum_1^d (1 - \cos p_\mu) [\xi^{-2} + 2 \sum_1^d (1 - \cos p_\mu)]} \quad (131)$$

which shows that, if $\mathbf{h} \rightarrow 0$, ξ diverges. More precisely if $d > 4$, the r.h.s. has a finite limit as $\xi \rightarrow \infty$, hence

$$d > 4 \quad \beta = \beta_c \quad M \sim h^{1/3} \quad \delta = 3 \quad (132)$$

a classical mean field value. If $d = 4$, $\xi^6 h^2 \sim \text{cst} \ln \xi$ and

$$d = 4 \quad \beta = \beta_c \quad M \sim \left(h \ln \frac{1}{h} \right)^{1/3} \quad \delta = 3 \quad (133)$$

Finally for $d < 4$, one finds

$$d < 4 \quad \beta = \beta_c \quad M \sim h^{(d-2)/(d+2)} \quad \delta = \frac{d+2}{d-2} \quad (134)$$

The spherical model therefore offers a simple laboratory to study departures from mean field theory due to fluctuations, between the two critical dimensions two and four. At the marginal upper critical dimension, the appearance of logarithms signals that new phenomena are taking place, and are naturally related to similar properties of Brownian motion. In two dimensions, the correlation length diverges exponentially as one approaches zero temperature, which is a point where a consistent continuum field theory can be reached (the nonlinear σ -model). The investigation of higher order corrections (in $1/n$) is a more tricky matter which relies on techniques similar to those developed in chapter 5.

(i) Obtain the behaviour of the free energy as a function of $\theta = \beta_c - \beta$, and h^2 in the vicinity of the critical point. Writing m^2 for ξ^{-2} one has from (129b)

$$F = F_0 + \frac{n}{2} \left\{ \ln \frac{1}{\beta} + 2d\beta - \int \frac{d^d p}{(2\pi)^d} \ln \left[2 \sum_1^d (1 - \cos p_\mu) \right] + f \right\}$$

$$f = -\theta m^2 + \frac{h^2}{\beta m^2} \quad (135a)$$

$$- \int \frac{d^d p}{(2\pi)^d} \left\{ \ln \left[\frac{m^2 + 2 \sum_1^d (1 - \cos p_\mu)}{2 \sum_1^d (1 - \cos p_\mu)} \right] - \frac{m^2}{2 \sum_1^d (1 - \cos p_\mu)} \right\}$$

with a small m^2 behaviour

$$f \underset{m^2 \rightarrow 0}{\sim} \theta m^2 + \frac{h^2}{\beta m^2} + \begin{cases} Km^4 & d > 4 \\ Km^4 \ln \frac{1}{m} & d = 4 \\ Km^d & 2 < d < 4 \end{cases} \quad (135b)$$

Here K is a (d -dependent) positive constant, and equation (129a) expresses that F is stationary with respect to m^2 , at fixed h and θ . Rederive the critical behaviour for small h , θ and m^2 .

(ii) One can apply this same $n \rightarrow \infty$ limiting technique to the study of the Lee-Yang edge singularity in a complex external field. Set $h = hu$ where u is a fixed real unit vector. A singularity occurs when the solution $\xi(\beta, h)$ of (129a) becomes multivalued, i.e. when

$$\int \frac{d^d p}{(2\pi)^d} \frac{1}{[1 + \xi^2 \sum_1^d (1 - \cos p_\mu)]^2} + \frac{2h^2 \xi^2}{\beta} = 0 \quad (136)$$

One indeed finds a pair of complex conjugate singularities at $\pm ih_c$. The singular part of the free energy is $(h \mp ih_c)^{3/2}$ in all dimensions greater than two. To leading order the critical exponent $\sigma = 1/2$, as in mean field, while for h_c one finds

$$d > 4 \quad h_c \sim (\beta_c - \beta)^{3/2}$$

$$d = 4 \quad h_c \sim \frac{(\beta_c - \beta)^{3/2}}{\ln [1/(\beta_c - \beta)]} \quad (137)$$

$$d < 4 \quad h_c \sim (\beta_c - \beta)^{\frac{1}{2}(d+2)/(d-2)}$$

This seems at first sight to contradict the idea that $d = 6$ is an upper critical dimension for the Lee-Yang edge singularity. However, a closer look at $1/n$ corrections reinstates $d = 6$ as a marginal dimension, where one observes departures from the value $\sigma = 1/2$. In essence, the Lee-Yang singularity has no particular sensitivity to the number of components as one (internal space) direction is singled out by the external field. Consequently, the limit $n \rightarrow \infty$ is not particularly suited to reveal the role of fluctuations.

3.4 Corrections to mean field

The mean field approximation has shown the existence of critical points where qualitative as well as quantitative properties are modified. We have encountered continuous transitions characterized by a divergence of the correlation length, where the free energy as well as its derivative, the internal energy (hence the entropy), are continuous. The transition is accompanied by a change of symmetry properties of pure phases. The exact solution of the two-dimensional Ising model, or the spherical model, have indicated that mean field predictions are not reliable in low dimensions, where fluctuations have to be taken into account. The spherical model confirms that, for the family of models under consideration, four is an upper critical dimension. This is the same

dimension where the corresponding continuous φ^4 theory is (ultraviolet) renormalizable. These phenomena are related as shall become amply clear in the following.

One is therefore led to inquire about corrections to mean field (and similarly $1/n$ corrections to the large n limit). It is natural to look for a small parameter in which to develop a systematic expansion. The most naive choice is the inverse of the coordination number, or equivalently $1/d$, even though the dimension d only takes integer values. The result will be to take into account self-consistently correlations between second, third, ... neighbours. In a similar spirit, one could study corrections to the Bethe approximation, although with different methods. This looks at first far from the desired goal, since it ignores collective long range excitations, which are presumed to be responsible for the deviations of critical exponents from their mean field values in low dimensions. Nevertheless, the method gives us some useful information. It enables one to locate accurately the critical points and gives quantitatively correct values outside the critical region. These data are useful for instance when compared to those of numerical simulations. It does not, however, modify the critical behaviour, which remains in the large distance limit the one of free fields (the Gaussian model). What is required to obtain a handle on long wavelength fluctuation effects, is a nonperturbative method, paradoxically emerging from a clever resummation of the perturbative series.

There are various ways to obtain $1/d$ corrections. Here we will illustrate one of them. In chapter 7, we shall present strong coupling expansions. Also we will limit ourselves to the Ising model, and derive the corrections from a study of fluctuations around a saddle point corresponding to mean field.

3.4.1 Laplace transform

If J_{ij} denotes a symmetric positive definite matrix, we can write the following Laplace transform formula

$$\exp\left(\frac{1}{2}\beta\sum_{i,j}\sigma_iJ_{ij}\sigma_j\right) = (\det\beta J)^{-1/2} \int \prod_i \frac{dH_i}{(2\pi)^{1/2}} \times \exp\left\{-\frac{1}{2\beta}\sum_{i,j}H_i(J^{-1})_{ij}H_j + \sum_i H_i\sigma_i\right\} \quad (138)$$

If the left-hand side is to be identified with a Boltzmann weight for the Ising model, we seem to encounter a difficulty. Choosing J_{ij} equal to 1 or 0 if sites are neighbours or not ($J_{ii} = 0$) leads to an operator with a symmetric spectrum of eigenvalues and the r.h.s. looks meaningless. This is reflected in the fact that the Gaussian integral is not convergent, nor is the square root of the determinant uniquely defined. This is however not a serious objection. Since $\sigma_i^2 = 1$, this can be corrected by shifting J_{ij} by a multiple of the identity at the price of multiplying all Boltzmann weights by a common factor. As we are only going to use the representation to perform an expansion around a saddle point, this shift would only complicate the expressions without producing any new information. We leave it to the careful reader to correct for this loose use of mathematics, and we proceed with (138) as it stands. Summing over $\sigma_i = \pm 1$ and dividing (conventionally) by 2^N , we obtain the Ising partition function in an external field h_i in the form

$$Z(\beta, h) = (\det\beta J)^{-1/2} \int \prod_i \frac{dH_i}{(2\pi)^{1/2}} \times \exp\left\{-\frac{1}{2\beta}\sum_{i,j}H_i(J^{-1})_{ij}H_j + \sum_i \ln \cosh(H_i + h_i)\right\} \quad (139)$$

The r.h.s. describes an assembly of independent spins interacting with a random (displaced) field H_i with a Gaussian distribution. We note that the weight of this probability measure involves β^{-1} instead of β . The quantities H_i are continuous variables, and the

saddle point method leads at once to the mean field equations

$$H_i = \beta \sum_j J_{ij} \tanh(H_j + h_j) \quad (140)$$

We expand the action in equation (139) around a solution $\bar{H}_i$, and treat the higher order terms in $(H_i - \bar{H}_i)$ perturbatively. Terms in $\ln Z$ can be classified according to the number of loops in the corresponding Feynman diagrams. The perturbative technique will be elaborated in chapter 5. For simplicity, we assume a constant external field h , in which case the uniform solution $\bar{H}$ is given by

$$\bar{H} = 2\beta d \tanh(\bar{H} + h) \quad (141)$$

To zeroth order (also called *tree level*), the free energy is given by the familiar expression

$$F_0 = \ln \cosh(\bar{H} + h) - \beta d [\tanh(\bar{H} + h)]^2 \quad (142)$$

with $\bar{H}$ a function of β and h , according to equation (141). The magnetization is, to this order,

$$M_0 = \frac{d}{dh} F_0(\bar{H}(h, \beta), h, \beta) = \tanh(\bar{H} + h) \quad (143)$$

Now perform in (139) the shift $H_i \rightarrow H_i + \bar{H}$ and expand the action to second order in H_i . With $Z_0 = e^{NF_0}$, we find

$$\begin{aligned} \frac{Z}{Z_0} &= (\det \beta J)^{-1/2} \\ &\int \prod_i \frac{dH_i}{(2\pi)^{1/2}} \exp -\frac{1}{2} H \times \left[(\beta J)^{-1} - \frac{1}{\cosh^2(\bar{H} + h)} \right] H \\ &= \left\{ \det \left[1 - \frac{1}{\cosh^2(\bar{H} + h)} \beta J \right] \right\}^{-1/2} \end{aligned} \quad (144)$$

The singular (ill-defined) quantity $(\det \beta J)^{-1/2}$ has disappeared, and the one loop correction to the free energy, call it F_1 , reads

$$\begin{aligned} F_1 &= -\frac{1}{2} \int \frac{d^d q}{(2\pi)^d} \ln \left[1 - \frac{2\beta}{\cosh^2(\bar{H} + h)} \sum_1 \cos q_\mu \right] \\ &= \frac{\beta^2 d}{2 \cosh^4(\bar{H} + h)} + O\left(\frac{\beta^4 d^2}{\cosh^6(\bar{H} + h)}\right) \end{aligned} \quad (145)$$

The expansion coefficient, corresponding to the topological number of loops, is analogous to $\hbar$ in the WKB approximation of quantum mechanics.

We can also reorganize these corrections to obtain a $1/d$ expansion. The expressions become more and more cumbersome as we take into account higher order terms in $(H_i - \bar{H})$ in the action. Let us confine ourselves here to the above one loop correction. For simplicity, write H for $\bar{H} + h$. Then the leading term (142) is

$$F_0 = V_0(H) + h \tanh H \quad (146)$$

$$V_0(H) = \ln \cosh H - H \tanh H + \beta d \tanh^2 H$$

The variables H and h are related by the saddle point condition

$$\frac{\partial}{\partial H} (V_0(H) + h \tanh H) = 0 \quad (147a)$$

and to this order the magnetization is given by

$$M_0 = \frac{\partial}{\partial h} (V_0(H) + h \tanh H) \quad (147b)$$

The first order correction

$$V_1(H) = -\frac{1}{2} \int \frac{d^d q}{(2\pi)^d} \ln \left[1 - \frac{2\beta}{\cosh^2 H} \sum_1 \cos q_\mu \right] \quad (148)$$

allows one to write the free energy as

$$\begin{aligned} F &= V_0 + V_1 + h \tanh H \\ dF &= \left(\tanh H + \frac{dV_1}{dH} \frac{dH}{dh} \right) dh \end{aligned} \quad (149)$$

where we have taken into account the lowest order saddle point condition (147a). Therefore, up to first order, the magnetization is given by

$$M = \tanh H + \frac{dV_1}{dH} \frac{dH}{dh} \quad (150)$$

It is now natural to perform a *Legendre transformation*, which amounts to using M instead of H as a variable. Expanding H in terms of M up to first order again, we have

$$\begin{aligned} H &= \operatorname{arctanh} M + H_1 + \dots \\ \tanh H &= M - \frac{dV_1}{dH} \frac{dH}{dh} \Big|_{H=\operatorname{arctanh} M} + \dots \end{aligned} \quad (151)$$

Set also

$$V(M) = -F + hM \quad (152)$$

with $V(M)$ called the effective potential, and

$$dV = h dM \quad (153)$$

Thus to leading order in the corrections

$$\begin{aligned} V &= -V_0(H) - V_1(H) + h \frac{dV_1}{d \tanh H} \frac{d \tanh H}{dh} + \dots \\ &= -[V_0(H) + V_1(H)]|_{H=\text{arctanh } M} \\ &\quad + \left(\frac{dV_0}{d \tanh H} \frac{dV_1}{d \tanh H} + h \frac{dV_1}{d \tanh H} \right) \frac{d \tanh H}{dh} \end{aligned}$$

By virtue of the fact that $dV_0/d \tanh H = -h$, the last term drops out, and we are left with

$$V(M) = -[V_0(H) + V_1(H)]|_{H=\text{arctanh } M} \quad (154)$$

Equation (153) expresses the fact that, in the absence of an external field h , the effective potential should be stationary. The zeroth order calculation suggests in fact that V should be at a minimum, if corrections are indeed sensible (that is small enough). This is the reason for the extra minus sign introduced in equations (152) and (154). One expects β_c to be defined by the condition that this minimum is for $M = 0$ as long as $\beta < \beta_c$, and that, beyond this point, $V(M)$ develops two new minima at $\pm M_{sp}$.

We can rewrite V as a function of M using

$$\begin{aligned} V_0(M) &= 2\beta d \frac{M^2}{2} - \frac{1}{2} \sum_{\pm} (1 \pm M) \ln(1 \pm M) \\ &= (2\beta d - 1) \frac{M^2}{2} - \frac{M^4}{12} + \dots \\ V_1(M) &= -\frac{1}{2} \int \frac{d^d q}{(2\pi)^d} \ln \left[1 - 2\beta(1 - M^2) \sum_1^d (1 - \cos q_\mu) \right] \\ &= -\frac{1}{2} \int \frac{d^d q}{(2\pi)^d} \ln \left[1 - 2\beta \sum_1^d (1 - \cos q_\mu) \right] \\ &\quad - M^2 \int \frac{d^d q}{(2\pi)^d} \frac{\beta \sum_1^d (1 - \cos q_\mu)}{1 - 2\beta \sum_1^d (1 - \cos q_\mu)} \end{aligned}$$

$$+ M^4 \int \frac{d^d q}{(2\pi)^d} \left[\frac{\beta \sum_1^d (1 - \cos q_\mu)}{1 - 2\beta \sum_1^d (1 - \cos q_\mu)} \right]^2 + \dots \quad (155)$$

The expansion has coefficients which are singular at the original critical point $2\beta_c^{(0)}d = 1$, as is expected in a perturbative method. In any case, the origin is an extremum of $V(M)$, but new minima show up when the coefficient of the M^2 term vanishes, i.e. for a new value β_c given consistently to this order, by

$$2\beta_c d = \int \frac{d^d q}{(2\pi)^d} \frac{1}{1 - 2d\beta_c \frac{1}{d} \sum_1^d (1 - \cos q_\mu)} = G(0, 2d\beta_c) \quad (156)$$

where G is the familiar Green's function for Brownian motion. We note the similarity with the large n result - equations (102) and (103) - which was $2d\beta_c = G(0, 1)$. From chapter 1, $G(0, 1)$ admits an expansion of the form $1 + 1/2d + \dots$ hence the leading correction in both cases is

$$2d\beta_c = 1 + \frac{1}{2d} + \dots \quad (157)$$

The sign is in agreement with the idea that the effect of fluctuations is to decrease the critical temperature, hence to increase β_c .

The perturbative expansion can be carried beyond quadratic terms. Show that for an $O(n)$ symmetric model the following d^{-1} expansion holds (Fisher and Gaunt, Abe)

$$\frac{1}{2d\beta_c} = 1 - \frac{1}{2d} - \frac{1}{(2d)^2} \left(2 - \frac{2}{n+2} \right) - \frac{1}{(2d)^3} \left(7 - \frac{8}{n+2} \right) + \dots \quad (158)$$

For $n \rightarrow \infty$ we recover the result (102)-(104), namely $1/2d\beta_c = 1 - \Pi(0)$, where $\Pi(0)$ is the probability that a Brownian particle returns to the origin, which becomes 0 as $d \rightarrow 2$. One therefore expects the series (158) to become singular for $d = 2$, but this remains an open problem. In particular, $d = 4$ should not be a singularity. We also notice that the value $n = -2$ seems to play a special role. Such a value corresponds to a free complex fermionic model, provided the continuation in the parameter n is adequately defined.

Let us now compute the susceptibility in the high temperature phase

$$\beta < \beta_c \quad \chi^{-1} = \frac{\partial^2 V}{\partial M^2} \Big|_{M=0} \quad (159)$$

The previous calculation yields the location of the zero β_c of χ^{-1} . By subtracting the two expressions, we get, by approximating β_c by its leading order value $1/2d$ in the denominator,

$$[2d(\beta_c - \beta)\chi]^{-1} = 1 - \int \frac{d^d q}{(2\pi)^d} \times \frac{\frac{1}{d} \sum_1^d (1 - \cos q_\mu)}{\left[1 - \frac{1}{d} \sum_1^d (1 - \cos q_\mu)\right] \left[1 - 2\beta \sum_1^d (1 - \cos q_\mu)\right]} + \dots \quad (160)$$

For $d > 4$, as $\beta \rightarrow \beta_c \sim 1/2d$, the integral converges, and the critical exponent γ is 1. But for $d \leq 4$, the (infrared) effect of fluctuations is catastrophic in this perturbative scheme. The integral develops logarithmic singularities at $d = 4$, and the higher order terms have an even more divergent behaviour. Thus one finds a criterion of upper critical dimensionality (Landau-Ginzburg) telling us how the mean field approximation breaks down under the effect of long wavelength fluctuations. These divergences will be resummed by the renormalization group method to produce the expected finite deviations to critical exponents.

One could proceed to compute various $1/d$ corrections for other quantities above $d = 4$, and one can also generalize the method to other models characterized by one or several order parameters. Other strong coupling expansions are also available.

We will also have to return to the implementation of mean field theory when dealing with systems which do not admit a local order parameter, the typical cases being gauge field models.

Effect of an inhomogeneous external field. Let F denote now the total free energy. Up to one loop order

$$\begin{aligned} F &= V_0 + V_1 + \sum_i h_i \tanh H_i \\ V_0 &= \sum_i \ln \cosh H_i - H_i \tanh H_i + \frac{1}{2} \beta \sum_{i,j} (\tanh H_i) J_{ij} (\tanh H_j) \\ V_1 &= -\frac{1}{2} \text{Tr} \ln [\delta_{ij} - \beta J_{ij} (1 - \tanh^2 H_j)] \end{aligned} \quad (161)$$

The relation between H and h is given by

$$\frac{\partial}{\partial H_i} \left\{ V_0 + \sum_i h_i \tanh H_i \right\} = 0 \quad (162)$$

Let M_i denote the derivative of F with respect to h_i

$$M_i = \frac{d}{dh_i} F(H_i(h), h_i) \quad (163)$$

then the Legendre transform Γ as a functional of M_i is defined through

$$\Gamma(M) + F(h) = \sum_i h_i M_i \quad (164)$$

and it generalizes the effective potential (which appears as the density of Γ if M is homogeneous). Up to one loop,

$$\begin{aligned} \Gamma(M) &= - (V_0 + V_1) |_{M_i = \tanh H_i} \\ &= -\frac{1}{2} \beta \sum_{ij} M_i J_{ij} M_j + \frac{1}{2} \sum_{i,\pm} (1 \pm M_i) \ln(1 \pm M_i) \\ &\quad + \frac{1}{2} \text{Tr} \ln [\delta_{ij} - \beta J_{ij} (1 - M_j^2)] \end{aligned} \quad (165)$$

and h_i is then obtained as

$$\frac{\partial \Gamma}{\partial M_i} = h_i \quad (166)$$

Generalize the Legendre transformation to all orders.

Before presenting the general analysis of the perturbation series, we will devote the next chapter to general considerations on scaling and models of renormalization.

Notes

Mean field is almost as old as statistical mechanics. A standard reference is W.L. Bragg, E.J. Williams, *Proc. Roy. Soc. A* 145, 699 (1934). Corrections to the Clausius Mossotti formula for dense systems discussed in J. De Boer, F. Van der Maesen and C.A. Ten Seldam, *Physica* 19, 265 (1953) and G. Stell and G.S. Rushbrooke, *Chem. Phys. Lett.* 24, 531 (1974) are based on the formalism developed by J.G. Kirkwood, *J. Chem. Phys.* 4, 592 (1936) and J. Yvon, *Actualités scientifiques et industrielles*, 543,

Hermann, Paris (1937). The case of polar media was investigated by L. Onsager, *J. Am. Chem. Soc.* **58**, 1496 (1936).

For an overall view on critical phenomena as well as their physical context, see H.E. Stanley *Introduction to Phase Transitions and Critical Phenomena*, Clarendon Press, Oxford (1971). The Landau approach is also discussed extensively in L.D. Landau and E.M. Lifshitz *Statistical Physics*, third edition revised by E.M. Lifshitz and L.P. Pitaevskii Pergamon Press, Oxford (1980). The Bethe approximation is introduced in H. Bethe, *Proc. Roy. Soc. A*-**216**, 45 (1935). See also G.S. Rushbrooke and H.I. Scoins, *Proc. Roy. Soc. A*-**230**, 74 (1953). Zero modes arising in a model with a spontaneous broken continuous symmetry are discussed in J. Goldstone, *Nuovo Cimento*, **19**, 154 (1961). Analogies between the mechanism of superconductivity and particle mass generation, a subject not covered here, is found in Y. Nambu and G. Jona-Lasinio, *Phys. Rev.* **122**, 345 (1961) and **124**, 246 (1961). An influential paper in the context of particle physics, introducing the σ -model, is the one by M. Gell-Mann and M. Lévy, *Nuovo Cimento* **16**, 705 (1960). Various aspects of symmetry breaking, and much more, are discussed in depth in the collected Erice lectures of S. Coleman *Aspects of Symmetry*, Cambridge University Press, Cambridge (1985).

The investigation of complex zeroes is the work of C.N. Yang and T.D. Lee, *Phys. Rev.* **87**, 404,410 (1952). Our presentation is inspired by D. Ruelle, *Phys. Rev. Lett.* **26**, 303 (1971), *Comm. Math. Phys.* **31**, 265 (1973). For further developments see J.D. Bessis, J.-M. Drouffe, P. Moussa, *J. Phys.* **A9**, 2105 (1976) and D.A. Kurtze, M.E. Fisher, *Phys. Rev.* **B20**, 2785 (1979). The zeroes in the complex temperature plane are but one minor item in the Boulder Lectures by M.E. Fisher, in *Lectures in Theoretical Physics*, VII-C, edited by W.E. Brittin, University of Colorado Press, Boulder (1964). See also C. Itzykson, R.B. Pearson, J.-B. Zuber, *Nucl. Phys.* **B220 (FS8)** 415 (1983), and numerical studies by E. Marinari, *Nucl. Phys.* **B235 [FS11]**, 123 (1984), and G. Bhanot, R. Salvador, S. Black, P. Carter and R. Toral, *Phys. Rev. Lett.* **59**, 803 (1987).

The spherical model was solved by T.H. Berlin and M. Kac, *Phys. Rev.* **86**, 821 (1952). For the connection with the large n -expansion, see H.E. Stanley, *Phys. Rev.* **176**, 718 (1968). The Coleman lectures quoted above contain a review as well as further

developments initiated by G. 't Hooft for models involving matrix valued fields, to be discussed in chapter 10.

The large d expansions quoted in the last section are from M.E. Fisher and D.S. Gaunt, *Phys. Rev.* **133A**, 224 (1964) and R. Abe, *Prog. Theor. Phys.* **47**, 62 (1972).

We have not treated the question of defects in ordered media, a legitimate part of the physics of phase transitions. Numerous aspects of this fascinating subject are covered in the book of M. Kléman, *Points, Lines and Walls in Liquid Crystals, Magnetic Systems and Various Ordered Media*, John Wiley and Sons, New York (1983), including references to many original works. For a review of the topology of defects, see N.D. Mermin, *Rev. Mod. Phys.* **51**, 591 (1979) and L. Michel, *Rev. Mod. Phys.* **52**, 617 (1980).

SCALING TRANSFORMATIONS AND THE XY-MODEL

Condensed media show a large variety of critical phenomena, ranging from critical opalescence at the end point of the liquid-gas coexistence curve, the Curie transition of ferromagnetic materials, to the superfluid transition of helium, the behaviour of solutions of polymers, the conductivity of random media, ... To this list should be added systems with local symmetries, such as those suggested by particle physics in order to understand quark confinement. Surprisingly all these phenomena can be classified into a few universality classes, characterized by specific large distance behaviour with the same critical exponents. In this chapter, we sketch the methods and ideas of the renormalization procedure. We illustrate the concepts with simple approximations in the language of classical spin models in the first section, treating in more detail the XY-model as an example in the second section.

4.1 Scaling laws. Real space renormalization

4.1.1 Homogeneity and scale invariance

The discussion given in previous chapters suggests that, close to a continuous transition, critical systems exhibit universal properties. Correlations at large distance are not sensitive to the details of microscopic interactions. Their behaviour is described by a specific dimensional analysis governed by some essential characteristics of the system, such as the dimension of space, the nature of the order parameter and the underlying symmetries.

The mean field approximation gave a first idea of a simple critical behaviour. Fluctuations have only a quantitative effect in dimension greater than the upper critical one. This dimension is generally equal to four for spin systems, and the role of

fluctuations becomes nontrivial in four and lower dimensions. With T_c being the critical temperature, we use as a control parameter the reduced variable

$$\theta = \left(\frac{T}{T_c} - 1 \right) \quad (1)$$

In the language of field theory, the bare mass squared is proportional to θ . The domain $\theta > 0$ corresponds to a symmetric disordered phase, and, for $\theta < 0$, the symmetry is spontaneously broken. In this case, any uniform boundary condition or any infinitesimal uniform external field allows one to isolate a pure phase in which a given direction for the spontaneous magnetization has been singled out among all those equivalent with respect to the symmetry group. A continuous transition is characterized by a spontaneous magnetization vanishing continuously at $\theta = 0$. In this phase, there generally remains a reduced subgroup of symmetry, the isotropy group around this privileged direction. The size of the domain near $\theta = 0$ in which the spacing a , or more generally the interaction range, can be neglected with respect to the correlation length ξ depends on the microscopic parameters, and can be therefore difficult to estimate in a concrete case. Our hypothesis is $\xi \gg a$. This correlation length can be visualized (for a discrete symmetry group) by considering a typical configuration near $\theta = 0$, and estimating the size of clusters of spins with the same orientation. At the critical point, ξ is infinite, and this length scale itself disappears. We also assume that correlations become isotropic near the critical point.

In order to simplify the reasoning, let us consider a scalar field and let us measure lengths in terms of a microscopic spacing a . Ignoring anisotropies, a correlation function

$$G(\mathbf{x}) = \langle \varphi(\mathbf{x}) \varphi(0) \rangle \quad (2)$$

satisfies *a priori*

$$\frac{G(r_2)}{G(r_1)} = \gamma \left(\frac{r_2}{r_1}, \frac{r_1}{a} \right)$$

with $r = |\mathbf{x}|$. The aim of the theory is to justify the existence of the critical limit $a \rightarrow 0$ for fixed r_1 and r_2 , in which case

$$\frac{G(r_2)}{G(r_1)} = \gamma \left(\frac{r_2}{r_1} \right) \quad r_1 \gg a \quad r_2 \gg a \quad (3)$$

The group law

$$\frac{G(r_3)}{G(r_1)} = \frac{G(r_3)}{G(r_2)} \times \frac{G(r_2)}{G(r_1)}$$

requires that $\gamma(\rho)$ be a homogeneous function equal to unity for $\rho = 1$. In other words, one should have

$$G(r_2) = G(r_1) \left(\frac{r_1}{r_2} \right)^{d-2+\eta} \quad \text{at } \theta = 0 \quad (4)$$

with the conventional definition of the critical exponent η . This exponent vanishes in the mean field approximation, and thus measures the deviation from the Gaussian model. Hence $\eta \neq 0$ corresponds to a nontrivial scale invariant field theory. In the disordered phase $\theta > 0$, very close to the critical point (so that $\xi \gg c$), we can distinguish three domains of lengths r , large with respect to the lattice spacing or to any short-distance characteristic scale

$$\begin{cases} (i) & a \ll r \ll \xi \\ (ii) & a \ll r \sim \xi \\ (iii) & a \ll \xi \ll r \end{cases}$$

In the first domain, the preceding considerations remain valid since the correlation length can be considered as infinite. This region is equivalent to the ultraviolet region of a renormalizable field theory, as will be shown in the next chapter. The second region is a delicate zone of transition, where one begins to deviate from a homogeneous behaviour. Finally, correlation functions decrease exponentially in the third region, diverging more and more from the scale invariant regime. The temperature, characterized by the parameter θ , is an essential or relevant parameter for the large-distance (infrared) behaviour. This behaviour shows an instability as θ departs from 0. In other words, as the distance increases, the system becomes more and more sensitive to a deviation from the critical temperature.

These considerations are summarized by the expression

$$G(r, \theta) \sim \frac{1}{r^{d-2+\eta}} g\left(\frac{r}{\xi(\theta)}\right) \quad (5)$$

The function g is regular at the origin, up to logarithmic terms in marginal cases, while it decreases exponentially at infinity.

As θ vanishes, the mean field approximation suggests a divergence of the correlation length according to a power law

$$\xi(\theta) \sim \xi_+ \theta^{-\nu} \quad \text{as } \theta \rightarrow 0 \quad (6)$$

which introduces a second exponent ν . The quantity ξ_+ is of course dependent on the microscopic structure of the interactions. The hypothesis (6) allows one to relate the length (or mass) scale to the temperature (in field theoretic language, the physical mass to the bare mass). Landau's theory predicts $\nu = \frac{1}{2}$, while $\nu = 1$ for the two-dimensional Ising model.

Generalizing these ideas, one is naturally led to a series of scaling laws for the various physical quantities, stemming from the *scaling hypothesis* that ξ is the only relevant intrinsic scale length. Let us consider the free energy per unit volume. It is expected that fluctuations contribute to its singular part as the ratio of the unit volume to the volume $\xi(\theta)^d$

$$F_{\text{sing}}(\theta) \sim \frac{1}{\xi(\theta)^d} \sim \theta^{\nu d} \quad (7)$$

This is the so-called *hyperscaling hypothesis*. As a consequence, the exponent α of the singular part of the specific heat C , proportional to the second derivative of F , is given by

$$\begin{aligned} C(\theta) &\sim \theta^{-\alpha} \\ \alpha &= 2 - \nu d \end{aligned} \quad (8)$$

A positive value of α implies a divergence of the specific heat, whereas a negative value corresponds to a cusp in its representative curve as a function of temperature.

In order to justify equation (7), Pippard and Ginsberg gave the following argument. Suppose that the free energy per unit volume has a singular part behaving as $\theta^{2-\alpha}$. At a temperature greater than the critical temperature ($\theta > 0$), a fluctuation corresponding to a departure from an ordered state in a range ξ produces an increase $\Delta F \sim \theta^{2-\alpha} \xi^d$ of the free energy. The probability $e^{-\Delta F}$ of such a fluctuation becomes negligible when $\Delta F \simeq 1$, and one gets the estimate $\xi \sim \theta^{-(2-\alpha)/d}$.

The logarithmic singularity of the specific heat encountered in the Ising model is interpreted as $\alpha = 0$. This is in agreement with (7), according to which $\nu = 1$. Similarly, the spherical model

verifies this relation for $2 < d < 4$, with

$$\alpha = \frac{d-4}{d-2} \quad \nu = \frac{1}{d-2} \quad (9)$$

Relation (7) is violated by the mean field approximation where we found $\alpha = 0$ and $\nu = \frac{1}{2}$, independently of d . The above reasoning is indeed incompatible with the one leading to Landau's approximation, where fluctuations are ignored. To estimate the upper critical dimension d_c , we can nevertheless use these mean field values to predict that fluctuations, which are negligible in large dimensions, should become important as the dimension d decreases below d_c given by $\alpha_{mf} = 2 - \nu_{mf}d_c$, i.e. $d_c = 4$. We recover here the Landau-Ginzburg criterion for the upper critical dimension.

The susceptibility, which is the second derivative of the free energy with respect to the external field when this external field vanishes, is proportional to the integral of the two-point correlation function. Using (5), one deduces that

$$\chi(\theta) = \int_{|x| < \xi(\theta)} d^d x \frac{1}{|x|^{d-2+\eta}} \sim \xi(\theta)^{2-\eta} \quad (10)$$

The corresponding critical exponent denoted γ is therefore given by

$$\begin{aligned} \chi(\theta) &\sim \theta^{-\gamma} \\ \gamma &= \nu(2 - \eta) \end{aligned} \quad (11)$$

In the presence of an external field H , the transition disappears. The magnetization $M(H, \theta)$ vanishes linearly with H for $\theta > 0$, with a slope $\chi(\theta)$. For $\theta < 0$, $M(H, \theta)$ tends towards the spontaneous magnetization $M(\theta)$ when $H \rightarrow +0$. As one approaches the critical temperature from below, this spontaneous magnetization vanishes as

$$M(\theta) \sim (-\theta)^\beta \quad \text{as } \theta \rightarrow -0 \quad (12)$$

At the critical temperature, the magnetization is singular when H vanishes, with an exponent δ^{-1}

$$M(H, \theta = 0) \sim H^{1/\delta} \quad \text{as } H \rightarrow 0 \quad (13)$$

Relations (12) and (13) suggest that the quantities H, M^δ and $(-\theta)^{\beta\delta}$ are related. The *equation of state* to be derived in the

next subsection expresses $HM^{-\delta}$ as a regular function of $\theta M^{-1/\beta}$

$$H = M^\delta f_H(\theta M^{-1/\beta}) \quad (14)$$

The first negative zero of f_H corresponds to the spontaneous magnetization, and its value at zero yields the proportionality coefficient in equation (13). For $\theta > 0$, M and H are linearly related when they vanish simultaneously, $M = \chi(\theta)H$. Hence the function $f_H(z)$ behaves as $z^{\beta(\delta-1)}$ for large z . This yields $H \sim M\theta^{\beta(\delta-1)}$. Comparing with (11), we obtain the relation

$$\gamma = \beta(\delta - 1) \quad (15)$$

This illustrates the general method for relating singular quantities through regular relations. If the specific heat singularity is the same above and below the critical temperature, the equation of state (14) can be reinterpreted as a relation for the singular part of the free energy expressed in terms of H and θ . For $\theta > 0$, one writes

$$F_{\text{sing}}(\theta, H) \sim \theta^{2-\alpha} f_F(H\theta^{-\beta\delta}) \quad (16)$$

the second derivative with respect to H gives the susceptibility, and hence

$$\gamma = 2\beta\delta + \alpha - 2 \quad (17)$$

The relations (15) and (17) allow the determination of β and δ .

To summarize, the two critical exponents, η for the behaviour of the Green function at the critical point, and ν for the correlation length in the critical domain yield the four other exponents (using the scaling hypothesis) which are

$$\begin{aligned} \alpha &= 2 - \nu d \\ \beta &= \frac{1}{2}\nu(d - 2 + \eta) \\ \gamma &= \nu(2 - \eta) \\ \delta &= \frac{d + 2 - \eta}{d - 2 + \eta} \end{aligned} \quad (18)$$

The goal of the renormalization group is to justify the above relations and to provide a tool for computing critical exponents and other universal quantities, which can also be obtained from perturbative expansions at low or high temperature. It seems that these two different approaches, including other numerical methods

(such as finite size scaling, Monte-Carlo renormalization,...) are all in reasonable agreement. Delicate experiments confirm the theoretical predictions.

Let us add a remark concerning the Lee-Yang singularity. The magnetization is a regular function of the temperature in the real domain $H \neq 0$. However, its analytic continuation for complex H is singular (for $\theta > 0$) on the imaginary axis, starting at $\pm iH_c$. The value H_c can be estimated from equation (16) in the critical region

$$H_c \sim \theta^{\beta\delta} = \theta^{\beta+\gamma} \quad (19)$$

For d large, $\beta = \frac{1}{2}$, $\delta = 3$ and $\gamma = 1$ and one recovers $H_c \sim \theta^{3/2}$. On the other hand, the nature of the singularity cannot be obtained from the behaviour in the critical domain. Another length scale is involved, which becomes infinite at the Lee-Yang point. Relation (19) only expresses a compatibility relation between the two singularities.

4.1.2 Recurrence relations in real space

The behaviour of systems near the critical point can be analyzed by studying the effect of a change in scale (dilatation). This is the essential point elaborated by the various methods of renormalization, which will lead us, in the framework of field theory, to the Callan-Symanzik equations.

We have just established relations between critical exponents, using directly the observables and assuming only homogeneity. To study a scaling transformation, we will proceed at first in configuration space, although the sophisticated methods of the continuous theory use momentum space. We start from the elementary spacing (characterizing the finite range of the forces). In order to implement a change in scale and to erase progressively the microscopic details, there exist several techniques (decimation, block variables, ...). None of these methods is completely satisfactory, although they offer the simplest approach conceptually.

On the example of the Ising model, we present first the block spin technique introduced by Kadanoff. Let us consider a new lattice with spacing λa , λ being an integer, grouping the spins in clusters. To each of these blocks, we associate a new dichotomic

variable σ_b . For instance, one can choose the majority rule

$$\sigma_b = \text{Sign} \sum_{i \in b} \sigma_i \quad (20)$$

with an additional convention in the case where this sum vanishes (when this can happen). If the correlation length is large with respect to the elementary spacing, one expects that fluctuations within a given block are inessential, as long as λa is much smaller than ξ . In the original variables, the statistical weight e^S is given in terms of the action

$$S_a = \beta \sum_{(i,j)} \sigma_i \sigma_j + H \sum_i \sigma_i \quad (21)$$

with H small and β near the critical point. One wants to define a new action $S_{\lambda a}$ for the block variables, leading of course to the same partition function. This action is formally simple to construct by performing the constrained sum

$$e^{S_{\lambda a}} = \sum_{\{\sigma_i = \pm 1\}} \left[\prod_b \delta \left(\sigma_b - \text{Sign} \sum_{i \in b} \sigma_i \right) \right] e^{S_a} \quad (22)$$

It is clear that the partition function is invariant

$$Z = \sum_{\{\sigma_b = \pm 1\}} e^{S_{\lambda a}} = \sum_{\{\sigma_i = \pm 1\}} e^{S_a} \quad (23)$$

In contradistinction with the decimation method (where some initial variables are held fixed while others are summed over), the block variables cannot be identified to some of the original ones. Hence, correlations are not *a priori* invariant in a change of scale $a \rightarrow \lambda a$.

The equation (22) leads to an intricate form for $S_{\lambda a}$, and the constrained summation is almost as difficult to perform as the original one. Nevertheless, $S_{\lambda a}$ has the same global symmetry as the initial system. Moreover, if the parameter λ could vary continuously, one would expect that $S_{\lambda a}$ is similar to S_a in the vicinity of $\lambda = 1$. On the lattice, the transformation $S_a \rightarrow S_{\lambda a}$ is difficult to study without additional hypotheses or approximations. Simplifying drastically and using the idea of universality, we assume that, up to a constant, $S_{\lambda a}$ is well approximated by nearest neighbour interactions and a linear

symmetry breaking term

$$S_{\lambda a} \cong \beta_{\lambda} \sum_{(b,b')} \sigma_b \sigma_{b'} + H_{\lambda} \sum_b \sigma_b + S_0 \quad (24)$$

Therefore, interactions between next-to-nearest neighbours or farther away sites, as well as interactions between more than two sites (including odd symmetry breaking terms) are neglected. This will be justified later, showing that these interactions become inessential after several iterations of the rescaling transformation. In a general reasoning, they should however be included, and thus the initial action might as well be assumed to depend on an infinite number of parameters. For our present purposes, we restrict ourselves to the approximation (24). Even more, in a first stage, we suppress the external field H , so that H_{λ} also vanishes. Hence, up to the constant term S_0 , the block variables lead to a new Ising model, with nearest neighbour interactions on a new length scale. The transformation reduces to

$$\begin{cases} \mathbf{x} \rightarrow \mathbf{x}_{\lambda} = \mathbf{x}/\lambda \\ \theta \rightarrow \theta_{\lambda} = \varphi(\lambda, \theta) \end{cases} \quad (25)$$

which expresses that the physical quantities of the initial model relative to distances x and temperature θ are related to the transformed model with distances x_{λ} and temperature θ_{λ} . The additive quantity S_0 and the transformation law $\theta_{\lambda} = \varphi(\lambda, \theta)$ are regular functions of the initial parameters. This is an essential point, the renormalization transformation does not deal with singular functions.

Relation (23) expresses the conservation of the partition function and allows the determination of the free energy per site

$$\lambda^d F(\theta) = F(\theta_{\lambda}) + F_0(\theta) \quad (26)$$

where $F_0(\theta)$ is a regular function of θ depending smoothly on λ . Note that the same function F appears on both sides. This is due to the fact that renormalization preserves the structure of the model (with modified parameters). Here this relation is only approximate since the number of parameters has been limited; it becomes exact, but almost useless, in the infinite dimensional space of parameters describing all short-range interactions.

The effect of iterations of the transformation can be analyzed qualitatively. When the temperature differs from the critical tem-

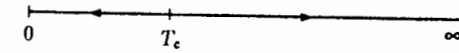


Fig. 1 Instability of the critical point with respect to the temperature.

perature, the block size increases and approaches the correlation length. Hence, in units of block lattice spacing, the new correlation length decreases, so that the deviation from the critical temperature increases. In particular, if $\beta < \beta_c$, β_{λ} decreases to zero, the high temperature attractive fixed point, as λ increases indefinitely. A similar scheme is valid for $\beta > \beta_c$, where renormalization drives the effective temperature to zero. Recall that, in the ordered phase, ξ describes the range of fluctuations around the spontaneous magnetization. In other words, renormalization amplifies the deviations from the critical temperature, which appears as a nontrivial, unstable infrared fixed point (figure 1). This property allows one to identify with a great sensitivity the critical point.

In the critical domain, assuming that it is possible to consider λ as a continuous parameter, one expects a behaviour near $\lambda = 1$ of the form

$$\theta_{\lambda} = \lambda^{y_{\theta}} \theta \quad (27)$$

with $y_{\theta} > 0$. The correlation length transforms as

$$\xi(\theta_{\lambda}) = \xi(\theta)/\lambda$$

One can iterate the transformation until $\xi(\theta_{\lambda})$ and θ_{λ} are of order unity which yields $\lambda \sim \theta^{-1/y_{\theta}}$. Hence

$$\xi \sim |\theta|^{-1/y_{\theta}} \quad \rightarrow \quad \nu = \frac{1}{y_{\theta}} \quad (28a)$$

The exponent is given by

$$y_{\theta} = \frac{1}{\nu} = \left. \frac{\lambda}{\theta_{\lambda}} \frac{\partial}{\partial \lambda} \theta_{\lambda} \right|_{\lambda=1} \quad (28b)$$

again assuming that λ is a continuous variable.

The hypothesis that the renormalization transformation is regular, even in the neighborhood of the critical point, leads to the same exponent ν above and below the critical temperature. Equation (28) is only meaningful if one can perform renormalization operations depending continuously on the dilatation parameter λ .

On a lattice, λ can only take discrete values, and, writing near the critical point $\theta_\lambda = \mu\theta + o(\theta)$, an estimate of y_θ is given by $\ln \mu / \ln \lambda$ with μ the derivative of the transformation at the critical point.

Assuming in equation (26) that the dependence of $F_0(\theta)$ on λ is not a serious problem, even when λ gets as large as $\theta^{-1/\nu}$ (admittedly a weak point of the argument), the singular part of the free energy satisfies the homogeneous equation

$$\lambda^d F_{\text{sing}}(\theta) = F_{\text{sing}}(\theta_\lambda) \quad (29)$$

and hence, according to (27), if we pick $\lambda \sim \theta^{-1/\nu}$,

$$F_{\text{sing}}(\theta) = f_\pm |\theta|^{2-\alpha} \quad (30)$$

which introduces the exponent α , given as before by

$$2 - \alpha = \nu d \quad (31)$$

The exponent α is the same above and below the critical temperature, although the amplitudes f_+ and f_- may differ.

This reasoning can be generalized in the presence of an external field H . Now $\beta_\lambda(\beta, H)$ and $H_\lambda(\beta, H)$ are both functions of β and H , the last one vanishing with H for reasons of symmetry. The critical point is characterized by the vanishing of both θ and H . We have in general to diagonalize the renormalization flow in the vicinity of the fixed point. We call again these variables θ and H , and hence we get in the critical domain

$$\begin{aligned} \theta_\lambda &= \lambda^{y_\theta} \theta \\ H_\lambda &= \lambda^{y_H} H \\ \lambda^d F_{\text{sing}}(\theta, H) &= F_{\text{sing}}(\theta_\lambda, H_\lambda) \end{aligned} \quad (32)$$

Previous hypotheses and definitions yield

$$F_{\text{sing}}(\theta, H) = |\theta|^{\nu d} F_{\text{sing}}(\pm 1, H/|\theta|^{\nu y_H}) \quad (33)$$

Differentiating with respect to H , one obtains the magnetization and the susceptibility

$$\begin{aligned} M &\sim (-\theta)^{\nu(d-y_H)} & \beta &= \nu(d-y_H) \\ \chi &\sim \theta^{-\nu(2-y_H-d)} & \gamma &= \nu(2y_H-d) \end{aligned} \quad (34)$$

In the present scheme, $y_\theta = 1/\nu$ and $y_H = \frac{1}{2}(d + \gamma/\nu)$ are the fundamental exponents which enable us to derive all the others. Combining relations (31) and (34), one gets

$$\alpha + 2\beta + \gamma = 2 \quad (35)$$

By differentiating equation (33) with respect to H , the magnetization can be re-expressed in terms of the field and the temperature as

$$H = M^{\nu y_H/\beta} f(\theta M^{-1/\beta}) \quad (36)$$

Comparing with (14), one gets

$$\begin{aligned} \delta &= \frac{\nu y_H}{\beta} = \frac{y_H}{d - y_H} \\ 2 - \alpha &= \nu d = \beta(1 + \delta) \end{aligned} \quad (37)$$

Scaling transformations can also be applied to correlation functions, taking into account the invariance of the partition function. For example, in the domain $\theta \geq 0$, one finds

$$\begin{aligned} Z^{-1} \delta^2 Z &= \sum_{\text{blocks}} \delta H_\lambda(\mathbf{x}'_1) \delta H_\lambda(\mathbf{x}'_2) G(\mathbf{x}'_1 - \mathbf{x}'_2, \theta_\lambda, H_\lambda) \\ &= \sum_{\text{sites}} \delta H(\mathbf{x}_1) \delta H(\mathbf{x}_2) G(\mathbf{x}_1 - \mathbf{x}_2, \theta, H) \end{aligned} \quad (38)$$

$$G(\mathbf{x}, \theta, H) = \lambda^{2(y_H-d)} G(\mathbf{x}/\lambda, \theta_\lambda, H_\lambda) \quad (39)$$

The prefactor arises from condition (38) and expresses the change of the unit of length as well as of the scale of the magnetic field. It is interpreted in the framework of field theory as a "wavefunction" renormalization, corresponding to a nontrivial transformation on the spin variables.

When iterating (39), $\mathbf{x}$ decreases and θ and H increase, so that one leaves the critical region. Assuming that correlations are isotropic in this domain, we have

$$G(\mathbf{x}, \theta, H) = \frac{1}{|\mathbf{x}|^{2(d-y_H)}} g(\theta |\mathbf{x}|^{1/\nu}, H |\mathbf{x}|^{y_H}) \quad (40)$$

In the absence of an external field, we recover the interpretation of $\xi \sim \theta^{-\nu}$ as a correlation length. At the critical temperature $\theta = 0$, one has rather $\xi \sim H^{-1/y_H}$. At the critical point $\theta = H = 0$, the correlation function is homogeneous in $|\mathbf{x}|$ and the critical exponent η verifies

$$d - 2 + \eta = 2(d - y_H) \quad (41)$$

or equivalently

$$\nu(2 - \eta) = \nu(2y_H - d) = \gamma \quad (42)$$

Summarizing, the renormalization method in real space introduces two fundamental exponents $y_\theta = 1/\nu$ and y_H which characterize the instability in even (θ) and odd (H) directions in coupling space. All other exponents are related to the previous ones

$$\begin{aligned} \alpha &= 2 - \nu d \\ \beta &= \nu(d - y_H) \\ \gamma &= \nu(2y_H - d) \\ \delta &= \frac{y_H}{d - y_H} \\ \eta &= 2 + d - 2y_H \end{aligned} \quad (43)$$

More generally, let us discuss, following Wegner, the behaviour of a set of couplings $\{K_i\}$ under a scale transformation λ which can be taken arbitrarily close to unity. The critical point is assumed to be at $K_i = 0$. Linearizing the transformation near this point, one writes

$$\lambda \frac{d}{d\lambda} K_i(\lambda) \Big|_{\lambda=1} = y_{ij} K_j \quad (44)$$

The eigenvectors of the matrix y_{ij} define linear combinations of the interactions, called *scaling* (or *operator*) *fields*. Their coefficient in the action have a simple power law behaviour under the renormalization transformation. They are classified in three categories

- i) *relevant* if the corresponding eigenvalue is positive,
- ii) *irrelevant* if it is negative,
- iii) *marginal* if it vanishes.

In particular, in our example, the deviation from the critical temperature and the magnetic field are *relevant*. Irrelevant couplings tend toward zero by renormalization transformation and can thus be neglected in the critical domain, where they only generate corrections to dominant terms. The case of marginal operators is more interesting and requires a study beyond the linear approximation.

This global approach allows one to sharpen the hypotheses of scale invariance theory when one describes the effects of renormalization in the space of coupling constants (figure 2). The critical manifold corresponds to the vanishing of relevant couplings. The

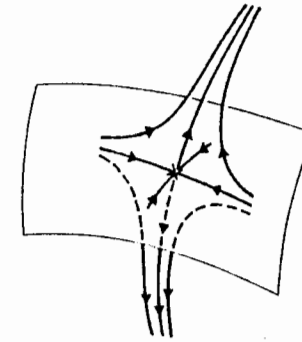


Fig. 2 A pictorial view of the renormalization group trajectories near the critical point.

codimension of this manifold is equal to the number of positive eigenvalues of the matrix y_{ij} . On this manifold, the renormalization transformations result in a contraction. Conversely, starting from a point out of this manifold, renormalization trajectories will finally run away. In other words, the concept of renormalization allows us to concentrate on a few relevant or marginal couplings, thus being intimately tied with the idea of universality.

Generalizing (39) to the correlations of n spins, show that, near the critical point,

$$G(\mathbf{x}_1, \dots, \mathbf{x}_n; \theta, H) = Z(\lambda)^{n/2} G(\mathbf{x}_1/\lambda, \dots, \mathbf{x}_n/\lambda; \theta_\lambda, H_\lambda) \quad (45)$$

$$Z(\lambda)^{1/2} = \lambda^{-(d-y_H)} = \lambda^{-(d-2+\eta)/2}$$

Here, $Z(\lambda)$ is the wave renormalization factor and should be distinguished from the partition function. This is again an example of conflicting traditions in what used to be distinct theoretical areas! In the case of a free field, we showed in chapter 1 that the continuous limit allows one to assign the dimension $(d-2)/2$ to the field in terms of wavenumbers (inverse length). The interpretation of the relation (46) is that the corresponding continuous field $\varphi(\mathbf{x})$ has an effective dimension

$$d_\varphi = d - y_H = \frac{d-2}{2} + \frac{\eta}{2} \quad (46)$$

The deviation from the canonical dimension $(d-2)/2$ is $\eta/2$, the so-called *abnormal dimension of the field* φ . The first equation (45) expresses that

$$\sum_i H_i \sigma_i \rightarrow \int H(x) \varphi(x) d^d x$$

is dimensionless; i.e. that the dimension of $H(x)$ is y_H .

4.1.3 Examples and approximations

We shall return later to effective numerical methods, and present here some rather drastic approximations designed to illustrate the renormalization transformations. We begin by a trivial case in which the decimation method (integration over a subset of variables) is possible and leads to the exact result.

One-dimensional decimation We take this opportunity to introduce the Potts model which generalizes the Ising model. Dynamical variables describe now q distinct states at each site. The nearest neighbour interaction energy assumes two possible values, distinguishing the case when the two states are identical (favoured case) from the one when they differ. The Ising model corresponds to $q = 2$. For a one-dimensional lattice of size L , the partition function reads

$$Z_L = \frac{1}{q^L} \sum_{\{\sigma\}} \exp \left(\beta \sum_{n=1}^L (2\delta_{\sigma_n, \sigma_{n+1}} - 1) \right) \quad (47)$$

where the Boltzmann weight is e^β if the neighbouring spins are equal, $e^{-\beta}$ if they differ. The partition function has been normalized to unity for $\beta = 0$. The global symmetry of the system is the permutation group π_q of q objects (with $\pi_2 = Z_2$), and the set of pure states can be identified with the elements of the cyclic group Z_q .

Each factor $\exp \beta(2\delta_{\sigma\sigma'} - 1)$ is expanded as

$$e^{\beta(2\delta_{\sigma\sigma'} - 1)} = \left(\frac{e^\beta + (q-1)e^{-\beta}}{q} \right) \times 1 + \left(\frac{e^\beta - e^{-\beta}}{q} \right) \times (q\delta_{\sigma\sigma'} - 1) \quad (48)$$

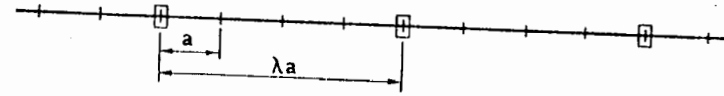


Fig. 3 Decimation on a one-dimensional lattice.

in the basis 1 and $q\delta_{\sigma\sigma'} - 1$. These matrices (with indices σ and σ') are orthonormal, in the following sense

$$\frac{1}{q} \sum_{\mu} [a_1 + b_1(q\delta_{\sigma\mu} - 1)][a_2 + b_2(q\delta_{\mu\sigma'} - 1)] = a_1 a_2 + b_1 b_2 (q\delta_{\sigma\sigma'} - 1) \quad (49)$$

The partition function appears as the trace of the matrix (48) to the power L , and the free energy is thus given by

$$F = \lim_{L \rightarrow \infty} \frac{\ln Z_L}{L} = \ln \frac{e^\beta + (q-1)e^{-\beta}}{q} \quad (50)$$

Of course, there is no transition. The known result $F = \ln \cosh \beta$ is recovered for $q = 2$.

In the decimation method (figure 3), one considers blocks of size λa , a being the lattice spacing and λ an integer. The spin variables located at sites with a coordinate which is a multiple of λa are kept fixed, while we sum over intermediate ones using formula (49). Generalizing the notation $t = \tanh \beta$ of the Ising model through

$$t = \frac{1 - e^{-2\beta}}{1 + (q-1)e^{-2\beta}} \quad (51)$$

one immediately obtains

$$t_\lambda \equiv t(\beta_\lambda) = t(\beta)^\lambda \quad (52)$$

This is the *exact* renormalization equation for this very simple case in which the decimation does not modify the form of the interaction. To be more precise, equation (52) is complemented by the transformation

$$\frac{Z_L(\beta)}{\left(\frac{e^\beta + (q-1)e^{-\beta}}{q} \right)^L} = \frac{Z_{L/\lambda}(\beta_\lambda)}{\left(\frac{e^{\beta_\lambda} + (q-1)e^{-\beta_\lambda}}{q} \right)^{L/\lambda}} \quad (53)$$

The fixed points of the transformation (52) are $t = 0$ ($\beta = 0$), attractive high temperature point, and $t = 1$ ($\beta = \infty$), repulsive low temperature fixed point which plays the role of the critical point. In the present case, we have a natural continuation in the variable λ . Hence, one can obtain the behaviour of the correlation length at low temperature. Indeed, near $t = 1$, $1 - t$ plays the role of θ in the general discussion, with

$$1 - t_\lambda \sim \lambda(1 - t) \quad (54)$$

and one deduces the expected result

$$\xi \sim \frac{1}{1 - t} \sim \frac{1}{q} e^{2\beta} \quad (55)$$

In higher dimension, the decimation procedure generates interactions beyond nearest neighbours. Truncations are then necessary, and it is rather difficult to assert their validity. Moreover, there is some difficulty in interpreting this type of transformation due to the absence of wavefunction renormalization.

It is then interesting to develop other approximation schemes in which the calculations are also rather simple. One of those has been proposed by Migdal and interpreted by Kadanoff.

Link displacement Consider the initial partition function

$$Z = \sum_{\{\sigma\}} e^{S(\sigma)} \quad (56)$$

Let $\Delta(\sigma)$ be a functional of the spins, with a vanishing mean value

$$\langle \Delta(\sigma) \rangle = Z^{-1} \sum_{\{\sigma\}} \Delta(\sigma) e^{S(\sigma)} = 0 \quad (57)$$

Due to the convexity inequality

$$Z_\Delta = \sum_{\{\sigma\}} e^{S(\sigma) + \Delta(\sigma)} = Z \langle e^{\Delta(\sigma)} \rangle \geq Z e^{\langle \Delta(\sigma) \rangle} = Z \quad (58)$$

The approximation replaces the action $S(\sigma)$ by a modified one $S(\sigma) + \Delta(\sigma)$, with a choice of $\Delta(\sigma)$ so as to generate recurrence relations. We analyze again the Potts model, but now in arbitrary dimension. The lattice is hypercubic. We introduce a transformation which treats separately the different directions, and therefore we allow distinct couplings $\beta_1, \dots, \beta_d$, for the interactions between

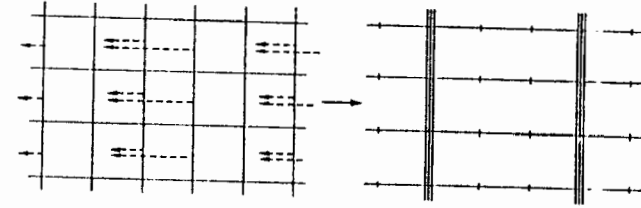


Fig. 4 Link displacement.

nearest neighbours on a link parallel to the direction $1, \dots, d$. The action reads

$$S(\sigma) = \sum_{\mathbf{x}, \mu} \beta_\mu (2\delta_{\sigma_{\mathbf{x}}, \sigma_{\mathbf{x}+\hat{\mu}}} - 1) \quad (59)$$

Let us now look for a perturbation Δ satisfying the condition (57) and such that the new action $S + \Delta$ can be interpreted as an action similar to S on a lattice with larger spacing. The approximation will then replace the r.h.s. of (58) by the l.h.s.. Due to translational invariance, the variables attached to two parallel links verify

$$\langle (2\delta_{\sigma_{\mathbf{x}}, \sigma_{\mathbf{x}+\hat{\mu}}} - 1) - (2\delta_{\sigma_{\mathbf{y}}, \sigma_{\mathbf{y}+\hat{\mu}}} - 1) \rangle = 0 \quad (60)$$

This equality allows one to displace parallel links. Choosing a direction μ , we displace links parallel to all directions $\nu \neq \mu$ in order to change the lattice spacing a_μ in direction μ to λa_μ (see figure 4). Links in any direction different from μ are replicated λ times. In the direction μ , it is now possible to use the technique of one-dimensional decimation previously described. With the same notation as in equations (51)–(52), the result of this transformation reads

$$\begin{cases} t(\beta'_\mu) = t(\beta_\mu)^\lambda \\ \beta'_\nu = \lambda \beta_\nu \quad \nu \neq \mu \end{cases} \quad (61)$$

This operation is performed consecutively for the different directions $\mu = 1, 2, \dots, d$. After this process, the lattice is again cubic with spacing λa , and the couplings β_μ are replaced by $\beta_\mu(\lambda)$ according to the Migdal-Kadanoff recurrence formulae

$$t(\lambda^{\mu-d} \beta_\mu(\lambda)) = [t(\lambda^{\mu-1} \beta_\mu)]^\lambda \quad 1 \leq \mu \leq d \quad (62)$$

This asymmetric treatment justifies the choice of distinct couplings in each direction. However, the transformation does not mix the different couplings. It is possible to imagine alternate procedures which restore the symmetry between directions. In any case, equation (62) is only approximate, except in one dimension, since we identify the left- and right-hand sides of inequality (58).

Write, in the same approximation, the multiplicative factor of the partition function.

Let $\{\beta_\mu^*\}$ be a nontrivial fixed point. The first component satisfies

$$t(\lambda^{1-d}\beta_1^*) = [t(\beta_1^*)]^\lambda \quad (63)$$

and from equation (62), one gets

$$\beta_\mu^* = \lambda^{-(\mu-1)}\beta_1^* \quad (64)$$

These critical values coincide when λ is analytically continued up to 1. Duality affords an interesting interpretation in two dimensions. We observe that the Potts model is self-dual, as is the Ising model. In the anisotropic case, the duality transformation reads

$$(\beta_1, \beta_2) \leftrightarrow (\tilde{\beta}_1, \tilde{\beta}_2) \quad (65)$$

$$e^{-2\tilde{\beta}_1} = t(\beta_2)$$

$$e^{-2\tilde{\beta}_2} = t(\beta_1) \quad (66)$$

The invariant curve corresponds to $\exp(-2\beta_1) = t(\beta_2)$ and is given by

$$(q-1)e^{-2\beta_1}e^{-2\beta_2} + e^{-2\beta_1} + e^{-2\beta_2} - 1 = 0 \quad (67)$$

Thus in two dimensions, for any choice of λ , the Migdal-Kadanoff fixed point lies on the self-dual line. The expressions (66)-(67) generalize those obtained in chapter 2. In chapter 6, we shall present a general study of duality. In particular, the critical isotropic point $\beta_1 = \beta_2 = \beta$ satisfies

$$e^{-2\beta} = t(\beta) = \frac{\sqrt{q}-1}{q-1} \quad (68)$$

Prove the self-duality of the 2-d Potts model and check that the fixed point given by equations (63)-(64) lies on the curve (67).

In terms of the variables y_1 and y_2 defined as

$$y_1 = e^{-2\beta_1}$$

$$y_2 = t(\beta_2) = \frac{1 - e^{-2\beta_2}}{1 + (q-1)e^{-2\beta_2}} \quad (69)$$

the Migdal-Kadanoff transformation reads

$$y_\mu(\lambda) = F_\lambda(y_\mu) \quad (\mu = 1, 2)$$

$$F_\lambda(y) = \left\{ \frac{[1 + (q-1)y]^\lambda - [1-y]^\lambda}{[1 + (q-1)y]^\lambda + (q-1)[1-y]^\lambda} \right\}^\lambda \quad (70)$$

The function $F_\lambda(y)$ is rational when λ is an integer, with $F_1(y) = y$. In the Ising case ($q = 2$) and for $\lambda = 2$, it reads

$$F_2(y) = \left(\frac{2y}{1+y^2} \right)^2 \quad (71)$$

We illustrate in figure 5 the typical instability characteristic of the critical point given by

$$y_1 = y_2 = y^* = F_\lambda(y^*) \quad (72)$$

This point depends on λ , but lies on the self-dual curve $y_1 = y_2$, as does the exact critical point. The transformation is at first meaningful only when λ is an integer. If, however, we perform a natural continuation for $\lambda \rightarrow 1$, we find the isotropic critical point

$$y_1 = y_2 = \frac{1 - y_1}{1 + (q-1)y_1} = \frac{\sqrt{q}-1}{q-1} \quad (73)$$

The critical exponents are rather unrealistic, and depend on the choice of λ . Near the fixed point, one can linearize the transformation as

$$(y_\lambda - y^*) \simeq F'_\lambda(y^*)(y - y^*) \quad (74)$$

According to the definitions (27-28), we can relate the coefficient in this relation to the exponent ν through

$$1 < F'_\lambda(y^*) = \lambda^{1/\nu}$$

$$\nu = \frac{\ln \lambda}{\ln F'_\lambda(y^*)} \quad (75)$$

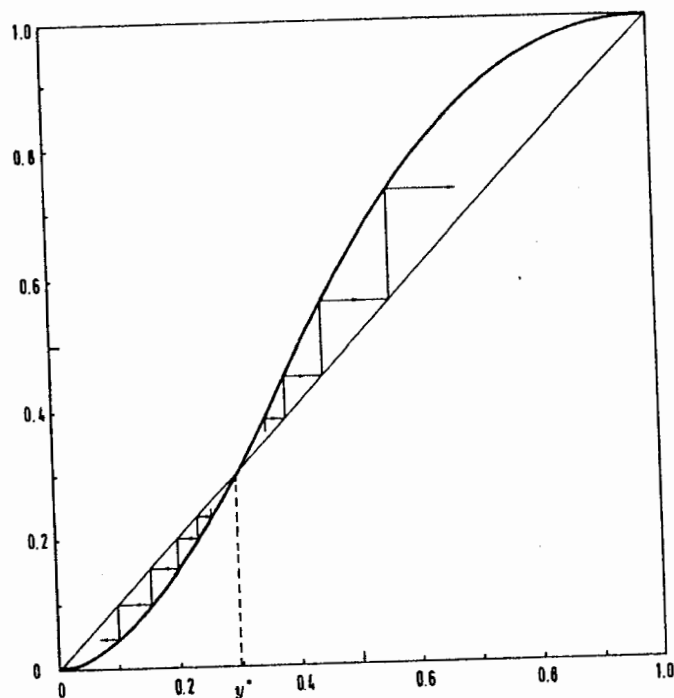


Fig. 5 The transformation $y \rightarrow F_2(y) = [2y/(1+y^2)]^2$ and its unstable fixed point $y^* \simeq 0.2956$, $F_2'(y^*) \simeq 1.6786$.

Let us compare this value with the one pertaining to the two-dimensional Ising model for which $\nu = 1$. For $q = 2$ and $\lambda = 2$, one finds

$$y^* = 0.2956 \quad F_2'(y^*) = 1.6786 \quad \nu = 1.3383 \quad (76)$$

This rather poor result does not improve in the limit $\lambda \rightarrow 1$. In this case, y^* tends to the familiar value

$$y^* = \sqrt{2} - 1 = 0.4142 \quad (77)$$

and one has

$$\begin{aligned} \nu^{-1} &= \frac{d}{d\beta} (\beta + \sinh \beta \cosh \beta \ln \tanh \beta) \Big|_{\beta^*} = 2 + \cosh 2\beta^* \ln \tanh \beta^* \\ &= 2 + \sqrt{2} \ln(\sqrt{2} - 1) \rightarrow \nu = 1.327 \end{aligned} \quad (78)$$

In all these decimation methods, the block variables are among the original ones and hence $G(\mathbf{x}, \beta) = G(\mathbf{x}/\lambda, \beta_\lambda)$ for the correlations. This implies

$$d - y_H = \beta/\nu = 0 \quad (79)$$

In spite of these shortcomings, the Migdal-Kadanoff approximation is nevertheless interesting. One may hope to construct systematic improvements due to its variational character. It is exact at the lower critical dimension ($d = 1$) and the general idea can be applied to more complex systems such as gauge models.

Majority rule As an illustration of the majority rule, we consider the Ising model on a two-dimensional triangular lattice, following Niemeijer and Van Leeuwen. Each site has six neighbours. Using the methods developed in chapter 2, the exact free energy per site, normalized to zero at low temperature, is found to be

$$\begin{aligned} F(\beta) &= \frac{1}{2} \int_{-\pi}^{\pi} \frac{d^2 \theta}{(2\pi)^2} \ln[(\cosh 2\beta)^3 + (\sinh 2\beta)^3 \\ &\quad - \sinh 2\beta(\cos \theta_1 + \cos \theta_2 + \cos(\theta_1 + \theta_2))] \end{aligned} \quad (80)$$

The critical point is

$$\beta_c = \frac{1}{4} \ln 3 = 0.2747 \quad (81)$$

and the critical exponents are the same as those on a square lattice.

Prove these results.

The triangular lattice is not self-dual. The duality transformation leads to a similar model on a honeycomb lattice where each site has three neighbours, with a new coupling β' such that

$$e^{-2\beta'} = \tanh \beta \quad (82)$$

An exact decimation procedure called the *star-triangle* transformation, allows one to recover a triangular lattice. Let us split

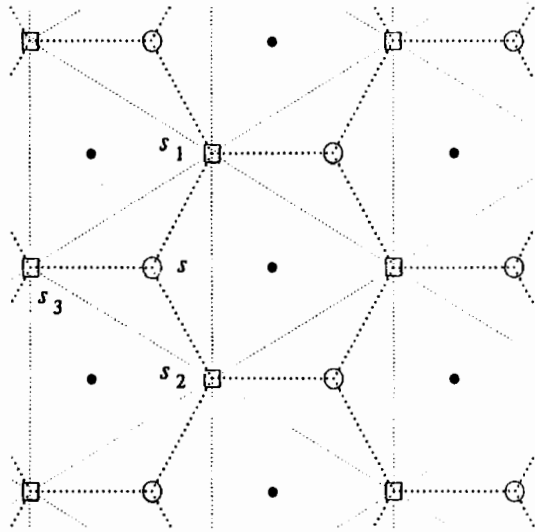


Fig. 6 The star-triangle transformation. The original sites appear as black dots. The dual sites on the honeycomb lattice are split into two families, represented by white circles and squares.

the dual honeycomb lattice, with twice as many sites as the original one, into two triangular sublattices. Now we can sum over the values of the spins s located on the first sublattice, which only interact with three neighbouring spins s_1, s_2, s_3 located on the second (figure 6)

$$\begin{cases} \frac{1}{2} \sum_{s=\pm 1} \exp[\beta' s(s_1 + s_2 + s_3)] = A \exp[\beta''(s_1 s_2 + s_2 s_3 + s_3 s_1)] \\ A^4 = \cosh 3\beta' (\cosh \beta')^3 \\ e^{4\beta''} = \frac{\cosh 3\beta'}{\cosh \beta'} = 2 \cosh 2\beta' - 1 = 2 \frac{\cosh 2\beta}{\sinh 2\beta} - 1 \end{cases} \quad (83)$$

The last relation may be rewritten in a more symmetric way

$$(e^{4\beta''} - 1)(e^{4\beta} - 1) = 4 \quad (84)$$

We have then a correspondence between two isomorphic models with respective couplings β and β'' . The critical point (if it

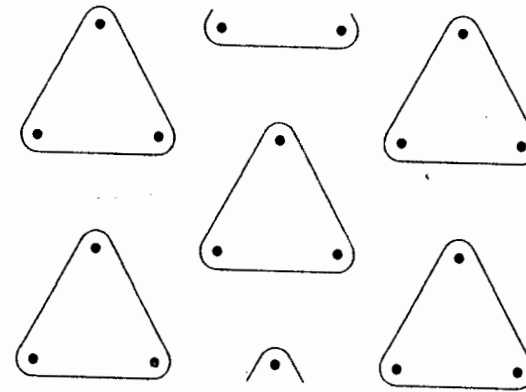


Fig. 7 Blocks of three spins.

exists and if it is unique, as is the case) is the fixed point of the transformation, and agrees with (81).

- (1) Write the corresponding relation between the free energies and check that it is satisfied by the expression (80).
- (2) Can the star-triangle transformation be generalized to the Potts model?

Let us now apply a renormalization transformation, according to the majority rule, with blocks containing three sites (an odd number which avoids ambiguities), as shown in figure 7. Larger blocks could of course also be considered. In this case, as the number of sites is divided by 3, the scaling factor is $\lambda = \sqrt{3}$. To each block is assigned the new dichotomic variable

$$\mu = \text{Sign}(\sigma_1 + \sigma_2 + \sigma_3) \quad (85)$$

The transformation is defined by assuming that the partition function is invariant. Upon inserting a product of factors $1 = \sum_{\mu} \delta_{\mu, \text{Sign}(\Sigma \sigma)}$, we write

$$Z = \sum_{\{\mu_a = \pm 1\}} e^{S'(\mu)} = \sum_{\{\mu_a = \pm 1\}} \sum_{\{\sigma_i = \pm 1\}} \left(\prod_a \delta_{\mu_a, \text{Sign}(\sigma_1^a + \sigma_2^a + \sigma_3^a)} \right) e^{S(\sigma)} \quad (86)$$

For fixed μ 's there are four configurations of the spins σ_i inside each block

$$\begin{cases} \sigma_1 = \sigma_2 = \sigma_3 = \mu \\ \sigma_1 = \sigma_2 = -\sigma_3 = \mu \\ \sigma_1 = -\sigma_2 = \sigma_3 = \mu \\ -\sigma_1 = \sigma_2 = \sigma_3 = \mu \end{cases} \quad (87)$$

The initial action is then split into two terms S_0 and S_1 . The first part S_0 contains the contribution of interactions between spins of the same block. The second part S_1 arises from interactions between spins of different blocks and is treated perturbatively. Introducing the notation

$$Z_0 = \sum_{\sigma} \prod_a \delta_{\mu_a, \text{Sign}(\sigma_1^a + \sigma_2^a + \sigma_3^a)} e^{S_0(\sigma)} = z^{N/3} \quad (88)$$

$$z = e^{3\beta} + 3e^{-\beta}$$

and

$$\langle A(\sigma) \rangle = Z_0^{-1} \sum_{\sigma} \prod_a \delta_{\mu_a, \text{Sign}(\sigma_1^a + \sigma_2^a + \sigma_3^a)} e^{S_0(\sigma)} A(\sigma) \quad (89)$$

one wants to compute

$$e^{S'(\mu)} = Z_0 \langle \exp S_1 \rangle \quad (90)$$

In the cumulant expansion of this mean value

$$\ln \langle \exp S_1 \rangle = \langle S_1 \rangle + \frac{1}{2} (\langle S_1^2 \rangle - \langle S_1 \rangle^2) + \dots \quad (91)$$

we only retain the first term. In this approximation, the block variables have only nearest neighbour interactions, as do the initial spins. For a given pair (a, b) , one finds the term

$$\beta \langle (\sigma_1^a + \sigma_2^a) \sigma_3^b \rangle = \beta (\langle \sigma_1^a \rangle + \langle \sigma_2^a \rangle) \langle \sigma_3^b \rangle$$

Since

$$\langle \sigma_i^a \rangle = \mu_a \frac{e^{3\beta} + e^{-\beta}}{e^{3\beta} + 3e^{-\beta}}$$

the renormalized coupling reads

$$\beta_\lambda = 2\beta \left(\frac{e^{3\beta} + e^{-\beta}}{e^{3\beta} + 3e^{-\beta}} \right)^2 \quad (\lambda = \sqrt{3}) \quad (92)$$

Higher order terms in (91) would generate longer-range interactions. Within the present approximation, the method reduces to

a coupling constant renormalization (92). Its fixed point is given by

$$\beta^* = \frac{1}{4} \ln(1 + 2\sqrt{2}) = 0.3356 \quad (93)$$

to be compared to (81). Near this fixed point, the expansion

$$(\beta_\lambda - \beta^*) = (\beta - \beta^*)[1 + 2\beta^*(8 - 5\sqrt{2})] + \dots$$

yields the critical exponent

$$(\sqrt{3})^{1/\nu} = 1 + 2\beta^*(8 - 5\sqrt{2}) = 1.6235 \quad \rightarrow \quad \nu = 1.1335 \quad (94)$$

This value is not very far from the exact one $\nu = 1$. One may hope that the approximation improves when further terms are kept in the expansion (91). Of course, one should then start from an action involving longer-range interactions with additional couplings. In any case, it is rather difficult to find the appropriate truncation without a thorough numerical study.

Sticking to this simple framework, let us analyze the effect of an external field H . A term $H \sum_i s_i$ is added to the action which is again split into S_0 and S_1 . Within each block, the contribution to the partition function is given by

$$z_H e^{H_\lambda^{(0)} \mu} = \sum_{\{\sigma_i = \pm 1\}} e^{\beta(\sigma_1 \sigma_2 + \sigma_2 \sigma_3 + \sigma_3 \sigma_1) + H(\sigma_1 + \sigma_2 + \sigma_3)} \delta_{\mu, \text{Sign}(\sigma_1 + \sigma_2 + \sigma_3)}$$

$$\rightarrow \begin{cases} z_H^2 = (e^{3H+3\beta} + 3e^{H-\beta}) (e^{-3H+3\beta} + 3e^{-H-\beta}) \\ e^{2H_\lambda^{(0)} \mu} = (e^{3H+3\beta} + 3e^{H-\beta}) / (e^{-3H+3\beta} + 3e^{-H-\beta}) \end{cases} \quad (95a)$$

with

$$\langle \sigma \rangle = \alpha_1 \mu + \alpha_2$$

$$\alpha_1 \pm \alpha_2 = \frac{e^{\pm 3H+3\beta} + e^{\pm H-\beta}}{e^{\pm 3H+3\beta} + 3e^{\pm H-\beta}} \quad (95b)$$

The relations (88) and (90) are replaced by

$$e^{S'(\mu)} \cong (z_H)^{N/3} \exp \left\{ \sum_a H_\lambda^{(0)} \mu_a + 2\beta \sum_{(ab)} (\alpha_1 \mu_a + \alpha_2) (\alpha_1 \mu_b + \alpha_2) \right\} \quad (96)$$

and the new couplings are

$$\begin{cases} \beta_\lambda = 2\beta \alpha_1^2 \\ H_\lambda = H_\lambda^{(0)} + H_\lambda^{(1)} \end{cases} \quad (97)$$

with $H_\lambda^{(0)}$ given by equation (95) and

$$H_\lambda^{(1)} = 12\beta\alpha_1\alpha_2 \quad (98)$$

The critical point corresponds to $\beta = \beta^*$ given by (93), and $H = 0$. Linearizing the transformation (which is already diagonal) in the vicinity of this point, leads to

$$\begin{cases} \beta_\lambda - \beta^* = \lambda^{1/\nu}(\beta - \beta^*) + \dots \\ H_\lambda = \lambda^{y_H} H + \dots \end{cases} \quad (99)$$

$$\begin{cases} \lambda^{1/\nu} = 1 + 2\beta^*(8 - 5\sqrt{2}) = 1.6235 \\ \lambda^{y_H} = \frac{3}{\sqrt{2}} + 3\beta^*(8 - 5\sqrt{2}) = 2.123 + 0.9353 = 3.0566 \end{cases} \quad (100)$$

We have distinguished in the last expression two contributions of respective orders 0 and 1. Using the value $\lambda = \sqrt{3}$, this yields the exponents

$$\begin{cases} y_H^{(0)} = 1.369 \\ y_H^{(1)} = 2.034 \end{cases} \quad (101)$$

In the present case $\nu = 1, \gamma = \frac{7}{4}$ correspond to the exact value $y_H = \frac{15}{8} = 1.875$. The result (101) is thus not so bad. Niemeijer and van Leeuwen, Braathen and Hemmer have performed the computation up to second order, including interactions between second and third neighbours. The coupling space is now three-dimensional, and $\lambda^{1/\nu}$ is identified with the largest eigenvalue of the linearized approximation near the fixed point. We only display the results

order	$y_0 = 1/\nu$	y_H	β^*
0		1.369	
1	0.822	2.034	0.3356
2	1.053	1.847	0.2514
exact	1.000	1.875	0.2747

(102)

which show a spectacular improvement. Of course, the preceding discussion should only be considered as a short introduction to numerous works which cannot be listed here.

Critical amplitudes Up to now we focused our attention on the calculation of the critical exponents. It is also interesting to compute the amplitude of singular terms. We limit ourselves to the simple example of a recurrence relation, with a fixed dilatation coefficient λ , of the type (26). This is rewritten as

$$F(\theta) = f(\theta) + \lambda^{-d} F(\theta_\lambda) \quad (103)$$

Near the critical point, θ measures the deviation from the critical temperature and θ_λ is its transform. The function $f(\theta)$ is assumed to be regular at $\theta = 0$. We work in the high temperature phase ($\theta > 0$) and normalize the free energy to $F(\infty) = f(\infty) = 0$. Near the origin, the fixed point of the map $\theta \rightarrow \theta_\lambda$, the linearized transformation is

$$\theta_\lambda = \mu_\lambda \theta + \mathcal{O}(\theta^2) \quad (104)$$

and we know that the critical exponent α is given by $\mu_\lambda^{2-\alpha} = \lambda^d$. The factor μ_λ is larger than unity, expressing the instability of the fixed point. Let n be an integer such that $n < 2 - \alpha < n + 1$ (for the time being, the inequalities are assumed to be strict). Let us iterate the transformation $\theta \rightarrow \theta_\lambda \rightarrow \theta_{\lambda^2} \rightarrow \dots \rightarrow \theta_{\lambda^n} \rightarrow \dots$. One obtains

$$F(\theta) = \sum_{k=0}^{p-1} \lambda^{-dk} f(\theta_{\lambda^k}) + \lambda^{-dp} F(\theta_{\lambda^p}) \quad (105)$$

As $p \rightarrow \infty$, θ_{λ^p} tends towards the high temperature stable point. The series converges and hence

$$F(\theta) = \sum_{k=0}^{\infty} \lambda^{-dk} f(\theta_{\lambda^k}) \quad (106)$$

The knowledge of an exact renormalization relation of the type (103) would allow the complete determination of the free energy. The same reasoning can also be applied in the low temperature phase (changing the normalization of F in order to ensure convergence). These hypotheses are of course very strong since we assume an exact renormalization formula. Let us nevertheless proceed within this framework. The estimation of (106) is rather difficult without numerical methods. As an approximation, one can replace the transformation $\theta \rightarrow \theta_\lambda$ by its linearized form. Better, we substitute a new variable $\tilde{\theta}(\theta)$ such that the renormalization transformation reads $\tilde{\theta}_\lambda \equiv \tilde{\theta}(\theta_\lambda) = \mu_\lambda \tilde{\theta}$. This

change of variable is always possible locally. We shall not discuss here the delicate question of its global domain of validity. In order to simplify the notation, we suppress the tilde on this variable and write

$$F(\theta) = \sum_{p=0}^{\infty} \lambda^{-dp} f(\mu_{\lambda}^p \theta) \quad (107)$$

In particular, we get $F(0) = f(0)/(1 - \lambda^{-d})$. As f is regular at $\theta = 0$, it can be expanded in a Taylor series as $f(\theta) = \sum_{k=0}^{\infty} f_k \theta^k$. Naively, one could think of inserting this expansion in (107) and inverting the order of the two summations. However, this is incorrect, since the summation over p does not converge when $k > n = [2 - \alpha]$ ($\mu_{\lambda}^k > \lambda^d$). Nevertheless, let us introduce the regular function

$$\varphi(\theta) = \sum_{k=0}^{\infty} f_k \frac{\theta^k}{1 - \mu_{\lambda}^k / \lambda^d} \quad (108)$$

Each term in this series is well-defined, since we exclude the case where $2 - \alpha$ is an integer. Moreover, if the radius of convergence of $f(\theta)$ is R , the expansion of φ converges in a disk of radius $\mu_{\lambda} R > R$. Let us now split f into a subtraction term $\sum_0^n f_k \theta^k$ and a remainder $f_1(\theta)$, so that the subtraction term leads to a convergent summation over p . Hence

$$F(\theta) = \sum_{k=0}^n f_k \frac{\theta^k}{1 - \mu_{\lambda}^k / \lambda^d} + \sum_{p=0}^{\infty} \lambda^{-dp} f_1(\mu_{\lambda}^p \theta) \quad (109)$$

In the last sum, each term is of order θ^{n+1} although we know that the sum has a singular behaviour in $\theta^{2-\alpha}$. In order to show this, we add and subtract in (109) the absolutely convergent term $\sum_{p=-\infty}^{-1} \lambda^{-dp} f_1(\mu_{\lambda}^p \theta)$. Substituting for $f_1(\theta)$ its Taylor expansion $\sum_{n+1}^{\infty} f_k \theta^k$, one obtains

$$\sum_{p=-\infty}^{-1} \lambda^{-dp} \sum_{k=n+1}^{\infty} f_k \mu_{\lambda}^{kp} \theta^k = - \sum_{k=n+1}^{\infty} f_k \frac{\theta^k}{1 - \mu_{\lambda}^k / \lambda^d}$$

the convergence being ensured by the fact that $\mu_{\lambda}^k > \lambda^d$ for $k \geq n+1$. Under these conditions,

$$F(\theta) = \varphi(\theta) + \sum_{p=-\infty}^{\infty} \lambda^{-dp} f_1(\mu_{\lambda}^p \theta) \quad (110)$$

This formula gives an explicit representation of F as the sum of a regular part φ , and a singular one which satisfies the homogeneous equation

$$F_{\text{sing}}(\theta) = \sum_{p=-\infty}^{\infty} \lambda^{-dp} f_1(\mu_{\lambda}^p \theta) = \lambda^{-d} F_{\text{sing}}(\mu_{\lambda} \theta) \quad (111)$$

and which behaves as $\theta^{2-\alpha}$. Hence

$$F_{\text{sing}}(\theta) = \theta^{2-\alpha} A(\theta) \quad (112)$$

$$A(\theta) = A(\mu_{\lambda} \theta) = \theta^{\alpha-2} \sum_{p=-\infty}^{\infty} \lambda^{-dp} f_1(\mu_{\lambda}^p \theta)$$

For a fixed renormalization factor λ , the amplitude A depends on θ and is in fact a periodic, bounded function of $\ln \theta$ with period $\ln \mu_{\lambda}$. Writing

$$A(\theta) = \sum_{s=-\infty}^{+\infty} A_s \exp\left(2i\pi s \frac{\ln \theta}{\ln \mu_{\lambda}}\right) \quad (113)$$

we have for the coefficients

$$A_s = \frac{1}{\ln \mu_{\lambda}} \int_1^{\mu_{\lambda}} \frac{d\theta}{\theta^{3-\alpha+2i\pi s/\ln \mu_{\lambda}}} \sum_{p=-\infty}^{\infty} \lambda^{-dp} f_1(\mu_{\lambda}^p \theta) \quad (114)$$

$$= \frac{1}{\ln \mu_{\lambda}} \int_0^{\infty} d\theta \theta^{\alpha-3-2i\pi s/\ln \mu_{\lambda}} f_1(\theta)$$

The last expression has been derived using the relation $\mu_{\lambda}^{2-\alpha} = \lambda^d$, so that each term of the sum is written as an integral over the interval $[\mu_{\lambda}^p, \mu_{\lambda}^{p+1}]$. Finally, the last integral in (114) can be integrated by parts until f_1 is replaced by the $(n+1)$ -th derivative of f . In general, (113) and (114) are only approximate, and one sees, on numerical examples, that the amplitude of the oscillations is very small, so that it is possible to replace $A(\theta)$ by A_0 . Of course, similar expressions can be deduced for the low temperature region.

When $2 - \alpha$ is an integer, the singularity does not disappear, but yields a logarithmic term. By continuity in $\varepsilon = n - 2 + \alpha \rightarrow 0$, one of the terms in φ becomes singular and one gets

$$F(\theta) = F_{\text{reg}}(\theta) - \frac{f_n \theta^n}{\varepsilon \ln \mu_{\lambda}} + \frac{\theta^n (1 - \varepsilon \ln \theta)}{\ln \mu_{\lambda}} \int_0^{\infty} \frac{d\theta' f_1(\theta')}{\theta'^{1+n-\alpha}} \quad (115)$$

$$\simeq F_{\text{reg}}(\theta) - \frac{f_n \theta^n}{\varepsilon \ln \mu_{\lambda}} + \frac{\theta^n (1 - \varepsilon \ln \theta)}{\ln \mu_{\lambda}} f_n \int_0^{\theta_0} \frac{d\theta'}{\theta'^{1-\varepsilon}}$$

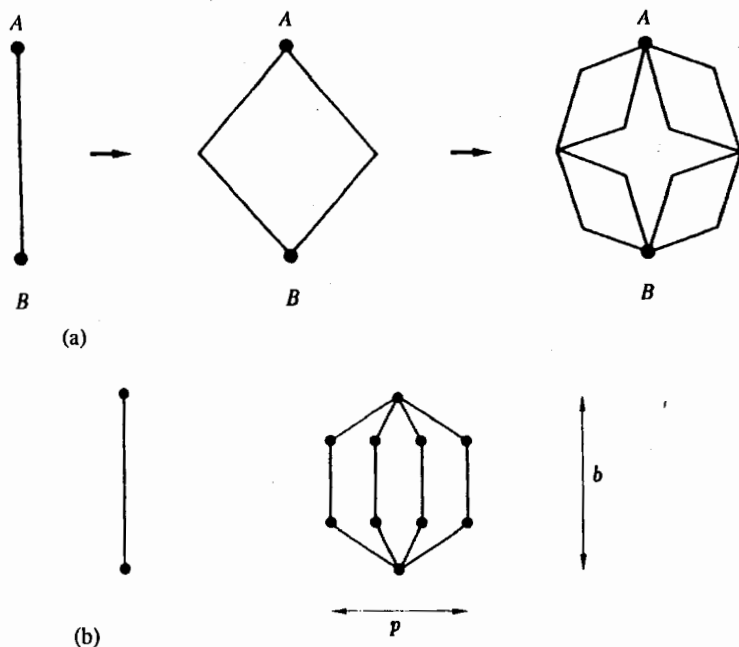


Fig. 8 (a) Recursive construction of a hierarchical lattice. (b) Generalization with p strings of b links. A dimension can be defined by comparing the increase in "length" $L' = bL$ between the two points AB to the increase in "volume", i.e. the number of bonds $N' = pbN$. If N scales like L^d , $d = \ln(pb)/\ln b$. Figure (a) corresponds to $d = 2$.

$$\simeq F_{\text{reg}}(\theta) - \frac{f_n}{\ln \mu_\lambda} \theta^n \ln |\theta|$$

The same reasoning can be applied for $\theta < 0$. Hence, while the amplitudes A_+ and A_- above and below the critical point seem to be unrelated when $2 - \alpha$ is not an integer, the logarithmic singularity has the same amplitude on both sides when $2 - \alpha$ is an integer. This is what has been observed in the two-dimensional Ising model.

Consider the *hierarchical* lattice obtained by the following recursive process. Starting from an elementary graph H_2 with two sites joined by one link, H_n is deduced from H_{n-1} by substituting four links for each one, according to the scheme depicted on figure 8a.

- i) Show the existence of a thermodynamical limit $n \rightarrow \infty$ for a q -states Potts model on H_n with a partition function

$$Z_n = \sum_{\sigma_i=1}^q \exp 2\beta \sum_{ij} \delta_{\sigma_i, \sigma_j} \quad (116)$$

- ii) Obtain by decimation a recurrence relation between Z_n and Z_{n-1} generalizing the Migdal-Kadanoff relation (71). Using $x = e^{2\beta}$ instead of $y = e^{-2\beta}$,

$$Z_n(x) = (2x + q - 2)^{2 \cdot 4^{n-2}} Z_{n-1}(x') \quad (117)$$

$$x' = T(x) = \left(\frac{x^2 + q - 1}{2x + q - 2} \right)^2$$

- iii) Study the critical exponents and amplitudes.
iv) To represent the behaviour of the partition function, one can plot the limiting set of its zeroes in the complex x -plane as $n \rightarrow \infty$. One obtains remarkable Julia sets indicating a very interesting structure as shown in figure 9. The linearization of the corresponding renormalization map near the fixed points, as described in the text above, yields very inaccurate amplitudes of the singular part of the free energy. The matter is discussed in the references.

4.2 The XY-model

Real space renormalization transformations meet spectacular success when applied to the XY-model in two dimensions. This is why we present it in this chapter. This model deals with two component classical spins (S_x, S_y) normalized to unit length $S_x^2 + S_y^2 = 1$, and interacting with their nearest neighbours. The corresponding global symmetry is $O(2)$. After some transformations, this model also describes a classical Coulomb gas, a set of fluctuating surfaces, or superfluid films. It undergoes a remarkable transition, such that the free energy and all its derivatives remain continuous (Kosterlitz and Thouless). There is no macroscopic spontaneous magnetization (Mermin and Wagner). The renormalization group arguments reveal the nature of the singularity, as can be checked by comparing with the exact solution of a related model.

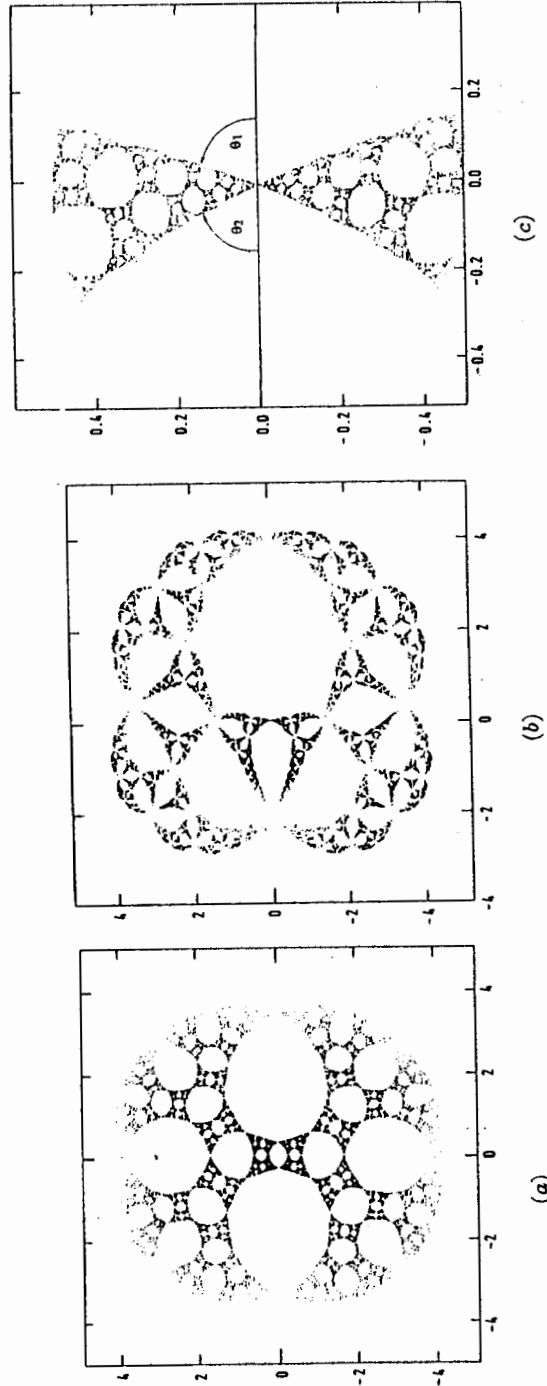


Fig. 9 (a) Julia set of zeros of the partition function in the $n \rightarrow \infty$ limit for the Ising case $q = 2$. The points shown in the figure are the images under the map T^{-1} of the unique zero of Z_1 . (b) Same as before for $q = 3$ with a bias designed to obtain a sharper view. (c) An enlargement of the Julia set for $q = 2$ near the ferromagnetic fixed point x_c showing two (equal) limiting angles θ_1 and θ_2 for the support of the zeroes.

Spin variables can also be described by angles, defined modulo 2π , with the partition function written as

$$Z = \int \prod_i \frac{d\theta_i}{2\pi} \exp \left(\beta \sum_{(ij)} \cos(\theta_i - \theta_j) \right) \quad (118)$$

The interaction is ferromagnetic, in the sense that the configurations with equal θ_i 's are favored. The global rotational symmetry appears here as a translation invariance $\theta_i \rightarrow \theta_i + \alpha$. We study this model on a two-dimensional square lattice.

4.2.1 High temperature behaviour

We have already considered this model in one dimension and shown that it is in a disordered phase at any temperature. One expects a similar disordered phase in two dimensions for sufficiently high temperature. This is confirmed by a strong coupling (high temperature) expansion (β small). Using a general procedure discussed in chapter 7, each factor is expanded on the basis of irreducible characters of the group $O(2)$, which amounts, in simpler terms, to the usual Fourier decomposition

$$\begin{aligned} e^{\beta \cos \theta} &= I_0(\beta) + \sum_{n=1}^{\infty} I_n(\beta) (e^{in\theta} + e^{-in\theta}) \\ &= I_0(\beta) \left(1 + \sum_{n \neq 0} t_n(\beta) e^{in\theta} \right) \end{aligned} \quad (119a)$$

where the coefficients are expressed in terms of modified Bessel functions

$$t_n(\beta) = t_{-n}(\beta) = \frac{I_n(\beta)}{I_0(\beta)} \quad (119b)$$

with the following behaviour for small or large β

$$t_n(\beta) \sim \begin{cases} \frac{\beta^n}{2^n n!} & \beta \rightarrow 0 \\ e^{-n^2/2\beta} & \beta \rightarrow \infty \end{cases} \quad (119c)$$

These coefficients generalize the quantity $\tanh \beta$ of the Ising model. They lie between 0 and 1 in the domain $0 \leq \beta < \infty$.

For a lattice with N sites and $2N$ links, one has

$$\frac{Z}{I_0^{2N}(\beta)} = \int \prod_i \frac{d\theta_i}{2\pi} \prod_{(ij)} \left(1 + \sum_{n_{ij} \neq 0} t_{n_{ij}}(\beta) e^{in_{ij}(\theta_i - \theta_j)} \right) \quad (120)$$

The interest of the expansion (119) is clear for small β , since the coefficients t_n decrease very rapidly with n and since integrations can be performed explicitly over each term in the expansion of the product (120). The result is

$$\frac{Z}{I_0^{2N}(\beta)} = \sum_{\substack{\{n_{ij}\} \\ (\partial n)_i = 0}} \prod_{(ij)} t_{n_{ij}}(\beta) \quad (121)$$

One sums over all configurations of relative integers n_{ij} , with the notations $t_0 \equiv 1$, $n_{ji} \equiv -n_{ij}$. These configurations satisfy the zero divergence condition $(\partial n)_i \equiv \sum_j n_{ij} = 0$. At high temperature, only finitely many links carry nonzero n_{ij} 's.

Similarly, one can compute the correlation function

$$\langle S_1 \cdot S_2 \rangle = \text{Re} \langle e^{i(\theta_1 - \theta_2)} \rangle \quad (122)$$

and more generally

$$\langle e^{im(\theta_1 - \theta_2)} \rangle = \frac{\sum_{\substack{\{n_{ij}\} \\ (\partial n)_i = m(\delta_{i1} - \delta_{i2})}} \prod_{(ij)} t_{n_{ij}}(\beta)}{\sum_{\substack{\{n_{ij}\} \\ (\partial n)_i = 0}} \prod_{(ij)} t_{n_{ij}}(\beta)} \quad (123)$$

The point 1 acts as a source and the point 2 as a sink of intensity m for the "field" n_{ij} . Analogous expressions are easily written for correlation functions involving more points.

For $m = 1$, the dominant term in the r.h.s. of equation (123) at high temperature is $N_{12} t_1(\beta)^{d_{12}}$, where d_{12} is the minimal distance on the lattice between points 1 and 2, and N_{12} is the number of corresponding paths of minimal length. Hence correlations decrease exponentially at high temperature to lowest order, and this extends to all orders because the high temperature series has a finite radius of convergence. The question arises whether this situation remains true at low temperature. In appendix A, we present the proof that a system with a continuous symmetry cannot have a nonvanishing mean value for the order parameter

in dimension two, in the absence of a symmetry breaking external field. This is to be contrasted with the case of a discrete symmetry. This phenomenon is due in essence to an excess of entropy over energy.

4.2.2 Low temperature expansion. Vortices.

At low enough temperature, the energetic term favours the alignment of spins, and one is tempted to expand $\cos(\theta_i - \theta_j)$ near $\theta_i - \theta_j = 0$. In this treatment, the periodic character of the angles is neglected, but the existence of an order, at least at short range, partly justifies this method. This point will be further studied later. The model is replaced by a so-called *spin wave* approximation. The partition function is the one of a pure Gaussian model

$$Z_{sw} = \int \prod_i \frac{d\theta_i}{(2\pi)^{1/2}} \exp -\frac{1}{2} \beta \sum_{(ij)} (\theta_i - \theta_j)^2 \quad (124)$$

where we have omitted a constant additive term in the action. The inverse propagator is the kernel of the quadratic form, i.e.

$$G^{-1}(\mathbf{x}, \mathbf{x}') = \int_{-\pi}^{\pi} \frac{d^2 \mathbf{q}}{(2\pi)^2} e^{i\mathbf{q} \cdot (\mathbf{x} - \mathbf{x}')} [4 - 2(\cos q_1 + \cos q_2)] \quad (125)$$

The rotational invariance of the spins, replaced here by the translation invariance of the angles, is associated with a zero mode. The corresponding infinite factor does not play a significant role in the thermodynamic limit; one can omit, for instance, the integration over one of the variables θ . The free energy per site associated to (124) is thus well-defined

$$F = \lim_{N \rightarrow \infty} \frac{1}{N} \ln Z_{sw} = -\frac{1}{2} \int_{-\pi}^{\pi} \frac{d^2 \mathbf{q}}{(2\pi)^2} \ln \{ \beta [4 - 2(\cos q_1 + \cos q_2)] \} \quad (126)$$

However, low energy fluctuation modes have a sufficiently high entropy to disorder the spins at large distances. This is the content of a theorem due to Hohenberg, Mermin and Wagner, and Coleman. Consider the expression of the correlation function

in the spin wave approximation

$$G_{sw}^{(p)}(x_1, x_2) = \langle \exp[ip(\theta_{x_1} - \theta_{x_2})] \rangle \quad (127)$$

$$= Z_{sw}^{-1} \int \prod \frac{d\theta_x}{2\pi} \exp \left(-\frac{1}{2}\beta \sum_{x, x'} \theta_x G_{x, x'}^{-1} \theta_{x'} + ip(\theta_{x_1} - \theta_{x_2}) \right)$$

which leads, after integration, to the expression

$$G_{sw}^{(p)}(x_1, x_2) = \exp \left(\frac{-p^2}{2\pi\beta} \Gamma(x_1 - x_2) \right) \quad (128)$$

The notation $\Gamma(x)$ denotes the integral

$$\Gamma(x) = \int_{-\pi}^{\pi} \frac{d^2 q}{2\pi} \frac{1 - e^{iq \cdot x}}{4 - 2(\cos q_1 + \cos q_2)} \quad (129)$$

Assuming an infrared regularization for the free field Green function $G(x)$, the well-defined quantity $\Gamma(x)$ is equal to the subtracted form $2\pi(G(0) - G(x))$. Observe that Γ appears in an exponential, since we compute mean values of exponentials.

The function $\Gamma(x)$ is well-defined for all x on a two-dimensional square lattice, vanishes for $x = 0$, is equal to $\pi/2$ for $|x| = 1$ and has an isotropic behaviour at large distances

$$\Gamma(x) \underset{|x|}{\sim} \ln(2\sqrt{2}e^\gamma |x|) + \mathcal{O}(1/|x|) \quad (130)$$

where γ is Euler's constant. The asymptotic behaviour given in equation (130) is an excellent approximation, since it gives $\Gamma(x) \simeq 1.6169$ for $|x| = 1$, to be compared with the exact value $\frac{1}{2}\pi \simeq 1.5708$.

To prove the estimate (130), it is sufficient to evaluate Γ for $x = (2p, 0)$. Hence

$$\begin{aligned} \Gamma(2p, 0) &= \int_{-\pi}^{\pi} \frac{dq_1}{2\pi} (1 - \cos 2pq_1) \int_{-\pi}^{\pi} dq_2 \frac{1}{4 - 2(\cos q_1 + \cos q_2)} \\ &= \frac{1}{2} \int_{-\pi}^{\pi} dq_1 \frac{1 - \cos 2pq_1}{\sqrt{(1 - \cos q_1)(3 - \cos q_1)}} \end{aligned}$$

This can be rewritten, with $q_1 = 2q$,

$$\Gamma(2p, 0) = 2 \int_0^{\frac{1}{2}\pi} dq \frac{\sin^2 2pq}{\sin q \sqrt{1 + \sin^2 q}} = \Gamma_0 + \Gamma_1$$

$$\begin{cases} \Gamma_0 = 2 \int_0^{\frac{1}{2}\pi} dq \frac{\sin^2 2pq}{\sin q} \\ \Gamma_1 = 2 \int_0^{\frac{1}{2}\pi} dq \frac{\sin^2 2pq}{\sin q} \left(\frac{1}{\sqrt{1 + \sin^2 q}} - 1 \right) \end{cases}$$

With the decomposition $\sin 2pq / \sin q = 2 \sum_{s=1}^p \cos(2s-1)q$, the first integral reads

$$\begin{aligned} \Gamma_0 &= 4 \int_0^{\frac{1}{2}\pi} dq \sin 2pq \sum_{s=1}^p \cos(2s-1)q \\ &= 2 \sum_{s=0}^{2p-1} \frac{1}{2s+1} = 2 \sum_{n=1}^{4p-1} \frac{1}{n} - \sum_{n=1}^{2p-1} \frac{1}{n} \\ &\sim 2 \ln(4p-1)e^\gamma - \ln(2p-1)e^\gamma + \mathcal{O}\left(\frac{1}{p}\right) \\ &\sim \ln 2p + 2 \ln 2 + \gamma + \mathcal{O}\left(\frac{1}{p}\right) \end{aligned}$$

To estimate Γ_1 , we can replace it by its limit when $p \rightarrow \infty$, taking $\sin^2 2pq$ for its average $\frac{1}{2}$, so that

$$\Gamma_1 \sim \int_0^{\frac{1}{2}\pi} dq \frac{1}{\sin q} \left(\frac{1}{\sqrt{1 + \sin^2 q}} - 1 \right)$$

Expanding now

$$(1 + \sin^2 q)^{-1/2} - 1 = \sum_{s=1}^{\infty} \frac{\sin^{2s} q}{s!} \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2} - s)}$$

one obtains

$$\begin{aligned} \Gamma_1 &\sim \sum_{s=1}^{\infty} \frac{\Gamma(\frac{1}{2})}{s! \Gamma(\frac{1}{2} - s)} \int_0^{\frac{1}{2}\pi} dq \sin^{2s-1} q = \sum_{s=1}^{\infty} \frac{\pi}{2s \Gamma(\frac{1}{2} + s) \Gamma(\frac{1}{2} - s)} \\ &= \frac{1}{2} \sum_{s=1}^{\infty} \frac{(-1)^s}{s} = -\frac{1}{2} \ln 2 \end{aligned}$$

Gathering these results, we find Γ up to order $1/p$ and obtain equation (130).

In terms of the lattice spacing, we define

$$r_0^{-1} = 2\sqrt{2}e^\gamma \quad (131)$$

so that $\Gamma(\kappa)$ is isotropic to a very good approximation

$$G_{sw}^{(p)}(\mathbf{x}) \underset{|\mathbf{x}| \gg 1}{\sim} \exp[-(p^2/2\pi\beta) \ln(|\mathbf{x}|/r_0)] = (r_0/|\mathbf{x}|)^{p^2/2\pi\beta} \quad (132)$$

The correlation function corresponds to the case $p = 1$. The spin wave approximation predicts a decrease according to a power law behaviour, qualitatively different from the high temperature exponential decrease. This suggests the existence of a transition at an intermediate temperature. The exponent characterizing the decrease of $G(\mathbf{x})$ varies with temperature and vanishes linearly at zero temperature. An isotropic power law behaviour is characteristic of a critical theory. We have thus here a whole continuous zone of critical temperatures. The underlying continuous field theory describes a free massless scalar field, with the subtractions which are implied by formulae (129)–(130) and which automatically arise when one computes invariant correlation functions. The result (132) shows the importance of large wavelength fluctuations which prevent ordering.

Fluctuations of the angles increase logarithmically at large distance. Hence the approximation neglecting their periodic character might become unjustified. Indeed, as Berezinskii, Kosterlitz and Thouless have shown, the topological effects resulting from this periodicity play an essential role. They combine with the spin wave fluctuations and increase disorder, so that correlation functions will acquire an exponential decrease beyond some critical temperature.

At a sufficiently large scale with respect to the lattice spacing, the microscopic structure of the latter can be neglected. Let us study the variation of the angle between the spin and a fixed direction as one follows a simple closed curve. Implicitly, fluctuations are supposed to be weak, so that this angle varies continuously along the path. If the spins are almost aligned, the variation of course vanishes. However, one can imagine configurations such that this angle varies by an integral multiple of 2π . This is a topological invariant, in the sense that it does not change when the curve or the configuration changes continuously. The simplest case corresponds to $n = \pm 1$ and describes a *vortex* of intensity $n = \pm 1$. For a continuous field of unit length, such a configuration should present at least one singular point. On a lattice, this concept is meaningless on a microscopic scale.

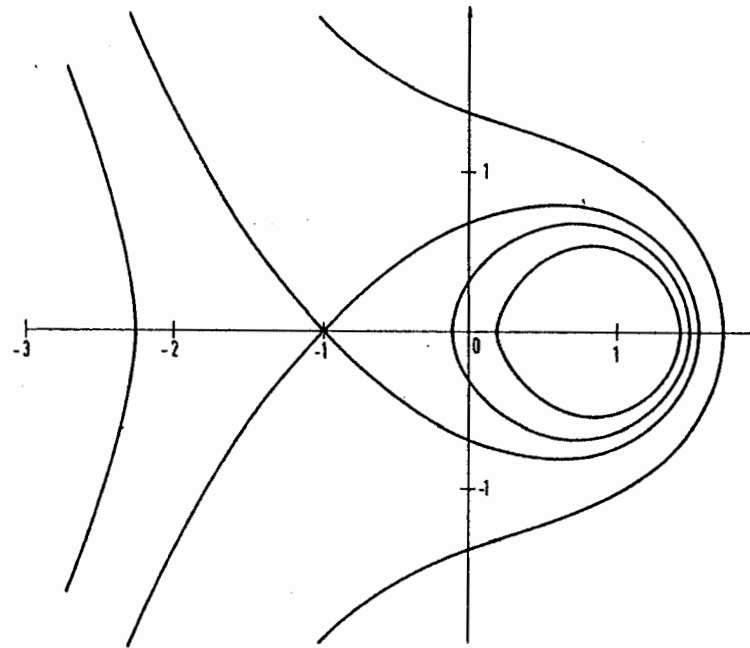


Fig. 10 A vortex pair of intensities ± 1 .

We assume now that all spins are aligned on the boundary of the system, so that the total vorticity (i.e. the sum of the n_i 's assigned to singular points) vanishes.

Figure 10 displays an example of two vortices with intensity ± 1 . Identifying the (x, y) plane with the complex plane $z = x + iy$, we choose $S_x + iS_y = e^{i\theta} = \exp(i \text{Arg}[i(z - z_0)/(z + z_0)])$. Lines are tangent to the vector field $\mathbf{S}(x, y)$, with $z_0 = 1$, and their equation is given by $y^2 = Ae^{-(x+1)} - (x-1)^2$, with $0 < A < \infty$. This particular example of a function $\theta(x, y)$, harmonic everywhere except at the points $\pm z_0$ (singularities of $\nabla \times \mathbf{S}$), is not arbitrary. It corresponds to an extremum of the classical action which is the continuous counterpart of (124)

$$S_{\text{class}} = \frac{1}{2}\beta \int d^2\mathbf{x} (\nabla \theta)^2 \quad (133)$$

The classical field equation is $\Delta\theta = 0$. We require that θ be constant at infinity. Due to the fundamental property of harmonic functions, either θ is constant or has singularities. Assuming only isolated singularities $z_i^\pm$ around which θ varies by $\pm 2\pi$ as one turns anticlockwise, one finds

$$e^{i\theta} = e^{i\theta_\infty} \prod_{i=1}^n \frac{(z - z_i^+)/|z - z_i^+|}{(z - z_i^-)/|z - z_i^-|} \quad (134)$$

Vortices of intensity ± 1 are centered at the points $z_i^\pm$. Intensities of higher magnitude are obtained when several z_i^+ (or z_i^-) coincide. The angle θ is a multivalued function, but $e^{i\theta}$ (or the spin) is well-defined everywhere, except at the singular points.

The classical action (133) is infinite for singular configurations such as (134). However, within the framework of the approximations that were made, it is consistent to regularize the action by excluding from the integration domain a small region – e.g. a disk of radius r_0 – centered on each singular point, and small with respect to the distances between vortices. Let D be the remaining integration domain. In order to simplify the notation, we relabel the points $z_i^\pm$ as z_j , denoting their intensities as q_j , and write

$$F = \rho e^{i\theta} = e^{i\theta_\infty} \prod_{i=1}^n \frac{z - z_i^+}{z - z_i^-} = e^{i\theta_\infty} \prod_{j=1}^{2n} (z - z_j)^{q_j} \quad (135)$$

In the domain D , θ and $\ln \rho$ are harmonic functions and, as a consequence of the Cauchy–Riemann equations, $(\nabla \theta)^2 = (\nabla \ln \rho)^2 = \nabla(\ln \rho \nabla \ln \rho)$. The classical action can now be evaluated for sufficiently small r_0 . If C_j is the circle surrounding z_j ,

$$\begin{aligned} S(\{q_j, z_j\}) &= \frac{1}{2}\beta \int_D d^2x (\nabla \theta)^2 = -\frac{1}{2}\beta \sum_j \oint_{C_j} ds \ln \rho n \cdot \nabla \ln \rho \\ &= -2\pi\beta \sum_{(ij)} q_i q_j \ln |z_i - z_j| + \sum_i q_i^2 \pi\beta \ln 1/r_0 \end{aligned} \quad (136)$$

In this formula, distances are expressed in terms of the lattice spacing. However, the complete expression is independent from this choice of unit, due to the relation $\sum q_j = 0$ which agrees with the scale invariance of the action.

Consider the case of a vortex pair, at distance r_{12} . Its statistical weight is $\exp(-2\pi\beta \ln r_{12}/r_0)$ (according to (136)). In a box of

linear dimension L , the global contribution of one vortex pair to the total free energy is, for dimensional reasons, of the form

$$\begin{aligned} \exp F_{\text{pair}} &\sim \int_{|x_1 - x_2| > r_0} d^2x_1 d^2x_2 \exp \left(-2\pi\beta \ln \frac{|x_1 - x_2|}{r_0} \right) \\ &\sim L^4 \exp(-2\pi\beta \ln L/r_0) \\ F_{\text{pair}} &\sim (4 - 2\pi\beta) \ln L \end{aligned} \quad (137)$$

If $2\pi\beta > 4$, this contribution is negligible in the limit $L \rightarrow \infty$. On the other hand, if $2\pi\beta < 4$, an instability appears. The creation of well-separated vortices is favoured and disorder increases. Hence a first estimate of the critical temperature is

$$2\pi\beta_c = 4 \quad (138)$$

More generally, choosing a length scale a , the action is rewritten near a stationary configuration as

$$S = S(q_i, x_i) + \frac{1}{2}\beta \int d^2x (\nabla \theta_{\text{fluc}})^2 \quad \text{with } \theta = \theta_{\text{conf}} + \theta_{\text{fluc}}$$

Summing over all vortex sectors yields an approximate partition function

$$\begin{aligned} Z &= Z_{\text{sw}} \sum_n \frac{1}{n!^2} \exp \left(-2\pi\beta n \ln \frac{a}{r_0} \right) \\ &\times \int \prod_j d^2x_j \exp \left\{ 2\pi\beta \sum_{(ij)} q_i q_j \ln \frac{|x_i - x_j|}{a} \right\} \end{aligned} \quad (139)$$

The (regularized) term $\pi\beta \ln a/r_0$ plays the role of the chemical potential for the $2n$ vortices of intensity $q_j = \pm 1$ located at x_j . The factor $n!$ for each species takes into account their indistinguishability, and the contribution of Gaussian fluctuations is factorized. Up to this term Z_{sw} , the partition function describes a gas of classical charged particles, with a Coulomb interaction (in two dimensions) and globally neutral. In this interpretation, the kinetic term has been integrated and absorbed into a redefinition of the chemical potential. The corresponding charges in this picture are $\sqrt{2\pi}q_j$. The preceding heuristic considerations can be translated in this interpretation. In a low temperature phase, charges are bound and one has a dielectric medium. Above the critical temperature, one has a plasma of free charges.

4.2.3 The Villain action

It is also possible to obtain a Coulomb model, starting from a slightly modified version of the action due to Villain. In the expansion (119)–(120) of the exponentiated action, the coefficients are replaced by their asymptotic form for large β , i.e. $\exp(-n^2/2\beta)$. The Boltzmann weight is replaced by the periodic Jacobi function, solution of the heat equation on a circle,

$$z(\theta) = \sum_{n=-\infty}^{+\infty} \exp\left(-\frac{n^2}{2\beta} + in\theta\right) = \sqrt{2\pi\beta} \sum_{p=-\infty}^{+\infty} \exp\left(-\frac{1}{2}\beta(\theta - 2\pi p)^2\right) \quad (140)$$

Prove that the two forms given in (140) are equivalent, using the Poisson summation formula.

$$f(x) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} e^{ikx} \tilde{f}(k) \quad \rightarrow \quad \sum_{n=-\infty}^{+\infty} f(n) = \sum_{p=-\infty}^{+\infty} \tilde{f}(2\pi p) \quad (141)$$

The second form given in (140) is useful for large β ; it appears as a sum of Gaussians of width $1/\sqrt{\beta}$ centered on the values $2\pi p$ (this ensures the periodicity in θ). The partition function reads

$$Z_V = \int \prod_i \frac{d\theta_i}{2\pi} \prod_{(ij)} z(\theta_i - \theta_j) \quad (142)$$

This form is suitable for a duality transformation. Integrating over the angles θ , one finds

$$Z_V = \sum_{\substack{\{n_{ij}\} \\ (\partial n)_i = 0}} \exp\left(-\frac{1}{2\beta} \sum_{(ij)} n_{ij}^2\right) \quad (143)$$

In this expression, a relative integer $n_{ij} \equiv -n_{ji}$ has been assigned to each link, and the condition $(\partial n)_i = 0$ means $\sum_j n_{ij} = 0$, as previously. On a two-dimensional square lattice, these conditions can be easily solved. To each plaquette (elementary square), or, equivalently, to its center which is a node of the dual lattice (isomorphic to the initial one), one assigns an integral variable m_a . Each link ij carries the difference between those variables

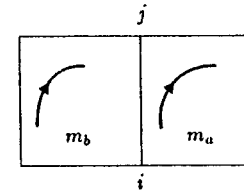


Fig. 11 Dual variables for the two-dimensional Villain model.

relative to the right and left adjacent plaquettes, $n_{ij} = m_a - m_b$ (figure 11).

In these variables, the Villain model is transformed by duality into a Gaussian discrete model

$$Z_V = \sum_{\{m_a\}} \exp\left(-\frac{1}{2\beta} \sum_{(ab)} (m_a - m_b)^2\right) \quad (144)$$

This solution of the equations $(\partial n)_i = 0$ in the form $n_{ij} = m_a - m_b$, leaves an arbitrariness $m_a \rightarrow m_a + m$; m is any integer. This ambiguity leads to an infinity which should be avoided, e.g. by omitting the summation over one of the variables m_a . The duality has here three effects. The global symmetry group $O(2)$ is replaced by its dual, the additive group of integers. The lattice is replaced by its dual, which is here isomorphic to the original one. Finally, β is changed to $1/\beta$.

Configurations can be visualized as surfaces in three dimensions joining plaquettes at height m_a above the lattice basis. The Gaussian discrete model describes thus a fluctuating surface, as does a so-called *SOS* (solid on solid) model in which the term $(m_a - m_b)^2$ is replaced by $|m_a - m_b|$, with an action which measures the deviation with respect to a planar surface. For small β , configurations with equal m_a are favoured. The cost of a unit step is $e^{-2/\beta}$. As β increases and approaches β_c , the surface, which was a smooth one up to microscopic fluctuations, changes structure in such a way that two distant points have uncorrelated heights. The surface is said to become “rough” (*roughening transition*). This model is used in particular to describe the shape of crystals in equilibrium with a solution.

For sufficiently large β , the discrete Gaussian model behaves nearly as its continuous counterpart, i.e. as a spin wave system. This can be exhibited by substituting to the sum over discrete variables an integration over continuous ones φ , according to

$$\sum_{m=-\infty}^{+\infty} f(m) = \int d\varphi f(\varphi) \sum_{m=-\infty}^{+\infty} \delta(\varphi - m) = \int d\varphi f(\varphi) \sum_{q=-\infty}^{+\infty} e^{2i\pi q\varphi} \quad (145)$$

$$Z_V = \sum_{\{q_a\}} \int \prod_a d\varphi_a \exp \left\{ -\frac{1}{2\beta} \sum_{(ab)} (\varphi_a - \varphi_b)^2 + 2i\pi \sum_a q_a \varphi_a \right\} \quad (146)$$

In this expression, the variables $\{\varphi_a\}$ describe continuous Gaussian fluctuations and the discrete ones $\{q_a\}$ the intensity of vortices. The factorization described in the preceding subsection is explicitly realized, and the formulae become identical after a change of notations $\varphi = \beta\theta_{\text{auc}}$. Global translational invariance of the fields φ is insured by the neutrality condition $\sum_a q_a = 0$. Other configurations have a vanishing weight in the thermodynamic limit.

This expression also allows one to understand the relation between the XY-model with yet another version, the so-called *sine-Gordon* model. The name (of somewhat dubious taste) arises from a nonlinear modification of the Klein-Gordon equation. For large β , vortices have an almost negligible effect, and the effect of the terms involving the variables q_a essentially introduce a periodic modulation in the integral over the fields φ_a . The action takes the sine-Gordon form

$$\frac{1}{2\beta} \sum_{(ab)} (\varphi_a - \varphi_b)^2 - \sum_a 2 \cos 2\pi \varphi_a + \dots$$

which is also suitable for a study of the XY-model.

Alternatively, integrating in equation (146) over the variables φ and taking into account neutrality, one obtains, without any approximation, a Coulomb gas of vortices

$$Z_V = Z_{\text{sw}} \sum_{\{q_a\}} \exp \left\{ 2\pi\beta \sum_{(ab)} q_a \Gamma(\mathbf{x}_a - \mathbf{x}_b) q_b \right\} \quad (147)$$

This is equivalent, to a very good approximation, to

$$Z_V = Z_{\text{sw}} \sum_{\{q_a\}} \exp \left\{ 2\pi\beta \sum_{(ab)} q_a \ln \frac{|\mathbf{x}_a - \mathbf{x}_b|}{a} q_b - \pi\beta \sum_a q_a^2 \ln \frac{a}{r_0} \right\} \quad (148)$$

The Villain model thus justifies the semiclassical derivation of the previous subsection.

As the effective chemical potential $\pi\beta \ln a/r_0$ is large (or the activity $y = \exp -\pi\beta \ln a/r_0$ is small), the first nontrivial configuration is a vortex pair of intensity ± 1 located respectively at $\mathbf{x}_1$ and $\mathbf{x}_2$. The corresponding correlation function is negative, since the intensities are opposite, and behaves as

$$\begin{aligned} \langle q(\mathbf{x}_1)q(\mathbf{x}_2) \rangle &= -y^2 \exp -2\pi\beta \ln(|\mathbf{x}_1 - \mathbf{x}_2|/a) \\ &\simeq -y^2 (a/|\mathbf{x}_1 - \mathbf{x}_2|)^{2\pi\beta} \end{aligned} \quad (149)$$

when $|\mathbf{x}_1 - \mathbf{x}_2| \gg a$ and $y \rightarrow 0$. According to the neutrality condition requiring that

$$\sum_{\mathbf{x}_1} \langle q(\mathbf{x}_1)q(\mathbf{x}_2) \rangle = 0$$

one also has

$$\langle q^2(0) \rangle \simeq y^2 \sum_{\mathbf{x} \neq 0} (a/|\mathbf{x}|)^{2\pi\beta} \quad (150)$$

As $\beta \rightarrow \infty$, few vortices are present and their correlations decrease rapidly with the relative distance. This describes a dielectric medium of neutral bound states. It is natural to perform an expansion in powers of y , considered as a parameter independent of β .

Let us now examine the effects of vortices on the spin correlations, already computed in the spin wave approximation (equations (128)–(132)).

4.2.4 Correlations

In the framework of the Villain model, let us compute a correlation

$$\langle e^{ip(\theta_1 - \theta_2)} \rangle = Z'_V / Z_V \quad (151)$$

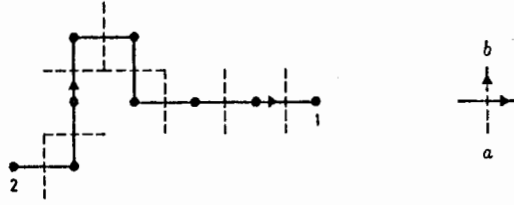


Fig. 12 The line C joining x_2 to x_1 in the computation of correlations. On the dual links, $\eta_{ab} = \pm 1$, according to the indicated orientation prescription.

In this ratio, Z'_V has a form similar to the sum (143) obtained after integration over the angles θ , i.e.

$$Z'_V = \sum_{\{n_{ij}\}} \exp\left(-\frac{1}{2\beta} n_{ij}^2\right) \quad (152)$$

$(\partial n)_i = p(\delta_{i,1} - \delta_{i,2})$

In order to define the dual expression, one has to solve the condition $(\partial n)_i = p(\delta_{i,1} - \delta_{i,2})$. Let us draw on the lattice an arbitrary line C joining the points 1 and 2, and oriented, e.g. from 2 to 1. On the links of the dual lattice, we define an antisymmetric function η_{ab} , vanishing if the link ab does not cross C and equal to ± 1 according to the relative orientations of C and ab if they cross (figure 12). Accordingly, it is possible to define integral variables m_a on the dual lattice sites such that

$$n_{ij} = m_a - m_b + p\eta_{ab} \quad (153)$$

As previously, summing over all these variables but one,

$$\begin{aligned} Z'_V &= \sum_{\{m_a\}} \exp -\frac{1}{2\beta} \sum_{(ab)} (m_a - m_b + p\eta_{ab})^2 \\ &= \sum_{\{q_a\}} \int \prod_a d\varphi_a \exp \left\{ -\frac{1}{2\beta} \sum_{(ab)} (\varphi_a - \varphi_b + p\eta_{ab})^2 + 2i\pi \sum_a q_a \varphi_a \right\} \end{aligned} \quad (154)$$

Integration over the φ variables is readily performed. Defining

$$\begin{aligned} \eta_a &= \sum_b \eta_{ab} \\ q'_a &= -\frac{p}{2i\pi\beta} \eta_a \end{aligned} \quad (155)$$

and using the fact that $\sum_a q'_a = 0$ (due to the antisymmetry of η_{ab}), one gets

$$\begin{aligned} Z'_V &= Z_{sw} \exp \left\{ -\frac{p^2}{2\beta} \sum_{(ab)} \eta_{ab}^2 \right\} \\ &\times \sum_{\{q_a\}} \exp \left\{ 2\pi\beta \sum_{(ab)} (q_a + q'_a) \Gamma(x_a - x_b) (q_b + q'_b) \right\} \end{aligned} \quad (156)$$

Taking into account the definitions (155) and the neutrality condition, the correlation factorizes as

$$\langle e^{ip(\theta_1 - \theta_2)} \rangle = G_{\text{wave}}(x_1 - x_2) G_{\text{Coulomb}}(x_1 - x_2) \quad (157)$$

with

$$\begin{cases} G_{\text{wave}}(x_1 - x_2) = \exp \left(-\frac{p^2}{2\beta} \sum_{(ab)} \eta_{ab}^2 - \frac{p^2}{2\pi\beta} \sum_{ab} \eta_a \Gamma(x_a - x_b) \eta_b \right) \\ G_{\text{Coulomb}}(x_1 - x_2) = \left\langle \exp ip \sum_a q_a \sum_b \Gamma(x_a - x_b) \eta_b \right\rangle_{\text{Coulomb}} \end{cases} \quad (158)$$

The last mean value involves the weights corresponding to (148). The label of the first factor G_{wave} is justified as follows. The exponent on the r.h.s. is written as $-(p^2/2\beta)\gamma(x_a - x_b)$ such that

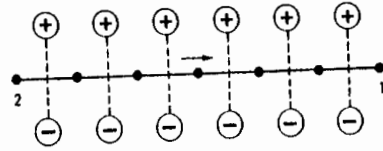
$$\gamma(x_a - x_b) = \sum_{(ab)} \eta_{ab}^2 - \sum_{a,b} \eta_a G(x_a - x_b) \eta_b \quad (159)$$

with

$$G(x) = \int_{-\pi}^{\pi} \frac{d^2 q}{(2\pi)^2} \frac{e^{iq \cdot x}}{4 - 2 \sum_{\mu=1}^2 \cos q_{\mu}}$$

The infrared divergence is automatically subtracted by virtue of the neutrality condition. The function η_a is a discretized version of a dipole potential on the curve C , as indicated in a particular case on figure 13. We perform the computations in this particular case with x_1 and x_2 along a principal axis of the lattice. With $r = |x_1 - x_2|$, one has

$$\gamma(r) = r - \sum_{x, x'=1}^{r-1} \{2G(x - x', 0) - G(x - x', 1) - G(x - x', -1)\}$$

Fig. 13 The dipole potential η_a .

Since

$$4G(x, y) - [G(x+1, y) + G(x-1, y) + G(x, y+1) + G(x, y-1)] = \delta_{x,0}\delta_{y,0}$$

this is also

$$\begin{aligned} \gamma(r) &= r - \sum_{x, x'=0}^{r-1} \{ \delta_{x, x'} - 2G(x-x', 0) + G(x-x'+1, 0) + G(x-x'-1, 0) \} \\ &= \sum_{x=0}^{r-1} \{ G(r-1-x, 0) - G(r-x, 0) + G(-x, 0) - G(-x-1, 0) \} \\ &= 2 \{ G(0) - G(r) \} = \Gamma(r)/\pi \end{aligned}$$

Returning to G_{wave} , one obtains as expected

$$G_{\text{wave}}(\mathbf{x}_1 - \mathbf{x}_2) = \exp \left[-\frac{p^2}{2\pi\beta} \Gamma(|\mathbf{x}_1 - \mathbf{x}_2|) \right] \quad (160)$$

In the evaluation of G_{Coulomb} , we use a cumulant expansion. To first order, $\langle q_a \rangle = 0$, and therefore

$$G_{\text{Coulomb}}(\mathbf{x}_1 - \mathbf{x}_2) \simeq \exp \left(-\frac{1}{2} p^2 \sum_{a,b} \langle q(\mathbf{x}_a) q(\mathbf{x}_b) \rangle \sum_{a'} \Gamma(\mathbf{x}_a - \mathbf{x}_{a'}) \eta_{a'} \sum_{b'} \Gamma(\mathbf{x}_b - \mathbf{x}_{b'}) \eta_{b'} \right) \quad (161)$$

The correlation $\langle q(\mathbf{x}_a) q(\mathbf{x}_b) \rangle$ has already been computed (equations (149)–(150)). Sums over a' or b' are estimated in the limit where $\mathbf{x}_a$ is far from the curve C , compared to the lattice spacing. In this limit, we use the formalism of analytic functions. If C is along the x axis,

$$\begin{aligned} \sigma_a &= \sum_{a'} \Gamma(\mathbf{x}_a - \mathbf{x}_{a'}) \eta_{a'} \simeq \int_{x_2}^{x_1} dx \frac{\partial}{\partial y} \ln |z_a - z|_{z=x} \\ &= - \int_{x_2}^{x_1} dx \frac{\partial}{\partial x} \text{Im} \ln(z_a - x) = -\text{Im} \ln \left(\frac{z_a - z_1}{z_a - z_2} \right) \end{aligned} \quad (162)$$

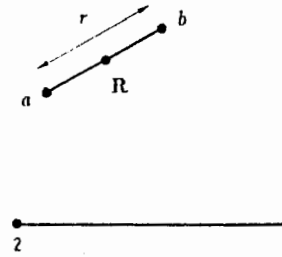


Fig. 14 Notations used in the computation of the effect of a vortex pair on the spin correlation function.

where $z_i = x_i$. Let z_R be the center of the segment ab . Pairs are strongly bound since the correlation $\langle q_a q_b \rangle$ decreases rapidly with the relative distance of the points ab . It is thus possible to expand $\sigma_{a,b}$ as

$$\sigma(\mathbf{R}) \pm \frac{1}{2} (\mathbf{x}_a - \mathbf{x}_b) \cdot \nabla \sigma(\mathbf{R})$$

Hence

$$\begin{aligned} G_{\text{Coulomb}}(\mathbf{x}_1 - \mathbf{x}_2) &\simeq \exp -\frac{1}{2} p^2 \sum_{a,b} \langle q(\mathbf{x}_a) q(\mathbf{x}_b) \rangle \sigma(\mathbf{x}_a) \sigma(\mathbf{x}_b) \\ &\simeq \exp \left(\frac{1}{8} p^2 \sum_{\mathbf{r}} \langle q(0) q(\mathbf{r}) \rangle \sum_{\mathbf{R}} [\mathbf{r} \cdot \nabla \sigma(\mathbf{R})]^2 \right) \end{aligned} \quad (163)$$

If (X, Y) stand for the coordinates of $\mathbf{R}$,

$$\frac{\partial \sigma}{\partial X} = \frac{Y - y_1}{|\mathbf{R} - \mathbf{x}_1|^2} - \frac{Y - y_2}{|\mathbf{R} - \mathbf{x}_2|^2} \quad \frac{\partial \sigma}{\partial Y} = -\frac{X - x_1}{|\mathbf{R} - \mathbf{x}_1|^2} + \frac{X - x_2}{|\mathbf{R} - \mathbf{x}_2|^2} \quad (164)$$

Furthermore, $\langle q(0) q(\mathbf{r}) \rangle$ is essentially a function of $r = |\mathbf{r}|$ only, so that one can perform an average over the orientation of $\mathbf{r}$,

$$\sum_{\mathbf{R}} [\mathbf{r} \cdot \nabla \sigma(\mathbf{R})]^2 \rightarrow \frac{1}{2} r^2 \sum_{\mathbf{R}} [\nabla \sigma(\mathbf{R})]^2$$

According to equation (164), we have

$$[\nabla \sigma(\mathbf{R})]^2 = [\nabla \Gamma(R - x_1) - \nabla \Gamma(R - x_2)]^2$$

As $\Delta\Gamma(R) = 2\pi\delta_{R,0}$,

$$\begin{aligned} & \frac{1}{2}r^2 \sum_R [\nabla \sigma(\mathbf{R})]^2 \\ &= -\frac{1}{2}r^2 \sum_R [\Gamma(\mathbf{R} - \mathbf{x}_1) - \Gamma(\mathbf{R} - \mathbf{x}_2)] \Delta[\Gamma(\mathbf{R} - \mathbf{x}_1) - \Gamma(\mathbf{R} - \mathbf{x}_2)] \\ &= 2\pi r^2 \Gamma(\mathbf{x}_1 - \mathbf{x}_2). \end{aligned}$$

Hence the final expression is

$$G_{\text{Coulomb}}(\mathbf{x}_1 - \mathbf{x}_2) \simeq \exp \left(\frac{1}{4} \pi p^2 \Gamma(\mathbf{x}_1 - \mathbf{x}_2) \sum_{|\mathbf{r}| > a} r^2 \langle q(0)q(\mathbf{r}) \rangle \right) \quad (165)$$

This result is very similar to the one given by the spin wave approximation (160), and differs by the coefficient of the exponent. Multiplying (160) and (165), one gets

$$\begin{aligned} \langle e^{ip(\theta_1 - \theta_2)} \rangle &\simeq \exp \left\langle -\frac{p^2}{2\pi\beta_{\text{eff}}} \Gamma(\mathbf{x}_1 - \mathbf{x}_2) \right\rangle \\ \frac{1}{\beta_{\text{eff}}} &= \frac{1}{\beta} - \frac{1}{2}\pi^2 \sum_{|\mathbf{r}| > a} r^2 \langle q(0)q(\mathbf{r}) \rangle \end{aligned} \quad (166)$$

The effect of taking into account the vortices is to increase the effective temperature β^{-1} , since equations (149)-(150) lead to

$$\frac{1}{\beta_{\text{eff}}} = \frac{1}{\beta} + \pi^3 y^2 \int_a^\infty \frac{dr}{a} \left(\frac{a}{r} \right)^{2\pi\beta-3} \quad y = \exp \left(-\pi\beta \ln \frac{a}{r_0} \right) \quad (167)$$

in which we reinstated the cutoff factor a . As a consequence, the correlation, which behaves as $r^{-p^2/2\pi\beta_{\text{eff}}}$, decrease more rapidly than in the spin wave approximation. The integral (167) has an infrared divergence for $2\pi\beta = 4$, a signal of the failure of the perturbative calculation when vortices become widely separated. This yields the same estimate as before for the critical temperature. We can now establish heuristically the renormalization group equations governing the critical behaviour.

Derive similar equations for higher order correlation functions.

4.2.5 Renormalization flow

We have just computed the effects of vortices on the correlation function to lowest order in the activity y , and found that the corresponding effective temperature increases. The parameters y and β are dimensionless, and are defined at the microscopic length scale a , the lattice spacing. The question is now to obtain the values y_λ and β_λ at a larger scale λa , in order that the correlation functions pertaining to this new scale retain their original values. We have to assume that y_λ remains sufficiently small, in order that the above calculation retains its validity. In formula (167), the integral from a to infinity is split into two parts, one from a to λa , and the second from λa to infinity. With $x = r/a$, this yields

$$\begin{aligned} \int_1^\infty dx x^{3-2\pi\beta} &= \int_1^\lambda dx x^{3-2\pi\beta} + \int_\lambda^\infty dx x^{3-2\pi\beta} \\ &= \frac{\lambda^{4-2\pi\beta} - 1}{4 - 2\pi\beta} + \lambda^{4-2\pi\beta} \int_1^\infty dx x^{3-2\pi\beta} \\ &\sim_{\lambda \rightarrow 1} \ln \lambda + [1 + (4 - 2\pi\beta) \ln \lambda] \int_1^\infty dx x^{3-2\pi\beta} \end{aligned}$$

Using this expression in the neighborhood of $\lambda = 1$, one finds

$$\frac{1}{\beta_{\text{eff}}} = \frac{1}{\beta} + \pi^3 y^2 \ln \lambda + \pi^3 y^2 [1 + (4 - 2\pi\beta) \ln \lambda] \int_1^\infty dx x^{3-2\pi\beta}$$

As $\lambda \rightarrow 1$, $\ln \lambda \sim \lambda - 1$ is small and the formula can be recast into a similar form

$$\frac{1}{\beta_{\text{eff}}} = \frac{1}{\beta_\lambda} + \pi^3 y_\lambda^2 \int_1^\infty dx x^{3-2\pi\beta_\lambda} \quad (168a)$$

with

$$\begin{aligned} \frac{1}{\beta_\lambda} &= \frac{1}{\beta} + \pi^3 y^2 \ln \lambda \\ y_\lambda &= y + (2 - \pi\beta)y \ln \lambda \end{aligned} \quad (\lambda \rightarrow 1) \quad (168b)$$

These relations are rewritten in differential form

$$\begin{cases} \lambda \frac{d}{d\lambda} \beta_\lambda^{-1} = \pi^3 y_\lambda^2 \\ \lambda \frac{d}{d\lambda} y_\lambda = (2 - \pi\beta_\lambda) y_\lambda \end{cases} \quad (169)$$

Although the preceding reasoning might look somehow dubious, since we obtain the variation of two parameters using a single

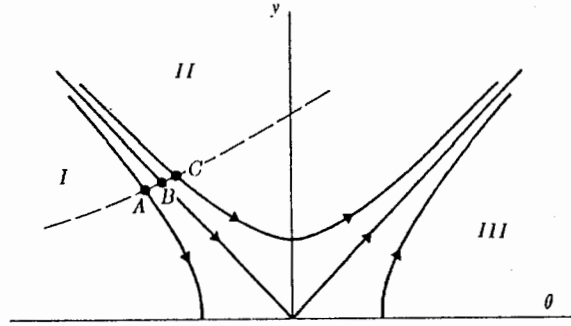


Fig. 15 The Kosterlitz-Thouless renormalization flow.

condition, the result agrees with a more elaborate analysis. The system (169) describes the renormalization flow of the bare coupling constants as the cut-off scale varies, and has been analyzed by Kosterlitz. Initial conditions are, for $\lambda = 1, \beta_1 = \beta$ and $y_1 = y = \exp(-\pi\beta \ln a/r_0)$. The system is only valid for small y_λ . Additional terms involving higher powers in y should become relevant outside this domain. The equations simplify when one introduces the reduced variables

$$\begin{aligned}\Theta_\lambda &= 2 - \pi\beta_\lambda \\ Y_\lambda &= 2\pi y_\lambda\end{aligned}\quad (170)$$

in which case

$$\begin{cases} \lambda \frac{d}{d\lambda} \Theta_\lambda = Y_\lambda^2 (1 - \frac{1}{2} \Theta_\lambda)^2 \simeq Y_\lambda^2 \\ \lambda \frac{d}{d\lambda} Y_\lambda^2 = 2\Theta_\lambda Y_\lambda^2 \end{cases} \quad (171a)$$

Near the fixed point $\Theta = 0, Y = 0$, the integral curves are hyperbolae

$$\Theta_\lambda^2 - Y_\lambda^2 = Cst \quad (171b)$$

In view of the second equation (171a), as λ grows, the axis $Y = 0$ is attractive for $\Theta < 0$ and repulsive for $\Theta > 0$. The renormalization flow is displayed in figure 15. In the domain I, renormalization drives the representative point to a critical one characterized by $Y_\infty = 0, \Theta_\infty \simeq -\sqrt{\Theta^2 - Y^2} < 0$. This is a

scale invariant low temperature fixed point, and the correlations decrease according to a power law $|x|^{-1/2\pi\beta_\infty}$. In regions II and III, the renormalization flow escapes rapidly from the domain of validity for the approximations. The effective activity for vortices Y_λ increases almost linearly with Θ , and one approaches a high temperature fixed point with free vortices (the plasma phase of the Coulomb gas, or the smooth phase in the language of surfaces), characterized by an exponential decay of correlations. The separating line is $Y = -\Theta$, along which the representative point flows towards the limiting point $Y = 0, \Theta = 0$. The figure also displays the initial conditions for $\lambda = 1, y = \exp(-\pi\beta \ln a/r_0)$, as a dashed curve

$$Y = 2\pi e^{(\Theta-2)\ln a/r_0} \simeq 2\pi e^{-2\ln a/r_0} (1 + \Theta \ln \frac{a}{r_0} + \dots) \quad (172)$$

in which $\ln a/r_0 = \ln 2\sqrt{2}e^\gamma \simeq 1.6$. This can be improved in a more detailed treatment. In any case, this is a nonuniversal curve, depending on the details of the model. For the Villain model, the intersection of the curve (172) with $Y = -\Theta$ leads to an estimate of the critical temperature

$$Y_c = -\Theta_c = \pi\beta_c - 2 \simeq \left(\ln \frac{a}{r_0} + \frac{e^{2\ln a/r_0}}{2\pi} \right)^{-1} \quad (173)$$

The parametrization in λ of the renormalization flow trajectories near the origin can also be found. In region I, we have $\Theta_\lambda^2 = Y_\lambda^2 + \Theta_\infty^2$. From the initial values $\Theta_1 = \Theta, Y_1 = Y$, one gets $\Theta_\infty^2 = \Theta^2 - Y^2$. Solving (171), one obtains ($\Theta_\infty < 0$)

$$(I) \quad \begin{cases} \Theta_\lambda = \Theta_\infty \frac{1 + (\lambda/\lambda_i)^{2\Theta_\infty}}{1 - (\lambda/\lambda_i)^{2\Theta_\infty}} \\ Y_\lambda = \sqrt{\Theta_\lambda^2 - \Theta_\infty^2} \\ \lambda_i^{-2\Theta_\infty} = \frac{\Theta - \Theta_\infty}{\Theta + \Theta_\infty} > 0 \end{cases} \quad (174)$$

As $\Theta_\infty < 0$, the terms depending on initial conditions decay as power laws in λ , which is characteristic of scale invariant theories.

In region II, let $Y_0 > 0$ be the height of the hyperbola above the point $\Theta = 0, Y_\lambda^2 - \Theta_\lambda^2 = Y_0^2 = Y^2 - \Theta^2$. One introduces a variable

ψ varying from ψ_i to $\frac{1}{2}\pi$, with

$$(II) \quad \begin{cases} \Theta_\lambda = Y_0 \tan \psi \\ Y_\lambda = \frac{Y_0}{\cos \psi} \end{cases} \quad \begin{cases} \lambda = \exp \frac{\psi - \psi_i}{Y_0} \\ \tan \psi_i = \frac{\Theta}{\sqrt{Y^2 - \Theta^2}} \end{cases} \quad (175)$$

The effective parameter Θ_λ varies monotonically from 0 to infinity, while Y_λ remains positive, increases indefinitely as $\lambda \rightarrow \infty$ and has a minimum if $\psi_i < 0$. The large positive values correspond to a domain where the approximation is not valid. However, if one starts from an initial value very close to the critical point, with $\Theta < 0, \psi_i < 0$, the renormalization flow can be followed in a region where Y_λ remains small, and the approximation can still be trusted. This is justified until one reaches a point close to the minimum of the curve Y_λ , i.e. with Y_0 corresponding to $\psi = 0$, up to a value of λ of the order of

$$\lambda \simeq e^{-\psi_i/Y_0} \quad (176)$$

This value of λ may be identified, up to a factor, with ξ/a , where ξ is the correlation length in the disordered phase near the end point of the critical line, since it defines the scale up to which the system remains close to the critical regime. When one approaches this point, Y_0 vanishes as

$$Y_0^2 \sim (\Theta - \Theta_c) \simeq (\beta_c - \beta) \quad (177)$$

while $\sin \psi_i = \Theta/Y$ is very close to -1, i.e. $\psi_i \sim -\frac{1}{2}\pi$. Hence,

$$\xi/a \simeq A \exp \frac{B}{\sqrt{\beta_c - \beta}} \quad (178)$$

where A and B are positive constants. Thus the correlation length increases faster than any power when one approaches the critical point. In this sense, the exponent ν is infinite, and the singular part of the free energy is expected to behave for $\beta_c > \beta$ as

$$F_{\text{sing}} \simeq \xi^{-2} \simeq Cst \exp \frac{-2B}{\sqrt{\beta_c - \beta}} \quad \beta \rightarrow \beta_c - 0 \quad (179)$$

In particular, all derivatives are continuous at β_c . This means that the XY transition is of infinite order.

The critical point is represented as point B in figure 15. On the one hand, the temperature is smaller than the estimate

$2\pi\beta_c = 4$. On the other hand, the effect of vortices is to damp the correlations. The two effects go in opposite directions and finally the critical correlation function behaves as $r^{-1/2\pi\beta_c} = r^{-1/4}$ up to logarithms. Following Kosterlitz, we may analyze the effect of changing the cutoff by an infinitesimal amount, $a \rightarrow a + \varepsilon a$, in the vortex contribution. According to (168b), we have for large $r = |\mathbf{x}|$

$$G_{\text{Coulomb}}(\mathbf{x}, a) \sim G_{\text{Coulomb}}(\mathbf{x}, a + \varepsilon a) \exp -\frac{\pi^2}{2} y^2 \varepsilon \ln \frac{r}{a} \quad (180)$$

Here, the dependence on a arises also from y and β . Equation (180) means that we trust the perturbative calculation only insofar as it provides us a means to estimate the effect in the change of scale, and we conceal the rest into the correlation computed with a cutoff $a + \varepsilon a$. This can in turn be iterated, provided we use the scale dependent quantities β_λ, y_λ until $\lambda a \sim r$, at which point the vortex contribution becomes of order unity, with a vortex separation of order of the renormalized lattice spacing. Using $Y_\lambda = 2\pi y_\lambda$ as above, this means that, provided the activity remains small, the total correlation function reads

$$G(\mathbf{x}) \simeq Cst \left(\frac{a}{|\mathbf{x}|} \right)^{1/2\pi\beta} \exp -\frac{1}{8} \int_1^{r/a} \frac{d\lambda}{\lambda} Y_\lambda^2 \ln \frac{r}{\lambda a} \quad (181)$$

the first factor being the spin wave contribution. Recall from equations (171) that, for small Y_λ and Θ_λ , $Y_\lambda^2 d\lambda/\lambda = d\Theta_\lambda$. Let us look at the critical theory, characterized by the fact that initial conditions satisfy $\Theta_1 + Y_1 = 0$, a relation which remains valid for any scale λ . We have therefore

$$\Theta^{-1} - \Theta_\lambda^{-1} = \ln \lambda \quad (182)$$

both Θ and Θ_λ being negative. In this case, the exponent in (181) takes the form

$$-\frac{1}{8} \int_1^{r/a} d\Theta_\lambda \ln \frac{|\mathbf{x}|}{\lambda a} \underset{r/a \rightarrow \infty}{\sim} \frac{\Theta}{8} \ln \frac{|\mathbf{x}|}{a} + \frac{1}{8} \ln \ln \frac{|\mathbf{x}|}{a} + Cst \quad (183)$$

It is good to note that changing the upper limit by a finite factor would not affect the leading behaviour. Working to leading order in Θ (as we did in the renormalization flow equations), we can (and should) approximate $(a/|\mathbf{x}|)^{1/2\pi\beta} = (a/|\mathbf{x}|)^{1/4(1-\Theta/2)} \simeq (a/|\mathbf{x}|)^{1/4+\Theta/8+\dots}$, so that, combining with (183), we find at

criticality

$$G(\mathbf{x}) \sim \text{Cst} \left(\frac{a}{|\mathbf{x}|} \right)^{1/4} (\ln |\mathbf{x}|)^{1/8} \quad (184)$$

Since $d = 2$, this means that the exponent η is given by

$$\eta = \frac{1}{4} \quad (185)$$

a result surprisingly identical to the one of Ising model, except for the logarithmic correction. The latter is seen to arise from the fact that the initial nonuniversal model is described by a curve in the (Y, Θ) plane, which does not go through the point $(0, 0)$. The logarithmic term reflects the presence of a marginal operator conjugate to the vortex activity.

If one uses ξ^{-1} instead of $\beta - \beta_c$ as a parameter to characterize the approach to the critical point (ν is infinite), one expects (cf. equation (42)) that the susceptibility in the disordered phase behaves as

$$\chi \sim \xi^{2-\eta} \sim \exp \frac{b}{\sqrt{\beta_c - \beta}} \quad \beta \rightarrow \beta_c - 0 \quad (186)$$

and diverges more rapidly than any power of $\beta_c - \beta$. Finally, the exponent δ describing the response to an external field is also given by scaling relations (which can be verified through a direct calculation)

$$\delta = \frac{d+2-\eta}{d-2+\eta} = 15 \quad (187)$$

again as in the Ising model.

The renormalization group ideas have therefore allowed Kosterlitz and Thouless to give a very precise description of the transition of the XY-model. There exists an exactly soluble model (in the family of six vertex models solved by Lieb) belonging to the same universality class which confirms the preceding analysis, as shown by van Beijeren.

Appendix 4.A Two-dimensional systems with continuous symmetry

4.A.1 Magnetization inequality

We present the proof given by Hohenberg, Mermin and Wagner concerning the vanishing of the order parameter for a system with a continuous symmetry, in dimension less than or equal to two. We only discuss a n -vector model with nearest neighbour ferromagnetic interactions. However, the following proof can be easily extended to a larger class of models (short-range ferromagnetic interactions, constraint $S^2 = 1$ released,...). The goal is to establish a strong enough inequality on the magnetization in the presence of an external uniform field. One uses Schwarz's inequality and the rotational invariance of the spin integration measure.

We consider a large, but finite lattice with N sites. Fourier analysis uses discrete momenta. The limit $N \rightarrow \infty$ will be implicitly taken at the end. The n -component spins $\mathbf{S}_\mathbf{x}$ interact through the action

$$S = \beta \left(\sum_{\langle \mathbf{x}, \mathbf{y} \rangle} \mathbf{S}_\mathbf{x} \cdot \mathbf{S}_\mathbf{y} + \sum_{\mathbf{x}} \mathbf{H} \cdot \mathbf{S}_\mathbf{x} \right) \quad (A.1)$$

and the statistical weight is $Z^{-1} \prod_{\mathbf{x}} d^n \mathbf{S}_\mathbf{x} \delta(S_\mathbf{x}^2 - 1) e^S$. The external field $\mathbf{H}$ is chosen along the n th coordinate axis in internal isotopic space, with $H = |\mathbf{H}|$. The magnetization

$$M(H) = \langle S_\mathbf{x}^n \rangle = \frac{1}{N} \left\langle \sum_{\mathbf{x}} S_\mathbf{x}^n \right\rangle \quad (A.2)$$

is positive, as can be seen easily (using e.g. a high temperature expansion, which is convergent on a finite lattice). With the positive statistical weight, any pair of finite sets of observables A_i and B_i (functions of the spin variables) satisfies the Schwarz inequality

$$\left| \left\langle \sum_i A_i B_i \right\rangle \right|^2 \leq \left\langle \sum_i A_i^2 \right\rangle \left\langle \sum_i B_i^2 \right\rangle \quad (A.3)$$

In order to use the rotational invariance of the measure, we consider the infinitesimal generators $L_{\alpha\beta} = S^\alpha \partial / \partial S^\beta - S^\beta \partial / \partial S^\alpha$.

They obey the following identities

$$\begin{aligned} \int d^n S \delta(S^2 - 1) L_{\alpha\beta} f(S) &= 0 \\ \int d^n S \delta(S^2 - 1) [g(S) L_{\alpha\beta} f(S) + f(S) L_{\alpha\beta} g(S)] &= 0 \end{aligned} \quad (A.4)$$

The second one follows from the first if one remarks that L is a first order differential operator. The first one is a consequence of the rotational invariance of the measure, expressed by the relation

$$0 = \frac{d}{d\theta} \int d^n S \delta(S^2 - 1) f(S^1, \dots, S^\alpha \cos \theta + S^\beta \sin \theta, \dots, S^\beta \cos \theta - S^\alpha \sin \theta, \dots, S^n)$$

We also use the notation $L_{x,i} = S_x^n \partial / \partial S_x^i - S_x^i \partial / \partial S_x^n$ for the rotations in the plane (n, i) , and $L_x \equiv \{L_{x,i}\}$, i varying here from 1 to $n-1$.

Inequality (A.3) is applied to

$$\begin{cases} A(k) = \sum_x e^{ik \cdot x} S_x^\perp \\ B(k) = - \sum_x e^{ik \cdot x} L_x(S) \end{cases} \quad (A.5)$$

In $B(k)$, the operator L_x acts on the action S . From equations (A.4), one deduces, whatever the functional C is,

$$\langle CB(k) \rangle = \left\langle \sum_x e^{ik \cdot x} L_x(C) \right\rangle \quad (A.6)$$

This relation allows one to recast the various terms of (A.3) in the form

$$\begin{aligned} \langle A(k) \cdot B(k) \rangle &= \sum_x \langle L_x \cdot S_x^\perp \rangle = (n-1) \sum_x \langle S_x^n \rangle = (n-1) NM(H) \\ \langle A(k) \cdot A(k) \rangle &= \sum_{x,y} e^{ik \cdot (x-y)} \langle S_x^\perp \cdot S_y^\perp \rangle \\ \langle B(k) \cdot B(k) \rangle &= - \sum_{x,y} e^{ik \cdot (x-y)} \langle L_x \cdot L_y(S) \rangle \\ &= \beta \left\{ N(n-1) HM(H) + \sum_{(x,y)} 2(1 - \cos k \cdot (x-y)) \right. \\ &\quad \left. \langle (n-1) S_x^n S_y^n + S_x^\perp \cdot S_y^\perp \rangle \right\} \end{aligned}$$

Hence the inequality

$$\begin{aligned} (n-1)^2 M^2(H) &\leq \beta \frac{1}{N} \sum_{x,y} e^{ik \cdot (x-y)} \langle S_x^\perp \cdot S_y^\perp \rangle \left((n-1) HM(H) \right. \\ &\quad \left. + \frac{1}{N} \sum_{(x,y)} 2(1 - \cos k \cdot (x-y)) \langle (n-1) S_x^n S_y^n + S_x^\perp \cdot S_y^\perp \rangle \right) \end{aligned}$$

Each factor in the r.h.s. is positive. In the second brace, the mean value (which is multiplied by a positive factor) can be bounded by

$$\langle (n-2) S_x^n S_y^n + S_x \cdot S_y \rangle \leq (n-2) + 1 = n-1$$

and one has

$$\frac{1}{N} \sum_{(x,y)} 2(1 - \cos k \cdot (x-y)) = 2 \sum_{\mu=1}^d (1 - \cos k_\mu)$$

Therefore, the inequality reads

$$\frac{\beta^{-1} (n-1) M^2(H)}{2 \sum_{\mu=1}^d (1 - \cos k_\mu) + HM(H)} \leq \frac{1}{N} \sum_{x,y} e^{ik \cdot (x-y)} \langle S_x^\perp \cdot S_y^\perp \rangle$$

We now take the thermodynamic limit and integrate over k in the Brillouin zone

$$\begin{aligned} \beta^{-1} (n-1) M^2(H) \int \frac{d^d k}{(2\pi)^d} \frac{1}{2 \sum_{\mu=1}^d (1 - \cos k_\mu) + HM(H)} \\ \leq \langle (S_x^\perp)^2 \rangle \leq 1 \end{aligned} \quad (A.7)$$

The result is valid for any integral dimension. Of course, this upper bound is quite loose, but is sufficient to exclude a spontaneous magnetization $M = \lim_{H \rightarrow 0} M(H)$ in dimensions 1 (as expected) and 2 (which is the required theorem). Indeed, the integral of the free propagator with a square mass $HM(H) \leq H$ remains finite as $H \rightarrow 0$ only when the dimension is strictly greater than 2. In dimension 1 or 2, it diverges, so that the magnetization should vanish in order to satisfy the inequality. More precisely, if $d = 1$, the integral is $\frac{1}{2} [HM(H) + H^2 M^2(H)/4]^{-1/2}$. Hence

$$0 \leq M(H) \leq \left(\frac{2\beta}{n-1} \right)^{2/3} H^{1/3} \left(1 + \frac{H^2}{4} \right)^{1/3} \quad (d=1) \quad (A.8)$$

and M vanishes with H , whatever the value of β . This result is not sufficiently strong to reproduce the known linear behaviour of M with respect to H , with a coefficient (susceptibility) diverging as $\beta \rightarrow \infty$. However, it is sufficient to eliminate the possibility of a phase with a nonvanishing order parameter. The interesting case leading to a new result is $d = 2$. The integral is logarithmically divergent as H vanishes, and, for sufficiently small H ,

$$0 \leq M(H) \leq \sqrt{\frac{4\pi\beta}{n-1}} \frac{1}{\sqrt{\text{Cst} - \ln H}} \quad (d=2) \quad (\text{A.9})$$

Again, a spontaneous magnetization is excluded at any temperature. We note that this inequality is uniformly valid both in H and β .

4.A.2 Correlation inequality

The ideas developed in the context of the two-dimensional XY-model have been used by McBryan and Spencer to deduce a polynomial decrease of the correlations in the absence of external field. The proof can be generalized to fields with more than two components; however, the corresponding bound is very weak, since in this case the correlations decrease exponentially, but is of course sufficient to establish the absence of a spontaneous magnetization.

Consider the two-point function, in the two-component case

$$0 \leq \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle = Z^{-1} \text{Re} \int \prod_y d\theta_y \times \exp \left\{ \beta \sum_{(\mathbf{x}', \mathbf{x}'')} \cos(\theta_{\mathbf{x}'} - \theta_{\mathbf{x}''}) + i(\theta_{\mathbf{x}} - \theta_{\mathbf{y}}) \right\} \quad (\text{A.10})$$

As above, we assume that the volume is large, but finite, and use periodic boundary conditions to be specific. For brevity, we denote by $G(\mathbf{x} - \mathbf{y})$ the lattice solution of $-\Delta G(\mathbf{x} - \mathbf{y}) = \delta_{\mathbf{x}, \mathbf{y}}$, with an implicit subtraction which will drop out in the subsequent estimates. One wants to make use of the fact that, for finite N , the integrand in (A.10) is not only periodic in each variable θ , of period 2π , but also analytic. One can therefore displace the contour of integration by setting $\theta_{\mathbf{x}'} \rightarrow \theta_{\mathbf{x}'} + i\psi_{\mathbf{x}'}$, where $\psi_{\mathbf{x}'}$ is

chosen as

$$\psi_{\mathbf{x}'} = \beta^{-1} [G(\mathbf{x}' - \mathbf{x}) - G(\mathbf{x}' - \mathbf{y})] \quad (\text{A.11})$$

The imaginary part $\psi_{\mathbf{x}'}$ is well-defined, and bounded in $\mathbf{x}'$ by Cst/β . After this substitution, we can majorize the phases by unity, and replace (A.10) by the weaker inequality

$$0 \leq \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle \leq Z^{-1} \times e^{-(\psi_x - \psi_y)} \int \prod_y d\theta_y e^{i\beta \sum_{(\mathbf{x}', \mathbf{x}'')} \cos(\theta_{\mathbf{x}'} - \theta_{\mathbf{x}''}) \cosh(\psi_{\mathbf{x}'} - \psi_{\mathbf{x}''})}$$

and hence

$$0 \leq \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle \leq e^{-(\psi_x - \psi_y)} \left\langle e^{\beta \sum_{(\mathbf{x}', \mathbf{x}'')} \cos(\theta_{\mathbf{x}'} - \theta_{\mathbf{x}''}) (\cosh(\psi_{\mathbf{x}'} - \psi_{\mathbf{x}''}) - 1)} \right\rangle$$

If β is sufficiently large, the uniform bound on ψ implies that

$$\beta \sum_{(\mathbf{x}', \mathbf{x}'')} (\cosh(\psi_{\mathbf{x}'} - \psi_{\mathbf{x}''}) - 1) \leq \frac{1}{2} \beta [1 + \mathcal{O}(\beta^{-2})] \sum_{(\mathbf{x}', \mathbf{x}'')} (\psi_{\mathbf{x}'} - \psi_{\mathbf{x}''})^2$$

Moreover,

$$\sum_{(\mathbf{x}', \mathbf{x}'')} (\psi_{\mathbf{x}'} - \psi_{\mathbf{x}''})^2 = - \sum_{\mathbf{x}'} \psi_{\mathbf{x}'} \sum_{\mathbf{x}''(\mathbf{x}')} (\psi_{\mathbf{x}''} - \psi_{\mathbf{x}'}) = \beta^{-1} (\psi_x - \psi_y)$$

The cosine in the exponential is bounded by 1, and hence, for $\beta > \beta(\varepsilon)$,

$$0 \leq \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle \leq \exp \left(-\frac{1}{2} (1 - \varepsilon) (\psi_x - \psi_y) \right)$$

or

$$0 \leq \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle \leq \exp \left(-\frac{1}{2} (1 - \varepsilon) [G(0) - G(\mathbf{x} - \mathbf{y})] / \beta \right) \quad (\text{A.12})$$

The quantity $G(0) - G(\mathbf{x}) = \Gamma(\mathbf{x})/2\pi$ behaves as $\ln(|\mathbf{x}|/r_0)/2\pi$ for large $|\mathbf{x}|$, and is in any case positive. For large separations, one obtains the realistic bound for the XY-model

$$0 \leq \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle \leq \left(\frac{r_0}{|\mathbf{x} - \mathbf{y}|} \right)^{(1-\varepsilon)/2\pi\beta} \quad \beta > \beta(\varepsilon) \quad (\text{A.13})$$

This inequality again excludes the possibility of a spontaneous magnetization, since $\lim_{|\mathbf{x}-\mathbf{y}| \rightarrow \infty} \langle \mathbf{S}_x \cdot \mathbf{S}_y \rangle = \langle \mathbf{S}_x \rangle^2 = 0$, at least at low temperature. The spontaneous magnetization is a nondecreasing function of β and thus always vanishes. The reasoning

can be generalized to $n \geq 3$ components and yields as an upper bound a polynomial decrease.

To summarize, these proofs are related to the impossibility of defining a genuine free massless field in two dimensions, due to the infrared divergence of fluctuations. This idea has been formalized by S. Coleman. This does not prevent one from using such a field, in any finite volume, and even some of its derived quantities such as derivatives or exponentials, in an infinite volume, as will be discussed in chapter 9 (in volume 2).

Appendix 4.B Phenomenological renormalization

In low enough dimension, in practice two or three, it is possible to develop a renormalization method in real space based on a hypothesis of quasiscale invariance for large, but finite systems. We shall return in much greater detail to this subject in chapter 9 (volume 2) devoted to conformal invariance in two-dimensional systems. Let us consider for instance the eigenvalues of the transfer matrix of a system with transverse size L^{d-1} . The logarithm of the largest eigenvalue is the free energy per volume unit, times L^{d-1} (up to a factor). Moreover, the ratio of the two largest eigenvalues (assumed nondegenerate) defines a correlation length, according to

$$\exp(-1/\xi_L) = \lambda_L^{(1)}/\lambda_L^{(2)} \quad (B.1)$$

As $L \rightarrow \infty$, ξ_L tends towards the correlation length of the infinite system ξ_∞ , at least for $\beta < \beta_c$. When the lengths L, ξ_L, ξ_∞ are large with respect to the lattice spacing, one expects a relation

$$\xi_L = \xi_\infty f(L/\xi_\infty) \quad (B.2)$$

where the function $f(x)$ approaches unity when $x \rightarrow \infty$ and ignores the lattice spacing. Conversely, if ξ_∞ increases indefinitely near a critical point, ξ_L remains finite, so that $f(x) \sim Ax$ when x vanishes. In the particular example of the Ising model where ξ_L and ξ_∞ depend on two essential parameters $\theta \sim (T - T_c)/T_c$ and H , one has

$$\frac{\xi_{L'}(\theta', H')}{\xi_L(\theta, H)} = \frac{\xi_\infty(\theta', H')}{\xi_\infty(\theta, H)} \frac{f(L'/\xi_\infty(\theta', H'))}{f(L/\xi_\infty(\theta, H))} \quad (B.3)$$

Let us choose θ' and H' , other parameters being kept fixed, in such a way that

$$\frac{L'}{\xi_\infty(\theta', H')} = \frac{L}{\xi_\infty(\theta, H)} \quad (B.4)$$

In the relation (B.3), the unknown function f can be eliminated and one finds

$$\frac{\xi_{L'}(\theta', H')}{\xi_L(\theta, H)} = \frac{\xi_\infty(\theta', H')}{\xi_\infty(\theta, H)} = \frac{L'}{L} \quad (B.5)$$

As L and L' are large (but finite), this formula establishes a relation between θ', H' and θ, H , depending on the finite lengths L and L' . If the function $\xi_L(\theta, H)$ is known, one gets a renormalization relation. Since ξ_∞ only diverges as $H \rightarrow 0, \theta \rightarrow 0$, the relation between (θ', L') and (θ, L) defined by

$$\frac{\xi_{L'}(\theta', 0)}{\xi_L(\theta, 0)} = \frac{L'}{L} = \frac{\xi_\infty(\theta', 0)}{\xi_\infty(\theta, 0)} \quad (B.6)$$

has a fixed point located at 0, with an asymptotic behaviour

$$\theta' = \left(\frac{L}{L'}\right)^{1/\nu} \theta \quad (B.7)$$

when L and L' tends simultaneously to infinity, with a fixed ratio. The effective correlation length ξ_L increases linearly with L at the fixed point $\theta = 0$. Similarly, we can retrieve the magnetic exponent by studying the behaviour at $\theta = 0$, where $\xi_\infty(0, H)$ diverges as $|H|^{-1/\nu_H}$ with $\nu_H = \frac{1}{2}(d + \frac{2}{\nu})$. The transformation (B.5) has $H = 0$ as a fixed point and generates a behaviour

$$H' = \left(\frac{L}{L'}\right)^{\nu_H} H \quad (B.8)$$

This method has both advantages and defects. On the positive side, it does not require the introduction of any new interaction by renormalization. On the other hand, one has just one equation to determine several variables (here θ' and H'). In the present case, one fortunately knows that the critical point lies at $H = 0$; it is hence possible to determine first the critical temperature, then the magnetic behaviour at this temperature.

The method requires the determination of $\xi_L(\theta, H)$. Hence, up to now, only two- (or perhaps three-) dimensional systems can be studied in practice.

Examine the convergence of the method for the two-dimensional Ising model, using two-dimensional strips of width L with periodic boundary conditions. Using the expressions of chapter 2, show that the correlation length is given by

$$\xi_L^{-1}(\beta) = \frac{1}{2}[(\gamma_1 + \gamma_3 + \dots + \gamma_{2L-1}) - (\gamma_0 + \gamma_2 + \dots + \gamma_{2L-2})]$$

$$\cosh \gamma_r = \frac{\cosh^2 2\beta}{\sinh 2\beta} - \cos \frac{r\pi}{L}$$
(B.9)

Defining $\beta_c(L, L')$ and $\nu(L, L')$ through

$$\frac{\xi_L(\beta_c)}{\xi_{L'}(\beta_c)} = \frac{L'}{L} \equiv u$$

$$\nu^{-1} = \frac{\ln \left(\frac{d\xi_L(\beta_c)}{d\beta} / \frac{d\xi_{L'}(\beta_c)}{d\beta} \right)}{\ln(L/L')} - 1$$
(B.10)

one finds

$$\beta_c(L, L') = \frac{1}{2} \ln(1 + \sqrt{2}) - \frac{\pi^3}{192} \frac{u(u+1)}{L^3} + \dots$$
(B.11)

$$\nu(L, L') = 1 - \frac{\pi^2 \ln 2}{24L^2} \frac{(u^2 - 1)}{\ln u} + \dots$$
(B.12)

This shows that even with strips of small width, one can get sensible results. The optimal choice of u lies near 1, that is $L = L' \pm 1$.

The technique can be applied to various two-dimensional problems, such as percolation, localization, XY-model, ... Equation (B.2) and its generalization to other observables needs a justification which can be given in the framework of the renormalization group, at least in dimension less than or equal to four. Indeed this relation is not valid any more in higher dimensions (in general above the upper critical dimension). This is to be expected, since the mean field approximation predicts a transition ignoring fluctuations, even for a finite size system.

We illustrate the technique on the spherical model studied in chapter 3, i.e. for the $n \rightarrow \infty$ limit of the n -vector model. We saw

that the correlation length of an infinite system satisfies

$$\beta = \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\xi_\infty^{-2} + 2 \sum_{\mu=1}^d (1 - \cos q_\mu)} \quad (B.13)$$

In a strip of width L , transverse components of the momentum take discrete values which are multiples of $2\pi/L$. Using Poisson's relation, a formula analog to (B.13) can be written for the correlation length of this system

$$\beta = \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\xi_L^{-2} + 2 \sum_{\mu=1}^d (1 - \cos q_\mu)} \sum_{\mathbf{m}_\perp} (2\pi)^{d-1} \delta^{(d-1)}(\mathbf{q}_\perp - \frac{2\pi}{L} \mathbf{m}_\perp)$$

$$= \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\xi_L^{-2} + 2 \sum_{\mu=1}^d (1 - \cos q_\mu)} \sum_{\mathbf{n}_\perp} e^{i\mathbf{q}_\perp \cdot \mathbf{n}_\perp L}$$
(B.14)

where the $(d-1)$ -dimensional vectors $\mathbf{m}_\perp$ and $\mathbf{n}_\perp$ have integer coordinates. Isolating in (B.14) the term $\mathbf{n}_\perp = 0$ and subtracting (B.13), one obtains

$$(\xi_\infty^{-2} - \xi_L^{-2}) \int \frac{d^d \mathbf{q}}{(2\pi)^d} \left(\sum'_{\mathbf{n}_\perp} \frac{e^{i\mathbf{q}_\perp \cdot \mathbf{n}_\perp L}}{\xi_L^{-2} + 2 \sum_{\mu=1}^d (1 - \cos q_\mu)} + \frac{1}{[\xi_\infty^{-2} + 2 \sum_{\mu=1}^d (1 - \cos q_\mu)] [\xi_L^{-2} + 2 \sum_{\mu=1}^d (1 - \cos q_\mu)]} \right) = 0$$
(B.15)

In the limit $\beta \rightarrow \beta_c$ and $L \rightarrow \infty$, ξ_∞ increases indefinitely and ξ_L is large. Under these conditions, one can replace $2 \sum_{\mu=1}^d (1 - \cos q_\mu)$ by $\mathbf{q}^2$, when the integrals are dominated by small momenta $\mathbf{q}$. This is indeed the case for $2 < d < 4$. This approximation gives, for the second term,

$$\int \frac{d^d \mathbf{q}}{(2\pi)^d} \sum'_{\mathbf{n}_\perp} \frac{e^{i\mathbf{q}_\perp \cdot \mathbf{n}_\perp L}}{\xi_L^{-2} + \mathbf{q}^2} = L^{2-d} \int_0^\infty dt e^{-t(L/\xi_L)^2} u(t) \quad (B.16)$$

We use the notation

$$u(t) = \sum'_{\mathbf{n}_\perp} \frac{e^{-\mathbf{n}_\perp^2/4t}}{(4\pi t)^{d/2}} \sim \begin{cases} \frac{2(d-1)}{(4\pi t)^{d/2}} e^{-1/4t} & t \rightarrow 0 \\ \frac{1}{(4\pi t)^{1/2}} & t \rightarrow \infty \end{cases} \quad (B.17)$$

The small and large t behaviour explicitly shows the convergence of the integral (B.16). With a similar approximation for the first integral in (B.15), one finds, in the critical region and for large L ,

$$\left[\left(\frac{L}{\xi_\infty} \right)^2 - \left(\frac{L}{\xi_L} \right)^2 \right] \left(\frac{\xi_L}{L} \right)^{4-d} \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} \times \int_0^1 dt \left[t \left(\frac{\xi_L}{\xi_\infty} \right)^2 + (1-t) \right]^{d/2-2} + \int_0^\infty dt e^{-t(L/\xi_L)^2} u(t) = 0 \quad (B.18)$$

This relation, which is of the form $\varphi(L/\xi_L, L/\xi_\infty) = 0$, can be solved as $\xi_L/\xi_\infty = f(L/\xi_\infty)$. The resulting function f has the required properties. In particular, at the critical point where ξ_∞ diverges, $\xi_L(\beta_c)$ grows linearly with L , according to

$$\xi_L(\beta_c) = sL \quad (B.19)$$

and s is given by the implicit equation

$$s^{2-d} \left\{ \frac{-\Gamma(1-d/2)}{(4\pi)^{d/2}} \right\} = \int_0^\infty dt e^{-t/s^2} u(t) \quad (B.20)$$

In particular, one finds in three dimensions

$$\sum'_{n_1, n_2} \frac{s}{\sqrt{n_1^2 + n_2^2}} \exp \left(-\frac{\sqrt{n_1^2 + n_2^2}}{s} \right) = 1 \quad s \simeq 0.539 \quad (d=3) \quad (B.21)$$

When the dimension d approaches 4, the l.h.s. of (B.20) develops a pole and s increases. As the integral diverges in the large t domain, one can replace $u(t)$ by its large t behaviour. Hence, with $d = 4 - \epsilon$,

$$\xi_L(\beta_c) = \left(\frac{1}{4\pi^2 \epsilon} \right)^{1/3} L \quad (d = 4 - \epsilon) \quad (B.22)$$

Above four dimensions, relation (B.2) is no more valid. Approximate scale invariance is broken and the cutoff factor, here the lattice spacing, plays a role. In equation (B.15), it is not possible anymore to replace the denominators $2 \sum_1^d (1 - \cos q_\mu)$ by q^2 , since this operation introduces ultraviolet divergences. For example, when $d = 4$, one finds

$$\left(\frac{L}{\xi_L} \right)^2 \ln \xi_L^2 - \left(\frac{L}{\xi_\infty} \right)^2 \ln \xi_\infty^2 = (4\pi)^2 \int_0^\infty dt e^{-t(L/\xi_L)^2} u(t) \quad (d=4) \quad (B.23)$$

At the critical temperature, the linear growth of $\xi_L(\beta_c)$ is modified by a logarithmic factor

$$\xi_L(\beta_c) = L \left(\frac{\ln L}{4\pi^2} \right)^{1/3} \quad (d=4) \quad (B.24)$$

Finally, as $d > 4$, there is no infrared divergence in the first integral (B.15) as $\xi_\infty^{-1} \rightarrow 0$ and $\xi_L^{-1} \rightarrow 0$. One can then replace these quantities by 0 in the integrand. Introducing the value of the Green function at the origin

$$g = \int \frac{d^d q}{(2\pi)^d} \frac{1}{\left[2 \left(1 - \sum_{\mu=1}^d \cos q_\mu \right) \right]^2} \quad (d > 4) \quad (B.25)$$

one obtains the following relation

$$\left(\frac{L}{\xi_\infty} \right)^2 = \left(\frac{L}{\xi_L} \right)^2 - \frac{1}{2g} \left(\frac{\xi_L}{L} \right)^{4-d} \quad (d > 4) \quad (B.26)$$

This is of the form

$$\frac{\xi_L}{\xi_\infty} = f \left(\frac{L}{\xi_\infty} L^{(d-4)/3} \right) \quad (d > 4) \quad (B.27)$$

In particular, for $\beta = \beta_c$, $\xi_L(\beta_c)$ grows with L faster than linearly

$$\xi_L(\beta_c) = L(2g L^{d-4})^{1/3} \quad (d > 4) \quad (B.28)$$

Finally, as $d \rightarrow \infty$, $\xi_L(\beta_c)$ diverges, as is suggested by the mean field approximation.

(1) Show that, for the spherical model, in the broken symmetry phase, the behaviour is given by

$$\xi_L(\beta) = 2(\beta - \beta_c) L^{d-1} \quad \beta > \beta_c \quad (B.29)$$

(2) Conversely, for a system with a discrete symmetry, it is expected that ξ_L grows exponentially with L for $\beta > \beta_c$. Why?

(3) Is the ratio $s = \xi_L(\beta_c)/L$ universal? Estimate this ratio for two-dimensional models.

As we shall see later, if a two-dimensional critical theory is conformally invariant, it is possible to find the form of the correlation function for various geometries, as for instance strips with periodic boundary conditions. One can use a conformal mapping of the complex plane $z = x + iy$ on the strip $w = w_1 + iw_2$, $-\frac{1}{2}L < w_1 < \frac{1}{2}L$ of the form

$$z = \exp -2i\pi w/L$$

$$z_a - z_b = -2i \sin \left(\pi \frac{w_a - w_b}{L} \right) \exp \left(-i \pi \frac{w_a + w_b}{L} \right)$$

This mapping is unique, up to a global transformation of the plane $z \rightarrow \alpha z + \beta$ which only modifies the correlation function by a multiplicative factor

$$\langle \varphi(x_a) \varphi(x_b) \rangle = \frac{\text{Cst}}{|x_a - x_b|^\eta} = \frac{\text{Cst}}{|z_a - z_b|^\eta}$$

If we choose in the strip $w_a = it$ (with real positive t), $w_b = 0$, one finds

$$\langle \varphi(t) \varphi(0) \rangle_{\text{strip}} \sim \frac{\text{Cst}}{|\exp(2\pi t/L) - 1|^\eta} \underset{t \rightarrow \infty}{\sim} e^{-2\pi \nu t/L} \quad (B.30)$$

Therefore, in two dimensions, $\xi_L(\beta_c) = L/2\pi\eta$ indeed grows linearly with L at the critical point, with a universal coefficient $s = 1/2\pi\eta$.

Notes

Ideas of scale invariance were developed by B. Widom, *J. Chem. Phys.* **43**, 3892, 3898 (1965); A.Z. Patashinskii and V.L. Pokrovskii, *Zh. Eksp. Teor. Fiz.* **50**, 439 (1966); L.P. Kadanoff, *Physics* **2**, 263 (1966), *Rev. Mod. Phys.* **39**, 395 (1967); M.E. Fisher, *Rept. Prog. Phys.* **30**, 615 (1967); K.G. Wilson, *Phys. Rev. D* **2**, 1473, 1478 (1970), *Phys. Rev. B* **4**, 3174, 3184 (1971); F.J. Wegner, *Phys. Rev. B* **5**, 4529 (1972) and *Lecture Notes in Physics* **37**, 171 (1973) Springer Verlag, Berlin.

A general view on critical phenomena prior to the explosion of the renormalization group is found in H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* Clarendon Press, Oxford (1971). Further important references will be quoted after the next chapter.

The q -state model is due to R.B. Potts, *Proc. Camb. Phil. Soc.* **48**, 106 (1952). The corresponding algebraic aspects were developed by H.N.V. Temperley and E.H. Lieb, *Proc. Roy. Soc. London A* **322**, 251 (1971).

For real space renormalization methods, see L. Kadanoff and A. Houghton, *Phys. Rev. B* **11**, 377 (1975); Th. Niemeijer and J.M.J. Van Leeuwen in Domb and Green vol. VI (1976); A.A. Migdal, *Z. Eksp. Teor. Fiz.* **69**, 810, 1457 (1975); L.P. Kadanoff, *Ann. Phys.* **100**, 359 (1976). Combinations of real space renormal-

ization and Monte-Carlo methods are treated in the volume *Real-space Renormalization*, T.W. Burkhardt and J.M.J. Van Leeuwen eds, Topics in Current Physics **30**, Springer Verlag, Berlin (1982). Julia sets of zeroes for partition functions on hierarchical lattices are taken from a joint work with B. Derrida, L. de Sèze and J.-M. Luck and are reviewed in *Critical Phenomena* 1983 Brasov school, V. Ceausescu *et al* eds, Birkhäuser, Boston (1985). Phenomenological renormalization discussed in the appendix originates in the work of M.E. Fisher and M.N. Barber, *Phys. Rev. Lett.* **28**, 1516 (1972) on finite size scaling and was applied by many authors following M.P. Nightingale, *Physica* **83A**, 561 (1976). For a field theoretical treatment, see E. Brézin, *J. Physique* **43**, 15 (1982).

The absence of spontaneous magnetization for two-dimensional systems with continuous symmetries is proved by P.C. Hohenberg, *Phys. Rev.* **158**, 383 (1967); N.D. Mermin and H. Wagner, *Phys. Rev. Lett.* **17**, 1133 (1966); N.D. Mermin, *Journ. Math. Phys.* **8**, 1061 (1967). The pathologies of massless two dimensional scalar fields as well as their positive aspects are studied by S. Coleman, *Comm. Math. Phys.* **31**, 259 (1973) and *Phys. Rev. D* **11**, 2088 (1975). Part of appendix A is based on the work of O. McBryan and T. Spencer, *Comm. Math. Phys.* **53**, 299 (1977).

The role of vortices in the XY-system was discussed by V.L. Berezinskii, *Soviet Physics JETP*, **32**, 493 (1971) and the nature of the transition by J.M. Kosterlitz, *J. Phys. C* **7**, 1046 (1974). Our discussion is patterned after the work of J.V. José, L.P. Kadanoff, S. Kirkpatrick and D.R. Nelson, *Phys. Rev. B* **16**, 1217 (1977) except for minor modifications. The spin wave plus Coulomb gas version is due to J. Villain, *J. Physique* **36**, 581 (1975). A systematic field theoretic analysis in the framework of the sine-Gordon model is provided by D.J. Amit, Y.Y. Goldschmidt and G. Grinstein, *J. Phys. A* **13**, 585 (1980), with references to previous works.

Interestingly, the same XY-model can be interpreted in various ways which allow experimental checks. It is in the same universal class as the solid on solid model of W.K. Burton, N. Cabrera and F.C. Franck, *Phil. Trans. Roy. Soc. A* **243**, 299 (1951) which describes the transition from a flat to a rough interface. As discussed in the text, it is also related to the dielectric plasma transition of the classical Coulomb gas; see S.T. Chui and

J.D. Weeks, *Phys. Rev.* **B14**, 4978 (1976); J.D. Weeks in *Ordering in Strongly Fluctuating Condensed Matter Systems*, T. Riste ed., Plenum Press, New York (1980). One can also mention the theory of superconducting films discussed for instance by B.I. Halperin and D.R. Nelson, *J. Low Temp. Phys.* **36**, 599 (1979) or the universal jump in the He^4 superfluid density, D.R. Nelson and J.M. Kosterlitz, *Phys. Rev. Lett.* **39**, 1201 (1977). The list is not limitative.

5

CONTINUOUS FIELD THEORY AND THE RENORMALIZATION GROUP

Continuous field models were originally systematically elaborated (and still are) in the context of particle physics. Most relevant to the study of critical phenomena is the derivation of a renormalization group flow, characterized by a set of Callan-Symanzik equations for Green functions, with regular coefficients derived from perturbation theory. This is supplemented by the idea due to Fisher and Wilson of an expansion in powers of the deviation from the strict renormalizability dimension, i.e. four in the case of the φ^4 model, but can also be presumed to work directly in the physical dimension three (and even possibly two). We devote this chapter to a general presentation and a survey of some applications. An appendix gives a short introduction to multicritical phenomena.

5.1 The Lagrangian and dimensional analysis

5.1.1 Introduction

We want to investigate universal critical properties. A discrete lattice model appears as a regularizing intermediate stage, which allows a precise meaning to be given to the functional integrals of the field theory. Following the scheme suggested by the mean field approximation, one is tempted to start directly from a continuous model, in which the lattice is replaced by a d -dimensional continuous Euclidean space, and also, even for models with a discrete symmetry, the dynamical variables are replaced by continuous fields. From the original formulation, we just retain the idea of a cutoff factor Λ large with respect to all momenta, which ensures that some possibly divergent expressions remain finite.

The (positive) statistical weight of a configuration is the exponential of an action which is the integral over space of a *local*

J.D. Weeks, *Phys. Rev.* **B14**, 4978 (1976); J.D. Weeks in *Ordering in Strongly Fluctuating Condensed Matter Systems*, T. Riste ed., Plenum Press, New York (1980). One can also mention the theory of superconducting films discussed for instance by B.I. Halperin and D.R. Nelson, *J. Low Temp. Phys.* **36**, 599 (1979) or the universal jump in the He^4 superfluid density, D.R. Nelson and J.M. Kosterlitz, *Phys. Rev. Lett.* **39**, 1201 (1977). The list is not limitative.

5

CONTINUOUS FIELD THEORY AND THE RENORMALIZATION GROUP

Continuous field models were originally systematically elaborated (and still are) in the context of particle physics. Most relevant to the study of critical phenomena is the derivation of a renormalization group flow, characterized by a set of Callan-Symanzik equations for Green functions, with regular coefficients derived from perturbation theory. This is supplemented by the idea due to Fisher and Wilson of an expansion in powers of the deviation from the strict renormalizability dimension, i.e. four in the case of the φ^4 model, but can also be presumed to work directly in the physical dimension three (and even possibly two). We devote this chapter to a general presentation and a survey of some applications. An appendix gives a short introduction to multicritical phenomena.

5.1 The Lagrangian and dimensional analysis

5.1.1 Introduction

We want to investigate universal critical properties. A discrete lattice model appears as a regularizing intermediate stage, which allows a precise meaning to be given to the functional integrals of the field theory. Following the scheme suggested by the mean field approximation, one is tempted to start directly from a continuous model, in which the lattice is replaced by a d -dimensional continuous Euclidean space, and also, even for models with a discrete symmetry, the dynamical variables are replaced by continuous fields. From the original formulation, we just retain the idea of a cutoff factor Λ large with respect to all momenta, which ensures that some possibly divergent expressions remain finite.

The (positive) statistical weight of a configuration is the exponential of an action which is the integral over space of a *local*

Lagrangian. The finiteness of the interaction range is replaced by the notion of *locality*, which means in practice the introduction of as limited a number of derivative terms as is possible. The Lagrangian may be considered as the relevant leading term in an expansion allowing the calculation of fluctuation effects around a mean field solution. Alternatively, it can be postulated *a priori* to arise from averages over domains large with respect to the microscopic scale, but small compared to the correlation lengths when they increase indefinitely.

These considerations establish a connection with continuous field theory. Conversely, using this analogy with statistical mechanics in the Euclidean domain, one can have access to global properties. In a first stage, we limit ourselves to a scalar model with Z_2 symmetry ($n = 1$), similar to the Ising model. Denoting by φ the fluctuating field, the Lagrangian, invariant under the substitution $\varphi \rightarrow -\varphi$, is expanded as

$$\mathcal{L}(\varphi) = \frac{1}{2}(\partial\varphi)^2 + \frac{1}{2}m_0^2\varphi^2 + \frac{1}{4!}g_0\varphi^4 + \dots \quad (1)$$

A multiplicative factor has been absorbed in the definition of the field φ , so that the coefficient of the derivative term (*kinetic term*) is normalized to $\frac{1}{2}$. Note the Euclidean invariance of the various terms, when φ is assumed to behave as a scalar field. Omitted terms contain higher powers or more derivatives of φ . We will have to justify the fact that they are indeed negligible in the critical domain.

In this construction, the coefficients m_0^2 and g_0 are regular functions of the basic parameters, in particular the temperature. By tuning these parameters, it is possible to increase indefinitely the correlation length. The dependence of the latter on the variable m_0^2 thus yields a linear temperature scale near the critical point and allows us to obtain the critical exponent ν .

Given the Lagrangian (1), the partition function Z is expressed as a functional integral

$$Z = \int \mathcal{D}\varphi \exp\left(-\int d^d x \mathcal{L}(\varphi)\right) \quad (2)$$

analogous to the corresponding quantity in the discrete case. This symbolic expression acquires a precise meaning, using a mathematical construction, for an integral dimension less than or equal to three. As we shall see below, the essential reason is

that a finite number of subtractions is sufficient to compensate the short distance singularities and to make the theory finite. From the present point of view, however, we assume the existence of a microscopic structure, characterized for instance by the lattice spacing, and the continuous theory is used to describe phenomena at large distances. Hence, possible divergences can be avoided from a physical point of view by introducing in the integrations over the momenta a cutoff factor $\Lambda \sim a^{-1}$. The renormalization theory will organize the potentially divergent expressions into a finite number of factors relating the macroscopic theory to the microscopic parameters. The exploitation of these relations leads to the renormalization group and finally allows the calculation of the critical behaviour. One should therefore consider that, besides the parameters m_0^2 and g_0 explicitly written in the Lagrangian, the model depends on the regularization procedure and on its cutoff parameter Λ .

When one neglects fluctuation effects, the Lagrangian (1) describes a second order transition in terms of the parameter m_0^2 . In the high temperature symmetric phase ($m_0^2 > 0$), the field fluctuates around a zero value. When $m_0^2 < 0$ (with of course $g_0 > 0$), there are fluctuations around a nonzero value of φ (spontaneous magnetization), corresponding to the minimization of $\frac{1}{2}m_0^2\varphi^2 + g_0\varphi^4/4!$, i.e. $M^2 \equiv \langle\varphi\rangle^2 = -6m_0^2/g_0$. It is easily seen that this corresponds to the mean field description studied in previous chapters.

Fluctuations induce a modification of critical properties when the dimension is less or equal to four. Wilson and Fisher have suggested an expansion in a parameter $\epsilon = 4 - d$, assumed small and positive. Continuity (and even more differentiability) properties of the critical exponents as functions of ϵ should be presupposed. This requires a precise definition of the dimensional continuation, an example of which has been given for the spherical model in chapter 3.

In the sequel, we shall use extensively a formal perturbative method in the coupling constant g_0 in order to justify the renormalizability properties, even though we cannot assume at first that it is a small parameter. As we shall see later, this is not true in the four-dimensional case. The only fully soluble model uses Gaussian fluctuations, which provide a first approximation to the deviations to mean field. For instance, if, in the high

temperature phase, one neglects g_0 and takes m_0^2 proportional to the deviation θ with respect to the critical temperature, the Fourier transform of the two point correlation function is

$$G_0(\mathbf{p}, \theta) = \int d^d \mathbf{x} e^{i\mathbf{p} \cdot \mathbf{x}} \langle \varphi(\mathbf{x}) \varphi(0) \rangle = \frac{1}{\mathbf{p}^2 + \theta} = \frac{1}{\theta} \left(\frac{1}{1 + \mathbf{p}^2/\theta} \right) \quad (3)$$

This can be rewritten

$$G_0(\mathbf{p}, \theta) = \theta^{-\gamma} g(p\theta^{-\nu}) \quad (4)$$

with

$$\gamma = 1 \quad \nu = \frac{1}{2} \quad (5)$$

being the critical exponents given by the mean field approximation. The susceptibility is indeed proportional to the correlation function at zero momentum. The function g , regular at the origin, has an argument of the form $p\xi$, ξ being the correlation length proportional to $\theta^{-\nu}$. Hence the study of the 2-point correlation function allows one to obtain two critical exponents, and the scaling properties – which will be justified in the framework of the continuous model – then allow the calculation of all other exponents. The effects of anharmonic terms for $g_0 > 0$ have now to be studied.

5.1.2 Generating functionals and dimensional analysis

The relative dimensionless probabilities are written $\exp(-S)$, where the action S , which is also dimensionless, is the energy divided by kT in the thermodynamical language. Dimensions are conventionally expressed in momentum units (to be more precise, in wavenumber units; the practice of using the language of quantum mechanics with $\hbar = 1$ is, however, in general use). Correspondingly, lengths have dimension -1 , so that the dimension of the Lagrangian is

$$[\mathcal{L}] = d \quad (6)$$

The coefficient of the kinetic term being a pure number, one is led to assign to the field a dimension

$$[\varphi] = \frac{1}{2}(d-2) \quad (7)$$

In particular, φ is dimensionless in space dimension 2 (as are the angles in the XY -model) and of dimension 1 for $d = 4$.

As a consequence,

$$[m_0] = 1 \quad [g_0] = 4 - d \quad (8)$$

More generally, the coupling constant g_{n_1, n_2} of an additional term $\partial^{n_2} \varphi^{n_1}$ has a dimension

$$[g_{n_1, n_2}] = -\delta_{n_1, n_2} = d - \frac{1}{2}n_1(d-2) - n_2 \quad (9)$$

The number δ_{n_1, n_2} is called the *canonical degree* of the corresponding operator. It should of course not be confused with the usual notation for the Kronecker symbol. It is common usage, by reference to quantum field theory, to speak of operator for a local quantity built out of fields and their derivatives, even though in the statistical context they might more naturally be referred to as local observables. The degree given by equation (9) is 2 in dimension two when n_2 vanishes, whatever n_1 is. In space dimension four, the terms explicitly written in the Lagrangian (1) are the only ones with a nonpositive degree. In dimension three, there is only one additional term φ^6 satisfying both Euclidean invariance and $\varphi \rightarrow -\varphi$ symmetry. However, this term is dominated by the φ^4 term for small fluctuations of φ around 0, except for the case where g_0 vanishes. The inclusion of such a term allows one to study a change from a second to a first order transition. This *tricritical* phenomenon is discussed in appendix A.

The microscopic scale parameter is the cutoff factor $\Lambda \sim 1/a$. Operators with a positive degree behave as $\Lambda^{-\delta}$ and seem therefore negligible as $\Lambda \rightarrow \infty$. This naive view will nevertheless have to be slightly corrected. For example, terms with a positive degree obviously play a role in the determination of the critical temperature, which is not universal. But the universal dominant critical behaviour is not sensitive to these operators. Hence, at least near $d = 4$, the large distance behaviour is governed by the φ^4 Lagrangian (1), with at most a redefinition of the *bare* parameters m_0^2 and g_0 . We will show indeed that renormalization acts as a filter, suppressing the irrelevant terms (with a positive degree) and keeping only a minimal number of relevant parameters.

This discussion which justifies the use of renormalizable local Lagrangians for the description of critical phenomena is quite natural in particle physics. Indeed, we have as yet no knowledge

of very high energy processes (i.e. at very small distances). The ability to construct efficient models describing phenomena at a given scale is based on a quasi-decoupling between this scale and those of much higher energies. The latter only yields the values of some parameters which cannot be determined in an effective low energy theory as, for instance, the fine structure constant of electrodynamics.

Correlation functions (or Green functions, or also Schwinger functions in this context) are obtained by coupling φ to an external source $h(\mathbf{x})$, adding to the Lagrangian a term

$$\mathcal{L}_{\text{source}} = -h(\mathbf{x})\varphi(\mathbf{x}) \quad (10)$$

The dimension of the external field $h(\mathbf{x})$ is

$$[h] = \frac{1}{2}(d+2) \quad (11)$$

With the additional source term (10), the partition function becomes a generating functional $Z(h)$ for the correlation functions

$$\begin{aligned} Z(h) &= \int \mathcal{D}\varphi \exp \left(- \int d^d \mathbf{x} \mathcal{L}(\varphi) + \int d^d \mathbf{x} h\varphi \right) \\ &= Z(0) \left\langle \exp \left(\int d^d \mathbf{x} h\varphi \right) \right\rangle \end{aligned} \quad (12)$$

The n -field (n -point) correlation functions are defined as

$$\langle \varphi(\mathbf{x}_1) \dots \varphi(\mathbf{x}_n) \rangle_h = Z(h)^{-1} \frac{\delta^n}{\delta h(\mathbf{x}_1) \dots \delta h(\mathbf{x}_n)} Z(h) \quad (13)$$

When $h \rightarrow 0$ and in the symmetric phase, only correlations involving an even number of points do not vanish. They are translationally invariant, and we will denote them by

$$\begin{aligned} \langle \varphi(\mathbf{x}_1) \dots \varphi(\mathbf{x}_n) \rangle &= Z(h)^{-1} \frac{\delta^n}{\delta h(\mathbf{x}_1) \dots \delta h(\mathbf{x}_n)} Z(h) \Big|_{h=0} \\ &= \int \prod_1^n \left(\frac{d^d \mathbf{p}_k}{(2\pi)^d} e^{-i\mathbf{p}_k \cdot \mathbf{x}_k} \right) (2\pi)^d \delta(\sum \mathbf{p}_k) G_n(\mathbf{p}_1, \dots, \mathbf{p}_n) \end{aligned} \quad (14)$$

Conversely,

$$Z(h)/Z(0) = \sum_{n=0}^{\infty} \frac{1}{n!} \int d^d \mathbf{x}_1 \dots d^d \mathbf{x}_n h(\mathbf{x}_1) \dots h(\mathbf{x}_n) \langle \varphi(\mathbf{x}_1) \dots \varphi(\mathbf{x}_n) \rangle \quad (15)$$

The momentum conservation distribution $\delta(\sum \mathbf{p})$ has been extracted, so that in momentum space G_n is defined on the linear manifold $\sum \mathbf{p} = 0$. Its canonical dimension is

$$[G_n] = d - n[h] = d - \frac{1}{2}n(d+2) \quad (16)$$

In particular, $[G_2] = -2$ in any dimension, as in the free field example (3).

The free energy $\ln Z(h)$ (for convenience, our definition differs by a factor $-1/\beta$ from usual thermodynamics) is the generating functional for the *connected* correlation functions. In the zero field limit,

$$\langle \varphi(\mathbf{x}_1) \dots \varphi(\mathbf{x}_n) \rangle_c = \frac{\delta^n}{\delta h(\mathbf{x}_1) \dots \delta h(\mathbf{x}_n)} \ln Z(h) \Big|_{h=0} \quad (17)$$

These connected functions have of course the same dimension as the general correlation functions. The connection between the two types is given by *cumulants* which allow the exponentiation of the expansion (15). The terminology comes from the fact that only connected diagrams appear in the perturbative expansion. These results will be recalled in chapter 7 (volume 2), and we assume in this chapter that they are familiar to the reader. The free energy is an extensive quantity (proportional to the volume) when fields are defined on a (large) finite domain.

From the free energy $\ln Z(h)$, one obtains the generating functional of *one particle irreducible* Green functions (or *vertex functions*) $\Gamma(\varphi)$ by applying a Legendre transformation. The corresponding variable (spatially varying magnetization) is commonly denoted by the same symbol φ as the integration field variable in equation (2). In the perturbative theory, the contributing diagrams cannot be disconnected by cutting an internal line, and propagators corresponding to the external lines are omitted (*amputation*). This is the origin of the qualification *one-particle irreducible*. The quantity $\Gamma(\varphi)$ is defined through

$$\Gamma(\varphi) = \sup_h \left\{ \int d^d \mathbf{x} h(\mathbf{x}) \varphi(\mathbf{x}) - \ln Z(h) \right\} \quad (18)$$

In practice, one only checks the extremality condition, so that the transformation

$$\Gamma(\varphi) + \ln Z(h) = \int d^d x h(x) \varphi(x) \quad (19)$$

$$\frac{\delta \ln Z(h)}{\delta h(x)} = \varphi(x) = \langle \varphi(x) \rangle_h$$

is reciprocal.

- (1) Obtain the Legendre transformation for a Gaussian integral.
- (2) Prove that the Legendre transformation (19) is reciprocal

$$\frac{\delta \Gamma(\varphi)}{\delta \varphi(x)} = h(x)$$

A nonzero limit of $\delta \ln Z(h)/\delta h$ as h becomes uniformly small with a fixed sign signals a spontaneous breaking of the Z_2 symmetry. The corresponding value of φ is the (uniform) spontaneous magnetization which occurs in the low temperature phase.

When h tends to be uniform, the free energy appears as the integral of a constant density. The same is true for $\Gamma(\varphi)$ when φ is uniform, and we write in this case

$$\Gamma(\varphi) = \int d^d x V(\varphi) \quad (20)$$

The effective potential $V(\varphi)$, which takes into account the fluctuations, is a generalization of the corresponding term $\frac{1}{2}m_0^2\varphi^2 + g_0\varphi^4/4!$ in the Lagrangian (1).

Returning to the general case, the vertex functions are obtained by expanding $\Gamma(\varphi)$ in increasing powers of φ , as one does with $\ln Z(h)$ to get the connected functions

$$\Gamma(\varphi) = \sum \frac{1}{n!} \int d^d x_1 \dots d^d x_n \varphi(x_1) \dots \varphi(x_n) \Gamma_n(x_1, \dots, x_n)$$

$$\Gamma_n(x_1, \dots, x_n) = \left. \frac{\delta^n \Gamma(\varphi)}{\delta \varphi(x_1) \dots \delta \varphi(x_n)} \right|_{\varphi=0}$$

$$= \int \prod_1^n \left(\frac{d^d p_k}{(2\pi)^d} e^{-i p_k \cdot x_k} \right) (2\pi)^d \delta(\sum p_k) \Gamma_n(p_1, \dots, p_n) \quad (21)$$

Functions and their Fourier transforms share abusively the same notation. In the expansion (21) around $\varphi = 0$, only terms with n even do not vanish. This is no longer the case if $\Gamma(\varphi)$ is expanded around another value (for example the spontaneous magnetization of the ordered phase).

The dimension of the argument φ of Γ is the same as the one of the field φ , i.e. $\frac{1}{2}(d-2)$, while $\Gamma(\varphi)$ is of course dimensionless, and hence

$$[\Gamma_n(x_1, \dots, x_n)] = \frac{1}{2}n(d+2) \quad (22)$$

$$[\Gamma_n(p_1, \dots, p_n)] = d - \frac{1}{2}n(d-2) = [G_n(p_1, \dots, p_n)] + 2n$$

The relation between the dimensions of the vertex functions and of the Green functions corresponds to the amputation of the external line factors, as can be seen by making the relation (19) explicit. In particular, $[\Gamma_2(p)] = 2$, $[\Gamma_4(p)] = 4 - d$. More generally, by comparing relations (9) and (22), one sees that vertex functions generalize the coupling constants of the Lagrangian.

It is also possible to introduce other source terms analogous to (10), coupled to more complex combinations than the field φ . In particular, as the control parameter is m_0^2 , one is naturally led to consider a source term coupled to $\frac{1}{2}\varphi^2$. The corresponding Green functions *with insertion of φ^2* reads

$$\langle \varphi(x_1) \dots \varphi(x_n) \frac{1}{2}\varphi^2(y_1) \dots \frac{1}{2}\varphi^2(y_r) \rangle = \int (2\pi)^d \delta(\sum p_k + \sum q_l)$$

$$G_{n,r}(p_k, q_l) \prod_1^n \left(\frac{d^d p_k}{(2\pi)^d} e^{-i p_k \cdot x_k} \right) \prod_1^r \left(\frac{d^d q_l}{(2\pi)^d} e^{-i q_l \cdot y_l} \right)$$

$$[G_{n,r}(p, q)] = [G_n] - 2r = d - \frac{1}{2}n(d+2) - 2r \quad (23)$$

The same dimensional considerations also apply to the connected functions. In particular, the specific heat is related to the connected function $G_{0,2}$ at zero momentum, the dimension of which is $d-4$. Vertex functions with insertions of φ^2 are defined by a Legendre transformation on the source of the φ field only. Their Fourier transforms $\Gamma_{n,r}(p, q)$ have a dimension

$$[\Gamma_{n,r}] = [G_n] - 2r = d - \frac{1}{2}n(d-2) - 2r \quad (24)$$

In all these expressions, the dimensions are called *canonical*, in order to distinguish them from the *anomalous* or *dynamical*

dimensions arising from nontrivial critical properties. The original cutoff factor Λ introduced by regularization will be absorbed by the renormalization theory into suitable large scale quantities.

We have only presented here the simplest Lagrangian model corresponding to the scalar φ^4 theory. The generalization to fields taking their values in a vector or matrix space, or a Grassmann algebra, does not present conceptual difficulties. Technical subtleties may however arise in a proper definition of the path integral measure, and this is the source of possible anomalies in the implementation of various symmetries.

5.2 The perturbative method

5.2.1 Diagrammatic series

The free field Gaussian model gives an accurate description of the critical behaviour when $d > 4$. Near $d = 4$, it is therefore natural, despite some difficulties which will indeed arise, to perform a perturbative expansion with respect to the anharmonic terms of the Lagrangian, i.e. in powers of the bare coupling constant g_0 . The analysis of this expansion will provide some clues about how to derive general properties. Let us split the Lagrangian and the action according to

$$\begin{aligned} \mathcal{L}_0 &= \frac{1}{2} \{ (\partial\varphi)^2 + m_0^2 \varphi^2 \} & S_0 &= \int d^d x \mathcal{L}_0(\varphi) \\ \mathcal{L}_I &= \frac{1}{4!} g_0 \varphi^4 & S_I &= \int d^d x \mathcal{L}_I(\varphi) \end{aligned} \quad (25)$$

and expand the generating functional in a power series in the bare coupling constant g_0 as

$$\begin{aligned} Z(h) &= \int \mathcal{D}\varphi \exp - \left\{ S_0 + S_I - \int d^d x h\varphi \right\} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \mathcal{D}\varphi \left(\int d^d x \mathcal{L}_I(\varphi) \right)^n \exp - \left\{ S_0 - \int d^d x h\varphi \right\} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\int d^d x \mathcal{L}_I \left(\frac{\delta}{\delta h} \right) \right)^n \int \mathcal{D}\varphi \exp - \left\{ S_0 - \int d^d x h\varphi \right\} \end{aligned}$$

The last expression involves a Gaussian functional integral which can be expressed in terms of the free propagator G_0 written in (3)

(with θ replaced by m_0^2). Indeed we recall that

$$\begin{aligned} \int \mathcal{D}\varphi \exp - \left\{ S_0 - \int d^d x h\varphi \right\} \\ = Z_0 \exp \frac{1}{2} \int d^d x d^d y h(x) G_0(x-y) h(y) \\ Z_0 = \det^{1/2} G_0(x-y) \end{aligned} \quad (26)$$

To define precisely the normalization factor Z_0 , one requires both a short-range regularization and an infrared cutoff (since $\ln Z_0$ is proportional to the volume). Of course, Z_0 is not $Z(h=0)$. Hence we get

$$\begin{aligned} Z(h) &= Z_0 \exp \left(- \int d^d x \mathcal{L}_I \left(\frac{\delta}{\delta h} \right) \right) \\ &\quad \exp \frac{1}{2} \int d^d x d^d y h(x) G_0(x-y) h(y) \end{aligned} \quad (27)$$

The unrenormalized perturbative series is obtained by expanding the first exponential and by performing, term by term, the differentiation operations. These implicitly contain Wick's theorem in the form

$$\begin{aligned} \langle \varphi(x_1) \dots \varphi(x_{2n}) \rangle_0 &= \frac{\delta^{2n}}{\delta h(x_1) \dots \delta h(x_{2n})} \frac{1}{2^n n!} \\ &\quad \left(\int d^d x d^d y h(x) G_0(x-y) h(y) \right)^n \Big|_{h=0} \\ &= \sum_{\text{distinct terms}} G_0(x_{a_1} - x_{a_2}) \dots G_0(x_{a_{2n-1}} - x_{a_{2n}}) \end{aligned} \quad (28)$$

where the last sum contains $(2n)!/2^n n!$ distinct terms (taking into account the invariance of G_0 under parity, $G_0(x) = G_0(-x)$).

It is convenient to express the perturbative rules directly in momentum space for the vertex functions. We recall that these functions are amputated from the external line factors corresponding to the arguments φ of $\Gamma(\varphi)$. To each term of (27) is associated a representative diagram combining vertices (from each monomial of $\mathcal{L}_I$) and lines (from each factor G_0 in (28)). Only the irreducible (and hence connected) diagrams are considered in the calculation of Γ_n .

For the φ^4 interaction, each vertex is associated with a factor $-g_0$ and a momentum conservation δ -function $(2\pi)^d \delta(\Sigma_p)$ for the

sum Σ_p of internal and external momenta entering the vertex. A line joining a vertex to itself (*tadpole*) does not contribute to Σ_p . Each internal line is associated with a d -dimensional momentum p with a propagation factor $G_0(p) = (p^2 + m_0^2)^{-1}$. One has to integrate over internal momenta with the measure $d^d p / (2\pi)^d$.

These Feynman rules are complemented by a combinatorial factor resulting, in some cases, in the incomplete cancellations of the factorials introduced in the definition of the interaction $(1/4!)$ and in the expansion of the exponential. This *symmetry factor* divides the above contribution and is the order of the symmetry group of the diagram. This group corresponds to the permutations of *internal* lines and vertices leaving the diagram invariant (external lines are tied to sources and should not be permuted). Note that tadpoles yield a factor $\frac{1}{2}$ due to these symmetries. If there is some hesitation while applying these rules, one is well advised to return to relation (27). Examples of symmetry factors can be found in the following exercise. From the coefficient of $\ln Z$ in (19), vertex functions have an additional overall factor -1 .

In a zero-dimensional theory, the partition function is the simple integral

$$z(g) = \int \frac{d\varphi}{(2\pi)^{1/2}} \exp - \left\{ \frac{\varphi^2}{2} + g \frac{\varphi^4}{4!} \right\} \quad (29)$$

It allows a verification of symmetry factors for vacuum diagrams, i.e. without external lines. Since all other factors in the Feynman rules are equal to unity, the coefficient of $(-g)^n$ is the sum of symmetry factors over all diagrams of order n

$$z(g) = \sum_{n=0}^{\infty} (-g)^n z_n \quad (30)$$

$$z_n = \frac{1}{4!^n} \int \frac{d\varphi}{(2\pi)^{1/2}} \frac{\varphi^{4n}}{n!} e^{-\frac{1}{2}\varphi^2} = \frac{4^n \Gamma(2n + \frac{1}{2})}{4!^n n! \Gamma(\frac{1}{2})} = \frac{(4n-1)!!}{4!^n n!}$$

This result can be checked using the symmetry factors of the diagrams displayed in figure 1. Indeed,

$$n=1 \quad \frac{1}{3} = \frac{1}{8}$$

$$n=2 \quad \frac{5 \cdot 7}{2^7 3} = \frac{1}{8^2 2!} + \frac{1}{2^4} + \frac{1}{2 \cdot 4!}$$

$$n=3 \quad \frac{5 \cdot 7 \cdot 11}{2^{10} 3} = \frac{1}{8^3 3!} + \frac{1}{2^4 8} + \frac{1}{8 \cdot 2 \cdot 4!} + \frac{1}{2^5} + \frac{1}{2^4 3} + \frac{1}{2^2 3!} + \frac{1}{2^4 3}$$

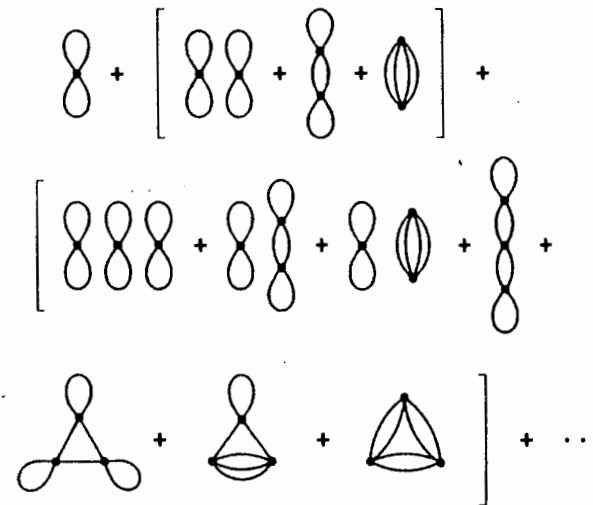


Fig. 1 Vacuum diagrams contributing to $Z(h=0)$ up to and including order g^3 .

and so on. In these expressions, the successive terms correspond respectively to the diagrams of figure 1.

The coefficient z_n increases as n^n , and, as a result, the series (30) is divergent, in fact asymptotic as $g \rightarrow 0$. The point $g=0$ is an essential singularity of $z(g)$ with a cut on the negative real axis, showing an instability for $g < 0$. These properties, which are clear for the simple integral (29), may be generalized to functional integrals. For the same reasons, one expects that the perturbative expansion leads to divergent, asymptotic series. However, signs alternate and there exist resummation methods (such as the one using a Borel transformation).

Study the perturbative expansion of the integral (29) with an external source added.

Let us now summarize the sets of diagrams for the computation of the various Green functions.

- i) **General correlation functions.** The contributions of external lines should be kept. Diagrams may be disconnected, but connected parts without external lines should be omitted (vacuum diagrams are eliminated by division by $Z(h=0)$).

The generating functional is

$$\frac{Z(h)}{Z(0)} = \sum_{n=0}^{\infty} \frac{1}{n!} \int \prod_{k=1}^n d^d x_k \prod_{k=1}^n h(x_k) G_n(x_1, \dots, x_n) \quad (31a)$$

ii) **Connected functions** obey the same rules, but only connected diagrams are to be considered. They are generated by

$$F(h) = \ln \frac{Z(h)}{Z(0)} = \sum_{n=0}^{\infty} \frac{1}{n!} \int \prod_{k=1}^n d^d x_k \prod_{k=1}^n h(x_k) G_n^c(x_1, \dots, x_n) \quad (31b)$$

iii) **Vertex functions.** No factor is included for the external lines of one-particle irreducible diagrams. The first diagrams of the 2- and 4-point vertex functions are displayed in figure 2 and their contribution will be computed in the following sections. The generating functional is given by the Legendre transform of $F(h)$

$$\begin{aligned} \Gamma(\varphi) &= \sup_h \left\{ \int d^d x h \varphi - F(h) \right\} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \prod_{k=1}^n d^d x_k \prod_{k=1}^n \varphi(x_k) \Gamma_n(x_1, \dots, x_n) \end{aligned} \quad (31c)$$

Besides the enumeration of diagrams and the determination of the combinatorial factors (a relatively easy task at low orders), one has to perform integrations. The computation of the corresponding Feynman integrals is in general difficult and involves many technical tricks. In most cases, these integrals cannot be expressed in terms of usual special functions.

5.2.2 Loop expansion

An n -point Green function depends, in its momentum representation, on n d -dimensional vectors, the sum of which vanishes due to translational invariance. If rotational invariance is preserved by the regularization procedure – which we assume to be the case – the function can be expressed in terms of the $n(n-1)/2$ invariants $p_i \cdot p_j$, $i \leq j \leq n-1$. For an integral dimension such that $n-1 \geq d+1$, there are necessarily additional linear relations

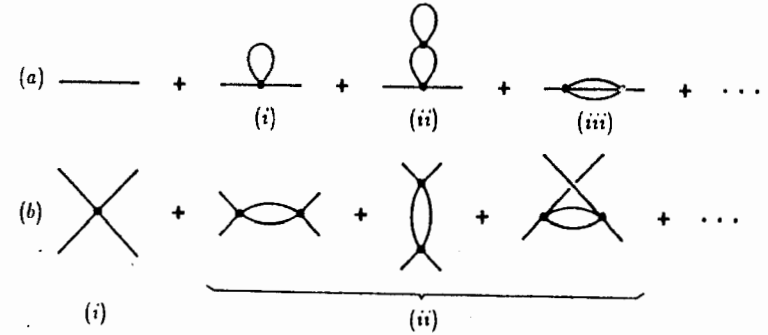


Fig. 2 First diagrams for the vertex functions. (a) 2-point function up to two-loop contributions, (b) 4-point function up to one-loop contributions.

between the vectors p_i which can be obtained by writing that the matrix $\{p_i \cdot p_j\}$, with $1 \leq i, j \leq n-1$, is at most of rank d .

Let us consider a connected n -point diagram of order v (with v vertices). In the perturbative theory around the fundamental state $\langle \varphi \rangle = 0$, the number n of external lines is necessarily even. Let l be the number of internal lines. Note that some of them are contracted on the same vertex (tadpole), with an associated factor

$$\int \frac{d^d p}{(2\pi)^d} \frac{1}{p^2 + m_0^2}$$

which is independent of the remaining part of the diagram and which converges only due to the presence of the cutoff factor. There exists a prescription (*Wick's ordering*) which eliminates these terms, but we will not use it here. With a φ^4 interaction, there are $4v$ lines incident on the vertices. Taking into account the fact that an internal line is incident on two vertices, we get

$$4v = n + 2l \quad (32a)$$

There are l vectorial integration variables, and v conservation rules at the vertices, one of which expresses the total external momenta conservation. Hence, the number B of independent (vectorial) integration variables, called the *number of loops*, is

$$B = l - v + 1 = v - \frac{1}{2}n + 1 \quad (32b)$$

Diagrams without loops are called *trees*. As an example, the lowest order diagrams for the vertex functions have been classified according to their number of loops in figure 2, and their contributions are

$$\begin{aligned}\Gamma_2(\mathbf{p}) &= \mathbf{p}^2 + m_0^2 + \frac{1}{2}g_0 \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\mathbf{q}^2 + m_0^2} + \dots \\ \Gamma_4(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4) &= g_0 \\ &- \frac{1}{2}g_0^2 \sum_{i=2}^4 \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{[(\mathbf{p}_1 + \mathbf{p}_i - \mathbf{q})^2 + m_0^2][\mathbf{q}^2 + m_0^2]} + \dots\end{aligned}\quad (33)$$

The number of loops B may be interpreted as the power of $\hbar$ in a semiclassical expansion. We have indeed $(\hbar)^{B-1} = \hbar^l \hbar^{-v}$, hence each vertex is assigned a factor $1/\hbar$ and each propagator a factor $\hbar$. This is equivalent to dividing the action by $\hbar$. The perturbative expansion of $Z(\hbar)$ appears as the semiclassical expansion of the functional integral around the saddle point $\varphi = 0$.

Let us compute the generating functional $\Gamma(\varphi)$ up to and including one-loop corrections. The partition function $Z(\hbar)$ is obtained by determining the saddle point of $\exp\{-S + \int d^d \mathbf{x} \hbar \varphi\}$, i.e.

$$(-\Delta + m_0^2)\varphi_c + \frac{1}{3!}g_0\varphi_c^3 = \hbar \quad (34)$$

One considers the solution which vanishes with $\hbar$

$$\varphi_c[\hbar] = \frac{1}{-\Delta + m_0^2}\hbar - \frac{1}{3!}g_0 \left(\frac{1}{-\Delta + m_0^2}\hbar \right)^3 + \dots \quad (35)$$

The next order is obtained by expanding the action around φ_c , and performing the corresponding Gaussian integral

$$\begin{aligned}\frac{Z(\hbar)}{Z(0)} &= \exp\left\{-S(\varphi_c[\hbar]) + \int d^d \mathbf{x} \hbar \varphi_c[\hbar]\right\} \\ &\times \det^{-1/2} \left[\frac{\delta^2 S(\varphi_c[\hbar])}{\delta \varphi(\mathbf{x}) \delta \varphi(\mathbf{y})} \right] \det^{1/2} \left[\frac{\delta^2 S_0}{\delta \varphi(\mathbf{x}) \delta \varphi(\mathbf{y})} \right]\end{aligned}\quad (36)$$

The Legendre transformation gives

$$\Gamma(\varphi) = S(\varphi) - \frac{1}{2} \text{Tr} \ln(G_\varphi G_0^{-1}) + \dots \quad (37)$$

where $G_\varphi(\mathbf{x}, \mathbf{y})$ is the inverse kernel of $\delta^2 S / \delta \varphi(\mathbf{x}) \delta \varphi(\mathbf{y})$. This reads

$$\begin{aligned}G_\varphi &= [-\Delta + m_0^2 + \frac{1}{2}g_0\varphi^2]^{-1} \\ G_\varphi G_0^{-1} &= 1 - G_0 \frac{1}{2}g_0\varphi^2 + G_0 \frac{1}{2}g_0\varphi^2 G_0 \frac{1}{2}g_0\varphi^2 + \dots \\ &= [1 + G_0 \frac{1}{2}g_0\varphi^2]^{-1}\end{aligned}\quad (38)$$

Hence formula (37) is rewritten as

$$\begin{aligned}\Gamma(\varphi) &= S(\varphi) + \frac{1}{2} \text{Tr} \ln(1 + \frac{1}{2}G_0 g_0 \varphi^2) \\ &= S(\varphi) + \sum_{p=1}^{\infty} \frac{(-1)^{p+1} g_0^p}{2p} \text{Tr} (\frac{1}{2}G_0 \varphi^2)^p \\ &= S(\varphi) + \frac{1}{2}g_0 \frac{1}{2!} \int \frac{d^d \mathbf{p}}{(2\pi)^d} \varphi(\mathbf{p}) \varphi(-\mathbf{p}) \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\mathbf{q}^2 + m_0^2} \\ &- \frac{1}{2}g_0^2 \frac{1}{4!} \int \prod_{k=1}^3 \frac{d^d \mathbf{p}_k}{(2\pi)^d} \varphi(\mathbf{p}_1) \varphi(\mathbf{p}_2) \varphi(\mathbf{p}_3) \varphi(\mathbf{p}_4 = -\mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3) \\ &\times \sum_{i=2}^4 \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{[(\mathbf{p}_1 + \mathbf{p}_i - \mathbf{q})^2 + m_0^2][\mathbf{q}^2 + m_0^2]} + \dots\end{aligned}\quad (39)$$

As $\varphi(\mathbf{x})$ tends towards a constant φ , $\Gamma(\varphi)$ becomes the integral of the effective potential $V(\varphi)$

$$\Gamma(\varphi) = \int d^d \mathbf{x} V(\varphi) \quad (40)$$

With the one-loop correction, $V(\varphi)$ is given by

$$V(\varphi) = \frac{1}{2}m_0^2\varphi^2 + \frac{1}{4!}g_0\varphi^4 + \frac{1}{2} \int \frac{d^d \mathbf{p}}{(2\pi)^d} \ln \left[1 + \frac{1}{\mathbf{p}^2 + m_0^2} \frac{1}{2}g_0\varphi^2 \right] + \dots \quad (41)$$

The steepest descent method, including the corrections due to the fluctuations around the saddle point $\varphi_c[\hbar]$, is thus identical to the perturbative theory. Terms with more loops are obtained from irreducible diagrams using the propagator G_φ given in (38). Apart from technical complications, this method can be generalized in various ways, e.g. to degenerate saddle points or to fields with internal degrees of freedom.

Besides the Green functions considered so far, one is also interested in mixed correlations including insertions of local operators, as already mentioned above. For instance, the insertion

of $\frac{1}{2}\varphi^2$,

$$\left\langle \varphi(x_1) \dots \varphi(x_n) \frac{1}{2} \varphi^2(y) \right\rangle = \int \prod_{k=1}^n \frac{d^d p_k}{(2\pi)^d} \frac{d^d q}{(2\pi)^d} (2\pi)^d \delta(\sum p_k + q) e^{-i(\sum p_k \cdot x_k + q \cdot y)} G_{n,1}(p_1, \dots, p_n; q) \quad (42)$$

arises formally from a limiting procedure

$$\left\langle \varphi(x_1) \dots \varphi(x_n) \frac{1}{2} \varphi^2(y) \right\rangle = \lim_{z \rightarrow 0} \frac{1}{2} \langle \varphi(x_1) \dots \varphi(x_n) \varphi(y+z) \varphi(y-z) \rangle \quad (43)$$

that is, in momentum representation,

$$G_{n,1}(p_1, \dots, p_n; q) = \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} G_{n+2}(p_1, \dots, p_n, \frac{1}{2}q + k, \frac{1}{2}q - k) \quad (44)$$

The insertion of $\frac{1}{2}\varphi^2$ at zero momentum is equivalent to a derivative with respect to m_0^2 . The insertion of more terms $\frac{1}{2}\varphi^2$ or other combinations of local operators leads to analogous expressions. The corresponding perturbative expansions are easily derived.

5.2.3 Evaluation of integrals and dimensional continuation

Several techniques have been developed for the evaluation of Feynman integrals and for the study of their analytical properties. We give here only some illustrative examples and focus on the dimensional continuation.

The introduction of vertex functions is a first step in reducing the complexity, since other Green functions can be deduced by algebraic operations. Note that there still remain articulated diagrams – i.e. which can be disconnected by omission of one vertex, see chapter 7 (in volume 2) – which lead to factorizable contributions.

As we are interested by dimensional continuation which, of course, makes sense only for scalars, it is important to use only scalar invariant combinations in the integrands. The identity

$$\frac{1}{(q^2 + m_0^2)^n} = \frac{1}{\Gamma(n)} \int_0^\infty d\alpha \alpha^{n-1} e^{-\alpha(q^2 + m_0^2)} \quad (45)$$

allows one to substitute exponentials for polynomials in the denominators. These exponentials are Gaussian with respect to

the momenta and the corresponding integrations can be carried out in any dimension. This method introduces conjugate variables α with inverse-square mass dimension. In a relativistic context, they are interpreted as proper times.

Let us take as an example the first (constant) correction to the 2-point function, which yields a shift of the critical temperature due to fluctuations. From equation (33), (diagram i of figure 2a)

$$\begin{aligned} l_2^n &= \frac{1}{2} g_0 \int \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m_0^2} = \frac{1}{2} g_0 \int \frac{d^d q}{(2\pi)^d} \int_0^\infty d\alpha e^{-\alpha(q^2 + m_0^2)} \\ &= \frac{g_0}{2(4\pi)^{d/2}} \int_0^\infty d\alpha \alpha^{-d/2} e^{-\alpha m_0^2} \end{aligned} \quad (46)$$

The Gaussian representation allows one to integrate over the internal momentum for an arbitrary dimension d . Rotational invariance has provided the area of the unit sphere

$$S_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \quad (47)$$

and the remaining radial integral yields in particular the power $\alpha^{-d/2}$. The dimension of $d\alpha \alpha^{-d/2}$ is the same as the original one of $d^d q (q^2 + m_0^2)^{-2}$. In the integral over the parameter α , two regions may generate divergences. The infrared region $\alpha \rightarrow \infty$ will give divergences if the masses vanish for $d \leq 2$. The ultraviolet region $\alpha \rightarrow 0$ is not sensitive to masses and gives singularities when $d \geq 2$. The integral is never convergent if initially $m_0^2 = 0$. However, with $m_0^2 > 0$, the integral is well-defined for $d < 2$ and can be analytically continued as a meromorphic function of d with simple poles for $d = 2, 4, \dots$

$$\Gamma_2^\alpha = \frac{g_0 m_0^{d-2}}{2(4\pi)^{d/2}} \Gamma(1 - d/2) \quad (48)$$

As the dimension of g_0 is $4 - d$, $g_0 m_0^{d-2} = (g_0 m_0^{d-4}) m_0^2$ has the correct dimension 2 of the irreducible 2-point function. For an even integral dimension, the expression (48) is meaningless due to ultraviolet singularities. Therefore, an ultraviolet regularization – a last remnant from the lattice – should be used. For this purpose, it is sufficient to remove from the integral over α a domain $(0, \Lambda^{-2})$. More generally, one introduces in the integral a function $\chi_\Lambda(\alpha)$ equal to unity as $\alpha \geq \Lambda^{-2}$ and vanishing with a sufficient number

of its derivatives at $\alpha = 0$. The arbitrariness on this function reflects the arbitrariness on the regularization procedure which should not modify the universal critical properties at distances large with respect to $a \sim \Lambda^{-1}$. With the above choice for the function $\chi_\Lambda(\alpha)$, i.e. Heaviside's step function $\theta(\alpha - \Lambda^{-2})$, we find

$$\Gamma_2^a \rightarrow \frac{g_0}{2(4\pi)^{d/2}} \int_{\Lambda^{-2}}^{\infty} d\alpha \alpha^{-d/2} e^{-m_0^2 \alpha} = \frac{g_0 m_0^{d-2}}{2(4\pi)^{d/2}} \times \left\{ \frac{2}{d-2} \left(\frac{\Lambda}{m_0} \right)^{d-2} e^{-m_0^2/\Lambda^2} - \frac{4}{(d-2)(d-4)} \left(\frac{\Lambda}{m_0} \right)^{d-4} e^{-m_0^2/\Lambda^2} + \frac{4}{(d-2)(d-4)} \int_{m_0^2/\Lambda^2}^{\infty} dx x^{2-d/2} e^{-x} \right\} \quad (49)$$

This expression can now be evaluated near $d = 4$ in the limit $\Lambda/m_0 \gg 1$. After expanding $\exp(-m_0^2/\Lambda^2)$, we get, with γ standing for Euler's constant (see below equation (59)),

$$\Gamma_2^c \underset{d \rightarrow 4}{\sim} \frac{g_0 m_0^{d-2}}{(4\pi)^{d/2}} \frac{1}{(d-2)} \times \left\{ (\Lambda/m_0)^{d-2} - (\Lambda/m_0)^{d-4} + 2 \frac{1 - (\Lambda/m_0)^{d-4}}{d-4} + \gamma + \dots \right\} \quad (50)$$

For $d = 4$, this is

$$\Gamma_2^a = \frac{g_0 m_0^2}{2(4\pi)^2} \left\{ \frac{\Lambda^2}{m_0^2} + \gamma - 1 - \ln \frac{\Lambda^2}{m_0^2} \right\} \quad (d = 4) \quad (51)$$

Thanks to the cutoff factor Λ , the pole at $d = 2$ translates into a quadratically ultraviolet divergent term, and the one at $d = 4$, into a logarithmically divergent term.

Of course, the above integral is particularly simple. However, the various phenomena just encountered generalize to more complex cases. Let us look for instance at the two-loop diagrams for the 2-point function. The two irreducible diagrams *ii* and *iii* of figure 2a yield the contributions

$$\Gamma_2^b = -\frac{1}{4} g_0^2 \int \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m_0^2} \int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 + m_0^2)^2} \quad (52)$$

$$= -\frac{1}{4} g_0^2 \frac{m_0^{2d-6}}{(4\pi)^d} \Gamma(1-d/2) \Gamma(2-d/2)$$

$$\begin{aligned} \Gamma_2^c(p) &= -\frac{1}{6} g_0^2 \int \frac{d^d q d^d k}{(2\pi)^{2d}} \frac{1}{[q^2 + m_0^2][k^2 + m_0^2][(q-p+k)^2 + m_0^2]} \\ &= -\frac{1}{6} g_0^2 \int_0^\infty d\alpha_1 d\alpha_2 d\alpha_3 e^{-m_0^2(\alpha_1 + \alpha_2 + \alpha_3)} \\ &\quad \int \frac{d^d q d^d k}{(2\pi)^{2d}} e^{-(\alpha_1 q^2 + \alpha_2 k^2 + \alpha_3 (q+k-p)^2)} \\ &= -\frac{g_0^2}{6(4\pi)^d} \int_0^\infty \frac{d\alpha_1 d\alpha_2 d\alpha_3}{(\alpha_2 \alpha_3 + \alpha_3 \alpha_1 + \alpha_1 \alpha_2)^{d/2}} \\ &\quad \exp \left\{ m_0^2(\alpha_1 + \alpha_2 + \alpha_3) + \frac{\alpha_1 \alpha_2 \alpha_3 p^2}{\alpha_1 \alpha_2 + \alpha_2 \alpha_3 + \alpha_3 \alpha_1} \right\} \end{aligned} \quad (53)$$

The homogeneity properties with respect to the variable α_i can be exploited by introducing a factor

$$1 = \int_0^\infty d\lambda \delta(\alpha_1 + \alpha_2 + \alpha_3 - \lambda)$$

Then α_i is changed to $\lambda \alpha_i$, and the integral over λ , which is convergent for $d < 3$, is performed. Hence

$$\begin{aligned} \Gamma_2^c(p) &= -\frac{g_0^2}{6(4\pi)^d} \int_0^1 \frac{d\alpha_1 d\alpha_2 d\alpha_3 \delta(\alpha_1 + \alpha_2 + \alpha_3 - 1)}{(\alpha_2 \alpha_3 + \alpha_3 \alpha_1 + \alpha_1 \alpha_2)^{d/2}} \\ &\quad \times \int_0^\infty d\lambda \lambda^{2-d} \exp \left\{ m_0^2 + \frac{\alpha_1 \alpha_2 \alpha_3 p^2}{\alpha_1 \alpha_2 + \alpha_2 \alpha_3 + \alpha_3 \alpha_1} \right\} \\ &= -\frac{g_0^2}{6(4\pi)^d} \Gamma(3-d) \int_0^1 \frac{d\alpha_1 d\alpha_2 d\alpha_3 \delta(\alpha_1 + \alpha_2 + \alpha_3 - 1)}{(\alpha_2 \alpha_3 + \alpha_3 \alpha_1 + \alpha_1 \alpha_2)^{d/2}} \\ &\quad \times \left\{ m_0^2 + \frac{\alpha_1 \alpha_2 \alpha_3 p^2}{\alpha_1 \alpha_2 + \alpha_2 \alpha_3 + \alpha_3 \alpha_1} \right\}^{d-3} \end{aligned} \quad (54)$$

Analyse the ultraviolet convergence of the diagram using dimensional arguments.

For $d \geq 3$, keeping $m_0^2 > 0$ prevents infrared divergences, but ultraviolet ones appear. At $d = 3$, the divergence is logarithmic and $\Gamma_2^c(p) - \Gamma_2^c(0)$ remains finite. In dimension four, the value at zero momentum is quadratically divergent (as is Γ_2^a) and the first derivative $d\Gamma_2^c(0)/dp^2$ is still logarithmically divergent. This leads to the so-called wavefunction renormalization which appears in φ^4 theory at the two-loop order.

Introducing the cutoff factor Λ in the α integrals, let us compute these divergent terms at $d = 4$

$$\Gamma_2^c(\mathbf{p}) = -\frac{1}{6(4\pi)^4} g_0^2 \{A + B\mathbf{p}^2 + C(\mathbf{p}^2)\} \quad (55)$$

where the finite function $C(\mathbf{p}^2)$ vanishes with its first derivative at the origin. Using the notation

$$\eta = \frac{m_0^2}{\Lambda^2} \quad (56)$$

we get from (53)

$$\frac{A}{m_0^2} = \eta^{-1} \int_1^\infty d\alpha_1 \int_1^\infty d\alpha_2 \int_1^\infty d\alpha_3 \frac{e^{-\eta(\alpha_1 + \alpha_2 + \alpha_3)}}{(\alpha_1 \alpha_2 + \alpha_2 \alpha_3 + \alpha_3 \alpha_1)^2} \quad (57)$$

The symmetric integral is computed in the sector $\alpha_3 > \alpha_2 > \alpha_1 > 1$. Introducing the variables $\alpha_3/\alpha_2 = z$, $\alpha_2/\alpha_1 = y$, $\alpha_1 = x$,

$$\frac{A}{m_0^2} = 6\eta^{-1} \int_1^\infty \frac{dx}{x^2} \int_1^\infty \frac{dy}{y} \int_1^\infty dz \frac{e^{-\eta x(1+y+yz)}}{(1+z+yz)^2} = 6 \int_\eta^\infty \frac{du}{u^2} e^{-u} \psi(u)$$

where $u = \eta x$ and

$$\psi(u) = \int_1^\infty \frac{dy}{y} \int_1^\infty dz \frac{e^{-uy(1+z)}}{(1+z+yz)^2}$$

The singular behaviour of A as $\eta \rightarrow 0$ arises from the small u integration domain where

$$\psi(u) = \psi_0 + \psi_1 u \ln u + \psi_2 u + \mathcal{O}(u^2 \ln^2 u)$$

A useful table of integrals is

$$\begin{aligned} \int_\eta^\infty \frac{du}{u^2} e^{-u} &= \frac{1}{\eta} + \ln \eta + \text{finite part} \\ \int_\eta^\infty \frac{du}{u} \ln u e^{-u} &= -\frac{1}{2} \ln^2 \eta + \text{finite part} \\ \int_\eta^\infty \frac{du}{u} e^{-u} &= -\ln \eta + \text{finite part} \end{aligned} \quad (58)$$

The final result is

$$\frac{A}{m_0^2} = 6\eta^{-1} \psi_0 - 3\psi_1 \ln^2 \eta + 6(\psi_0 - \psi_2) \ln \eta + \text{finite part}$$

with

$$\psi_0 = \int_1^\infty \frac{dy}{y} \int_1^\infty dz \frac{1}{(1+z+yz)^2} = \int_1^\infty dy \frac{1}{y(1+y)(2+y)} = \frac{1}{2} \ln \frac{4}{3}$$

To evaluate ψ_1 and ψ_2 , let us study

$$\begin{aligned} \frac{d\psi(u)}{du} &= \psi_1 \ln u + (\psi_1 + \psi_2) + \mathcal{O}(u \ln^2 u) \\ &= - \int_1^\infty dy \int_1^\infty dz \frac{1+z}{[1+z(1+y)]^2} e^{-uy(1+z)} \end{aligned}$$

The last integral is split into two terms φ_1 and φ_2 , corresponding to the decomposition

$$\begin{aligned} \frac{1+z}{[1+z+yz]^2} &= \frac{y}{(1+y)[1+z+yz]^2} + \frac{1}{(1+y)[1+z+yz]} \\ \varphi_1(u) &= - \int_1^\infty dy \frac{y}{1+y} \int_1^\infty dz \frac{e^{-uy(1+z)}}{(1+z+yz)^2} \\ &\xrightarrow{u \rightarrow 0} - \int_1^\infty dy \frac{y}{1+y} \int_1^\infty \frac{dz}{(1+z+yz)^2} = 2 \ln \frac{2}{3} + \frac{1}{2} \\ \varphi_2(u) &= - \int_1^\infty \frac{dy}{1+y} \int_1^\infty \frac{dz}{1+z+yz} e^{-uy(1+z)} \\ &= - \int_1^\infty \frac{dy}{(1+y)^2} e^{-uy + \frac{uy}{1+y}} \int_{uy(2+y)/(1+y)}^\infty \frac{e^{-v} dv}{v} \end{aligned}$$

Using Euler's constant

$$\gamma = -\Gamma'(1) = - \int_0^\infty dt \ln t e^{-t} \quad (59)$$

one finds

$$\begin{aligned} \varphi_2(u) &= \int_1^\infty \frac{dy}{(1+y)^2} \exp\left(-uy + \frac{uy}{1+y}\right) \\ &\quad \left\{ \exp\left(-\frac{uy(2+y)}{1+y}\right) \ln \frac{uy(2+y)}{1+y} + \gamma \right\} + \mathcal{O}(u \ln^2 u) \\ &= (\ln u + \gamma) \int_1^\infty \frac{dy}{(1+y)^2} + \int_1^\infty \frac{dy}{(1+y)^2} \ln \frac{y(2+y)}{1+y} \\ &\quad + \mathcal{O}(u \ln^2 u) \end{aligned}$$

Therefore one gets

$$\varphi_2(u) = \frac{1}{2} (\ln u + \gamma - 1 + 3 \ln 3 - \ln 2) + \mathcal{O}(u \ln^2 u)$$

Gathering all these results,

$$\begin{aligned} \frac{d\psi}{du} &= \psi_1 \ln u + \psi_1 + \psi_2 + \mathcal{O}(u \ln^2 u) \\ &= \varphi_1(u) + \varphi_2(u) = \frac{1}{2} \ln u + \frac{1}{2} (\gamma - \ln 3 + 3 \ln 2) + \mathcal{O}(u \ln^2 u) \end{aligned}$$

Hence

$$\psi_1 = \frac{1}{2} \quad \psi_2 = \frac{1}{2}(\gamma - 1 - \ln 3 + 3 \ln 2)$$

The first divergent coefficient finally reads

$$\frac{A}{m_0^2} = (3 \ln \frac{1}{3}) \frac{\Lambda^2}{m_0^2} - \frac{3}{2} \ln^2 \frac{\Lambda^2}{m_0^2} + 3[\ln 2 - 1 + \gamma] \ln \frac{\Lambda^2}{m_0^2} + \text{finite part} \quad (60)$$

Now we turn to the coefficient B in equation (55). One has

$$\begin{aligned} B &= - \int_{\Lambda^{-2}}^{\infty} \frac{d\alpha_1 d\alpha_2 d\alpha_3 \alpha_1 \alpha_2 \alpha_3}{(\alpha_1 \alpha_2 + \alpha_2 \alpha_3 + \alpha_3 \alpha_1)^3} e^{-m_0^2(\alpha_1 + \alpha_2 + \alpha_3)} \\ &= -6 \int_{\eta}^{\infty} \frac{dx}{x} e^{-x} \int_1^{\infty} dy e^{-xy} \int_1^{\infty} dz \frac{ze^{-xyz}}{(1+z+yz)^3} \quad (61) \\ &= -6 \int_{\eta}^{\infty} \frac{dx}{x} e^{-x} \bar{\psi}(x) = 6\bar{\psi}(0) \ln \eta + \text{finite part} \end{aligned}$$

where $\bar{\psi}(0)$ is finite

$$\begin{aligned} \bar{\psi}(0) &= \int_1^{\infty} dy \int_1^{\infty} dz \frac{z}{(1+z+yz)^3} = \int_1^{\infty} dz \int_1^{\infty} \frac{z dy}{(1+z+zy)^3} \\ &= \frac{1}{2} \int_1^{\infty} \frac{dz}{(1+2z)^2} = \frac{1}{12} \end{aligned}$$

Consequently,

$$B = -\frac{1}{2} \ln \frac{\Lambda^2}{m_0^2} + \text{finite part} \quad (62)$$

This completes the computation of singular terms of $\Gamma_2^c(p^2)$ in four dimensions.

(1) Show that, for $d = 3$, one can write

$$\Gamma_2^c(p^2) = -\frac{g_0^2}{6(4\pi)^3} \{ \bar{A} + \bar{B}p^2 + \bar{C}(p^2) \} \quad (63)$$

$$\bar{A} = 2\pi \ln \frac{\Lambda^2}{m_0^2} + \text{finite part} \quad \bar{B} = -\frac{2\pi}{27} m_0^{-2} \quad \bar{C} = \mathcal{O}(p^4)$$

In this case, only $\bar{A}$ is logarithmically divergent, and the terms within the braces are dimensionless since g_0 has the dimension of a mass.

(2) If $m_0^2 = 0$, the integral defining $\Gamma_2^c(p)$ converges for $2 < d < 3$. Show that its analytic continuation reads

$$\Gamma_2^c(p^2, m_0^2 = 0) = -\frac{1}{6(4\pi)^d} g_0^2 (p^2)^{d-3} \frac{\Gamma(3-d)\Gamma^3(d/2-1)}{\Gamma(3(d/2-1))} \quad (64)$$

which behaves as

$$\Gamma_2^c(p^2, m_0^2 = 0) \underset{d \rightarrow 4}{\sim} -\frac{1}{6(4\pi)^d} g_0^2 p^2 \frac{(p^2)^{d-4}}{2(d-4)} \quad (65)$$

Let us call $G(p^2)$ the coefficient of $-g_0^2 p^2 / 6(4\pi)^d$. The quantity $G(p^2) - G(P^2)$ has a finite limit as $d \rightarrow 4$

$$\lim_{d \rightarrow 4} G(p^2) - G(P^2) = \frac{1}{2} \ln \frac{p^2}{P^2}$$

This result may be compared to the expression (62) for the coefficient B . In the four-dimensional regularized theory, we have indeed

$$\Gamma_2^c(p^2, m_0^2, \Lambda^2) = -\frac{1}{6(4\pi)^4} \{ A(0, m_0^2, \Lambda^2) + p^2 G(p^2, m_0^2, \Lambda^2) \}$$

with

$$G(p^2, m_0^2, \Lambda^2) = \frac{1}{2} \ln \frac{m_0^2}{\Lambda^2} + g\left(\frac{p^2}{m_0^2}\right)$$

for large Λ^2 , with g finite and dimensionless just as G . In the region $m_0^2 \ll p^2, P^2 \ll \Lambda^2$, we introduce $G(p^2, m_0^2, \Lambda^2) - G(P^2, m_0^2, \Lambda^2) = g(p^2/m_0^2) - g(P^2/m_0^2)$. The preceding result shows that the limit $m_0^2 \rightarrow 0$ can be taken, and this implies that, for $p^2 \gg m_0^2$

$$g(p^2/m_0^2) \rightarrow \frac{1}{2} \ln(p^2/m_0^2)$$

In this domain $m_0^2 \ll p^2 \ll \Lambda^2$,

$$G(p^2, m_0^2, \Lambda^2) \sim \frac{1}{2} \ln \frac{m_0^2}{\Lambda^2} + \frac{1}{2} \ln \frac{p^2}{m_0^2} = \frac{1}{2} \ln \frac{p^2}{\Lambda^2}$$

Let us also compute the 4-point function at the one-loop order. The contributions of the corresponding diagrams are successively (figure 2b)

$$\Gamma_4^a = g_0 \quad (66)$$

$$\Gamma_4^b(p_1, \dots, p_4) = -\frac{1}{2} g_0^2 \sum_{i=2}^4 \int \frac{d^d q}{(2\pi)^d} \frac{1}{[q^2 + m_0^2][(p_1 + p_i - q)^2 + m_0^2]} \quad (67)$$

Let us denote the integral $F(\mathbf{p}^2 = (\mathbf{p}_1 + \mathbf{p}_i)^2)$. It reads

$$F(\mathbf{p}^2) = \frac{1}{(4\pi)^{d/2}} \int_0^\infty \frac{d\alpha_1 d\alpha_2}{(\alpha_1 + \alpha_2)^{d/2}} \exp - \left\{ m_0^2(\alpha_1 + \alpha_2) + \frac{\alpha_1 \alpha_2 \mathbf{p}^2}{\alpha_1 + \alpha_2} \right\} \\ = \frac{\Gamma(2 - d/2)}{(4\pi)^{d/2}} \int_0^1 \frac{d\alpha}{[m_0^2 + \mathbf{p}^2 \alpha(1 - \alpha)]^{2-d/2}} \quad (68)$$

In the case $m_0^2 = 0$, one gets

$$F(\mathbf{p}^2, m_0^2 = 0) = (\mathbf{p}^2)^{d/2-2} \frac{\Gamma(2 - d/2) \Gamma(d/2 - 1)}{(4\pi)^{d/2} \Gamma(d - 2)} \quad (69)$$

and one notices an infrared divergence at $d = 2$. Near $d = 4$, the quantity $F(\mathbf{p}^2, m_0^2 = 0) - F(\mu^2, m_0^2 = 0)$ has a limit

$$\lim_{d \rightarrow 4} F(\mathbf{p}^2, m_0^2 = 0) - F(\mu^2, m_0^2 = 0) = -\frac{1}{(4\pi)^2} \ln \frac{\mathbf{p}^2}{\mu^2} \quad (70)$$

The same logarithmic divergence can be computed for $d = 4$ in the regularized theory where

$$F(\mathbf{p}^2, m_0^2, \Lambda^2) = \frac{1}{(4\pi)^2} \int_{\Lambda^{-2}}^\infty \frac{d\alpha_1 d\alpha_2}{(\alpha_1 + \alpha_2)^2} \exp - \left\{ m_0^2(\alpha_1 + \alpha_2) + \frac{\alpha_1 \alpha_2 \mathbf{p}^2}{\alpha_1 + \alpha_2} \right\} \quad (71)$$

As $\Lambda^2 \rightarrow \infty$,

$$F(\mathbf{p}^2 = 0, m_0^2, \Lambda^2) = \frac{1}{(4\pi)^2} \left\{ \ln \frac{\Lambda^2}{m_0^2} - \ln 2 - 1 - \gamma + \mathcal{O}\left(\frac{m_0^2}{\Lambda^2} \ln \frac{\Lambda^2}{m_0^2}\right) \right\} \quad (72)$$

The dimensionless function F can be cast into the form

$$F(\mathbf{p}^2, m_0^2, \Lambda^2) = \frac{1}{(4\pi)^2} \ln \frac{\Lambda^2}{\mathbf{p}^2} + f\left(\frac{m_0^2}{\mathbf{p}^2}\right)$$

The combination $F(\mathbf{p}^2, m_0^2, \Lambda^2) - F(\mu^2, m_0^2, \Lambda^2)$ has a finite limit for $m_0^2 \rightarrow 0$

$$\lim_{m_0^2 \rightarrow 0} F(\mathbf{p}^2, m_0^2, \Lambda^2) - F(\mu^2, m_0^2, \Lambda^2) = -\frac{1}{(4\pi)^2} \ln \frac{\mathbf{p}^2}{\mu^2}$$

and this expression agrees with (70). This example shows explicitly the relations between ultraviolet behaviour and massless theory.

Rather than giving more examples, let us summarize the method. One first uses a parametric representation of the propagators. Then the Gaussian integration over the momenta is performed. Finally, the ultraviolet behaviour as $\alpha \rightarrow 0$ is analyzed in each sector. The large α behaviour is controlled by internal masses such as m_0^2 .

5.2.4 Group theoretical factors.

Let us look at the n -vector model with an $O(n)$ internal symmetry, described by dynamical variables $\Phi = \{\Phi_1, \dots, \Phi_n\}$ with an interaction $(\Phi^2)^2$. The correlation functions now depend not only on external momenta, but also on indices $i_1, \dots, i_n$ attached to the external fields. They are symmetric under simultaneous permutations of these indices and of the momenta. The contributions of diagrams contain extra polynomials in n of degree at most equal to the number of loops. They arise from contractions and summations on the internal indices. For instance, in the symmetric phase, the two-point function is diagonal in this internal space (see diagrams of figure 2a)

$$\Gamma_{2,ij}(\mathbf{p}^2) = \delta_{ij} \left\{ \mathbf{p}^2 + m_0^2 + \right. \\ + g_0 \frac{n+2}{6} \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\mathbf{q}^2 + m_0^2} \\ - g_0^2 \frac{(n+2)^2}{36} \int \frac{d^d \mathbf{q}}{(2\pi)^d} \frac{1}{\mathbf{q}^2 + m_0^2} \int \frac{d^d \mathbf{k}}{(2\pi)^d} \frac{1}{(\mathbf{k}^2 + m_0^2)^2} \\ - g_0^2 \frac{n+2}{18} \int \frac{d^d \mathbf{q} d^d \mathbf{k}}{(2\pi)^{2d}} \frac{1}{[\mathbf{q}^2 + m_0^2][\mathbf{k}^2 + m_0^2][(\mathbf{q} + \mathbf{k} - \mathbf{p})^2 + m_0^2]} \\ \left. + \dots \right\} \quad (73)$$

The coefficients have been obtained by contracting indices of internal and external lines, following Feynman's rules. The zero-dimensional model yields sum rules on these coefficients. For instance, for the 2-point function,

$$\langle \Phi_i \Phi_j \rangle = \frac{\delta_{ij}}{\Gamma} = \frac{\int d^n \Phi \Phi_i \Phi_j \exp - \left\{ \frac{1}{2} \Phi^2 + \frac{1}{4!} g_0 (\Phi^2)^2 \right\}}{\int d^n \Phi \exp - \left\{ \frac{1}{2} \Phi^2 + \frac{1}{4!} g_0 (\Phi^2)^2 \right\}} \quad (74)$$

hence

$$\begin{aligned}\Gamma &= \frac{1}{2}n \frac{\sum_{p=0}^{\infty} \frac{1}{p!} \left(-\frac{g_0}{6}\right)^p \Gamma(n/2 + 2p)}{\sum_{p=0}^{\infty} \frac{1}{p!} \left(-\frac{g_0}{6}\right)^p \Gamma(n/2 + 2p + 1)} \\ &= 1 + g_0 \frac{n+2}{6} - g_0^2 \frac{(n+2)(n+4)}{36} + \dots\end{aligned}\quad (75)$$

which coincides with (73) when integrals are replaced by unity.

The representation

$$\Gamma = 1 + \frac{g_0}{3} \frac{\int_0^{\infty} dx x^{n/2+1} \exp\left(-\left(x + \frac{g_0}{6}x^2\right)\right)}{\int_0^{\infty} dx x^{n/2} \exp\left(-\left(x + \frac{g_0}{6}x^2\right)\right)}$$

shows that Γ can be expressed as a continuous fraction in the variable $u = g_0/6$

$$\Gamma = 1 + \frac{(n+2)u}{1 + \frac{(n+4)u}{1 + \frac{(n+6)u}{1 + \dots}}}\quad (76)$$

This gives

$$\begin{aligned}\Gamma &= 1 + (n+2)u - (n+2)(n+4)u^2 + (n+2)(n+4)(2n+10)u^3 \\ &\quad - (n+2)(n+4)[(2n+10)^2 + (n+6)(n+8)]u^4 + \dots\end{aligned}\quad (77)$$

In particular $\Gamma = 1$ for $n = -2$ and it is a rational fraction in g_0 for any $n = -2p$.

For the 4-point function, (figure 2b) we have

$$\begin{aligned}\Gamma_{4,i_1 i_2 i_3 i_4}(p_1, p_2, p_3, p_4) &= \frac{1}{3}g_0(\delta_{i_1 i_2} \delta_{i_3 i_4} + \delta_{i_1 i_3} \delta_{i_2 i_4} + \delta_{i_1 i_4} \delta_{i_2 i_3}) \\ &\quad - \frac{1}{9}g_0^2 \sum_{3 \text{ terms}} \left((2 + \frac{1}{2}n)\delta_{i_1 i_2} \delta_{i_3 i_4} + \delta_{i_1 i_3} \delta_{i_2 i_4} + \delta_{i_1 i_4} \delta_{i_2 i_3}\right) \\ &\quad \times \int \frac{d^d q}{(2\pi)^d} \frac{1}{[q^2 + m_0^2][(p_1 + p_2 - q)^2 + m_0^2]} + \dots\end{aligned}\quad (78)$$

In particular, the four-point function is proportional to the combination $\sum \delta_{i_1 i_2} \delta_{i_3 i_4}$, for zero external momenta

$$\begin{aligned}\Gamma_{4,i_1 i_2 i_3 i_4}(0, 0, 0, 0) &= (\delta_{i_1 i_2} \delta_{i_3 i_4} + \delta_{i_1 i_3} \delta_{i_2 i_4} + \delta_{i_1 i_4} \delta_{i_2 i_3}) \times \\ &\quad \left\{ \frac{g_0}{3} - g_0^2 \frac{n+8}{18} \int \frac{d^d q}{(2\pi)^d} \frac{1}{(q^2 + m_0^2)^2} + \dots \right\}\end{aligned}\quad (79)$$

Show that, in zero dimension, with $u = \frac{1}{6}g_0$

$$\Gamma_{4,i_1 i_2 i_3 i_4} = (\delta_{i_1 i_2} \delta_{i_3 i_4} + \delta_{i_1 i_3} \delta_{i_2 i_4} + \delta_{i_1 i_4} \delta_{i_2 i_3}) \left\{ \Gamma^2 - \Gamma^3 \frac{\Gamma - 1}{(n+2)u} \right\}\quad (80)$$

where Γ is defined in (74). The quantity between braces can be developed as

$$\Gamma^2 - \Gamma^3 \frac{\Gamma - 1}{(n+2)u} = \frac{g_0}{3} - g_0^2 \frac{n+8}{18} + g_0^3 \frac{3n^2 + 46n + 140}{108} + \dots\quad (81)$$

There are of course numerous models with various internal symmetries. We will in particular encounter systems in which the dynamical variables are matrices.

5.2.5 Power counting

We want to generalize the observations made in the previous particular cases, which deal with the ultraviolet behaviour of the perturbative integrals. We assume $m_0^2 \neq 0$. In the Fourier representation, let us study the effect of a dilatation of all the momenta by the same factor. Consider a diagram with n external lines and v vertices. From (32) the number of independent integration variables is $B = l - v + 1 = 1 + v - \frac{1}{2}n$, if l denotes the number of internal lines.

In the case of a φ^4 interaction, a factor g_0 with a dimension $4 - d$ is associated to each vertex. The following discussion can be extended in an obvious way to the general case of local interactions. Similarly, the dimension two of the propagator associated to internal lines is to be modified accordingly if we study nonscalar fields. Let us also recall that the dimension of the vertex function Γ_n is given by

$$[\Gamma_n] = d_n = d - \frac{1}{2}n(d - 2)\quad (82)$$

The integrand has l quadratic denominators and B d -dimensional infinitesimal volume elements. A necessary (but not sufficient) condition for ultraviolet convergence is that the degree

$$\omega = Bd - 2l = d - \frac{1}{2}n(d - 2) - v(4 - d) = d_n - v(4 - d)\quad (83)$$

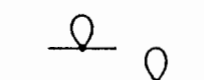
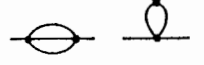
be negative. We assume of course $B \geq 1$ (i.e. $v \geq \frac{1}{2}n$) (there should be at least one integrall). The interpretation of equation

(83) is that the degree ω of the integral is such that its product with g_0^v has the degree of Γ_n . The diagram is *superficially convergent* if ω is negative. This is the case at sufficiently high orders for $d < 4$. Thus only a finite number of superficially divergent diagrams arises, corresponding to small values of n and v , for $d < 4$. The theory is then called *super-renormalizable*.

In four dimensions, the degree ω is independent of the number of vertices v and nonnegative for $n \leq 4$. The theory is called *renormalizable*. In higher dimension, the theory is not renormalizable and ultraviolet divergences become more and more severe with increasing perturbative order. We have already seen that the large-distance behaviour becomes trivial in this case.

The vertex functions with a nonnegative superficial degree of divergence to lowest nontrivial perturbative order ($d \geq n$) are called *primarily divergent*. These are the 2- and 4-point functions for $d = 4$, and only the 2-point function for $d = 2$ or 3.

Classify the superficially divergent diagrams in 2 and 3 dimensions. Study more carefully the nature of the divergence.

	$\omega = 0$	$\omega = 1$
$d = 2$		
$d = 3$		

For a scalar model, an integral with a nonnegative superficial degree of divergence is divergent. However, the converse statement might in general be false. For instance, functions with more than 4 points may be divergent at $d = 4$, because some subintegrals may diverge. The main aspect of the perturbative discussion is the study of this point. The results are the following.

- i) A divergence occurs *if and only if* there is at least one irreducible subdiagram with a nonnegative superficial degree. This corresponds to insertions of irreducible 2-point functions (*self-energy*) for $d < 4$. In dimension four, 4-point irreducible sub-diagrams should also be considered.

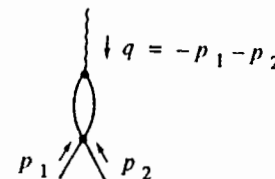


Fig. 3 First diagram contributing to $\Gamma_{2,1}$.

- ii) If the degree of a diagram together with those of all its subdiagrams is negative, the corresponding integral converges as expected.

These ultraviolet divergences are at first controlled by introducing, at an intermediate stage, a regularization governed by a cutoff factor Λ , already present in lattice versions. Divergences yield positive powers of Λ and $\ln \Lambda$ in the computation of integrals. Then a consistent subtraction procedure is required, leading to finite results in the limit $\Lambda \rightarrow \infty$.

We have also to consider vertex functions with insertions of composite operators, such as $\frac{1}{2}\varphi^2$. Let us recall the dimension of the function $\Gamma_{n,r}$ with r such insertions

$$[\Gamma_{n,r}] = d_n - 2r \equiv d - \frac{1}{2}n(d-2) - 2r \quad (84)$$

The corresponding superficial degree of divergence to order v of a perturbative contribution is

$$\omega = d_n - 2r - v(4-d) \quad (85)$$

The superficial degree of divergence is improved, i.e. decreases, with increasing number of insertions. For instance, $\Gamma_{2,1}$ is logarithmically divergent to all orders in four dimensions, as exemplified in the expression of the one loop diagram (figure 3)

$$\Gamma_{2,1}^a(p_1, p_2, q) = -\frac{1}{2}g_0 \int \frac{d^4k}{(2\pi)^4} \frac{1}{[k^2 + m_0^2][(q+k)^2 + m_0^2]} \quad (86)$$

Recall that $\Gamma_{2,1}(p, -p, 0)$ is the derivative of the 2-point function with respect to m_0^2 , $\partial \Gamma_2^a / \partial m_0^2$.

The function $\Gamma_{0,2}$ occurs in the determination of the specific heat. Its superficial degree of divergence to order v is

$$\omega = -(1+v)(4-d) \quad (87)$$

It is therefore ultraviolet convergent to any order for $d < 4$, logarithmically divergent to all orders (including 0) in four dimensions. Note that $\Gamma_{0,1}$ is a divergent constant related to the internal energy. However, it does not play any role in the contribution to subdiagrams since there is no external lines to be attached to a φ field.

This discussion also generalizes to insertions of more complex composite operators.

5.2.6 Perturbative renormalization

In dimension smaller than four, ultraviolet divergences only occur in the 2-point function, which also appears in subdiagrams pertaining to other correlation functions. In dimension four, there are also divergences in the 4-point function governing the coupling constant. We will proceed to renormalization according to the four-dimensional scheme. In lower dimension, some subtractions may appear unnecessary; however, they allow a smooth dimensional continuation and also play a role in the control of the infrared behaviour for $d < 4$.

In an abstract approach to renormalization, one defines subtractions in the integrands of the perturbative expressions in such a way as to obtain finite results without any auxiliary regularization. This technique has been developed by Bogoliubov, Parasiuk, Hepp and Zimmermann. It is, however, more instructive to use first a regularized theory. The perturbative expansion (considered as a series in the topological parameter counting the number of loops) is reexpressed in terms of *renormalized* (or *physical*) quantities. Here, these are the mass (the inverse correlation length) and the value of the 4-point vertex function for given values of the external momenta (the renormalized coupling constant). These renormalized quantities m and g are substituted for the *bare* quantities m_0 and g_0 initially present in the Lagrangian. Moreover, the field is rescaled by a factor $Z_1^{1/2}$: $\varphi = Z_1^{1/2} \varphi_R$. This process is called *wavefunction renormalization*. The corresponding correlation functions will remain finite when Λ increases indefinitely, and

all divergences are then concentrated into the relations between m , g , Z_1 and m_0 , g_0 , Λ .

The *renormalized* Green functions (correlations of φ_R fields) depend on external momenta and on m and g . These two constants are defined using normalization conventions. In a massive theory, these conventions fix the value of some vertex functions and of some of their derivatives at zero momenta. That this is sufficient to define a finite theory is due to a point that we have not yet emphasized, i.e. that all potentially divergent terms are at most polynomials in the external momenta, as indicated by power counting. An important consequence is that the counterterms which appear in the Lagrangian when it is expressed in terms of m , g , and φ_R are local. The locality property is a crucial feature of field theory, in that it allows control of the renormalized quantities in terms of finitely many parameters.

If one wants to study directly the massless critical theory, the above conventions should be changed, owing to potential infrared divergences. One has to fix renormalization conditions at nonzero momenta.

In the massive theory, we impose for instance the three normalization conditions

$$\left\{ \begin{array}{l} \Gamma_2^R(0) = m^2 \\ \frac{d\Gamma_2^R(0)}{dp^2} = 1 \\ \Gamma_4^R(0, 0, 0, 0) = m^{4-d}g \end{array} \right. \quad (88)$$

The normalization of Γ_4^R is such that the renormalized coupling constant g is dimensionless, in contradistinction with g_0 . In four dimensions, these three conditions correspond precisely to the three primitive divergences, the quadratic and logarithmic ones of the 2-point function Γ_2 and the logarithmic one of the 4-point function Γ_4 . The relation between *bare* regularized and *renormalized* Green functions reads for very large cutoff factor

$$\left\{ \begin{array}{l} G_n(p, m_0, g_0, \Lambda) = Z_1^{n/2} G_n^R(p, m, g) \\ \Gamma_n(p, m_0, g_0, \Lambda) = Z_1^{-n/2} \Gamma_n^R(p, m, g) \end{array} \right. \quad (89)$$

The first relation extends also to connected functions. The wavefunction renormalization factor Z_1 depends of course on Λ . Equations (88) and (89) allow the determination of the quantities

m_0^2 , g_0 and Z_1 in terms of m^2 , g and Λ as perturbative series in the number of loops

$$\begin{cases} m_0^2 = m_0^{2(0)} + m_0^{2(1)} + m_0^{2(2)} + \dots \\ g_0 = g_0^{(0)} + g_0^{(1)} + g_0^{(2)} + \dots \\ Z_1 = Z_1^{(0)} + Z_1^{(1)} + Z_1^{(2)} + \dots \end{cases} \quad (90)$$

The above normalization conventions have been chosen in such a way that the bare and renormalized constants are identical to lowest order ($B = 0$), except for a trivial scaling factor in the coupling constant.

$$\begin{cases} m_0^{2(0)} = m^2 \\ g_0^{(0)} = m^{4-d}g \\ Z_1^{(0)} = 1 \end{cases} \quad (91)$$

Hence, to this order, the functions Γ_2^R and Γ_4^R are

$$\begin{aligned} \Gamma_2^R &= p^2 + m^2 \\ \Gamma_4^R &= m^{4-d}g \end{aligned} \quad (92)$$

As a consequence, normalization conditions are homogeneous at higher orders. These operations can be summarized by expressing the Lagrangian in terms of the renormalized field, Z_1 , m and g . It takes the form

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} Z_1 [(\partial\varphi_R)^2 + m_0^2 \varphi_R^2] + \frac{1}{4!} g_0 Z_1^2 \varphi_R^4 \\ &= \frac{1}{2} [(\partial\varphi_R)^2 + m^2 \varphi_R^2] + \frac{1}{4!} m^{4-d} g \varphi_R^4 + \\ &\quad \sum_{k=1}^{\infty} \left[Z_1^{(k)} \frac{1}{2} (\partial\varphi_R)^2 + (Z_1 m_0^2)^{(k)} \frac{1}{2} \varphi_R^2 + \frac{1}{4!} (g_0 Z_1^2)^{(k)} \varphi_R^4 \right] \end{aligned} \quad (93)$$

This expression contains counterterms of order $B = 1, 2, \dots$, the effect of which has to be included order by order. For instance, counterterms of order 1 are to be taken into account at the one-loop order, while the irreducible subdiagrams (here, vertex and propagators) are to be computed at a strictly lower order, here 0. For higher orders, this method of including counterterms explicitly solves a complex combinatorial problem, implying internal subtractions in the diagrams.

Let us compute as an example the renormalized quantities to order one.

$$\begin{aligned} \Gamma_2^{R(1)}(p^2) &= \frac{1}{2} m^{4-d} g \int_{\Lambda} \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} + (Z_1 m_0^2)^{(1)} + Z_1^{(1)} p^2 \\ \Gamma_4^{R(1)}(p_1, p_2, p_3, p_4) &= -\frac{1}{2} m^{8-2d} g^2 \\ &\quad \times \sum_{i=2}^4 \int_{\Lambda} \frac{d^d q}{(2\pi)^d} \frac{1}{[q^2 + m^2][(q - p_1 - p_i)^2 + m^2]} + (g_0 Z_1^2)^{(1)} \end{aligned} \quad (94)$$

To first order, the renormalization conditions (88) require

$$\Gamma_2^{R(1)}(0) = 0, \quad \frac{d}{dp^2} \Gamma_2^{R(1)}(0) = 0, \quad \Gamma_4^{R(1)}(0, 0, 0, 0) = 0$$

hence

$$Z_1^{(1)} = 0$$

$$(Z_1 m_0^2)^{(1)} = m_0^{2(1)} = -\frac{1}{2} m^{4-d} g \int_{\Lambda} \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} \quad (95)$$

and using our earlier results (71) and (72),

$$\begin{aligned} (g_0 Z_1^2)^{(1)} &= g_0^{(1)} = \frac{3}{2} m^{8-2d} g^2 \int_{\Lambda} \frac{d^d q}{(2\pi)^d} \frac{1}{(q^2 + m^2)^2} \\ &= \begin{cases} \frac{3}{2} m^{4-d} g^2 \frac{1}{(4\pi)^{d/2}} \Gamma(2-d/2) & \text{for } d < 4 \\ \frac{3}{2} g^2 \frac{1}{(4\pi)^2} (\ln \frac{\Lambda^2}{m^2} - \ln 2 - \gamma - 1) & \text{for } d = 4 \end{cases} \end{aligned} \quad (96)$$

Finally,

$$\Gamma_2^{R(1)}(p) = 0$$

$$\begin{aligned} \Gamma_4^{R(1)}(p_1, p_2, p_3, p_4) &= -\frac{1}{2} m^{8-2d} g^2 \sum_{i=2}^4 \int_{\Lambda} \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} \\ &\quad \times \left\{ \frac{1}{(q - p_1 - p_i)^2 + m^2} - \frac{1}{q^2 + m^2} \right\} \end{aligned} \quad (97)$$

In dimension less than four, $g_0^{(1)}$ remains finite. The subtracted integrals define convergent renormalized functions and are expressed in terms of the physical quantities m and g . At the one-loop level, there is no wavefunction renormalization in the φ^4 theory; it appears only at the two-loop level, as we have already seen.

One can proceed similarly for higher orders. The counterterms defined at lower orders compensate the internal divergences of the diagrams and the normalization conditions define the new counterterms which compensate the global divergences. This is a recursive procedure which defines both the counterterms of the bare Lagrangian and the renormalized functions correctly subtracted. In a super-renormalizable theory, this process is finite and counterterms are polynomials in the coupling constant. It continues indefinitely in a renormalizable theory.

It is a rather delicate matter to analyze the procedure to all orders. The main result is that the successive terms of the expansion for the correlation functions are subtracted convergent integrals. Furthermore their structure is independent from the original regularization procedure. Indeed, one can develop an algorithm for the construction of their integrands using only the normalization conditions. The regularization has the advantage of yielding expressions for g_0 , m_0 and Z_1 in terms of g , m and Λ , at least when it is consistently implemented. Of course, the cutoff factor disappears in the renormalized theory. As a consequence, the latter depends on one dimensionful parameter less than the bare regularized theory. The proof of these statements is at the heart of the renormalization theory and requires a sophisticated mathematical apparatus which will not be developed here.

It is useful to comment on the effects of the regularization and of the renormalization on the symmetries of a model. It is clear, for instance, that if a bare theory is seemingly invariant under dilatations, this symmetry is broken by regularization. We will study this effect in the renormalized theory, using the equations governing changes of scale. On the other hand, the global symmetry $\varphi \rightarrow -\varphi$ of the scalar case, and more generally the $O(n)$ symmetry of the n component vector model are broken neither by regularization nor by renormalization. There will exist a symmetric phase and spontaneous symmetry breaking will be a consequence of dynamics. The question of scale invariance is but one aspect of the quest for *anomalies* to apparent symmetries in the bare Lagrangian, or equivalently in the classical theory. Such anomalies may occur when functional integrals are more precisely defined.

The introduction of composite operators in the correlation functions leads to new divergences. Let us examine the insertions

of $\frac{1}{2}\varphi^2$. In four dimensions, $\Gamma_{2,1}$ and $\Gamma_{0,2}$ are logarithmically divergent. The function $\Gamma_{2,1}$ is dimensionless and is normalized using its value at order 0

$$\Gamma_{2,1}^R(p=0) = 1 \quad (98)$$

where p represents symbolically all momenta. The relation between bare and renormalized functions introduces a new divergent constant Z_2 , which we define through

$$\Gamma_{2,1}(p, g_0, m_0, \Lambda) = Z_1^{-1} Z_2 \Gamma_{2,1}^R(p, g, m) \quad (99)$$

Write the corresponding relation for correlation functions. The function $G_{2,1}$ is such that

$$G_{2,1} = G_2 \Gamma_{2,1} G_2 = Z_1 G_2^R Z_1^{-1} Z_2 \Gamma_{2,1}^R Z_1 G_2^R = Z_1 Z_2 G_{2,1}^R$$

When the momentum associated to $\frac{1}{2}\varphi^2$ vanishes, one has

$$G_{2,1}(p_1, p_2 = -p_1, q = 0) = -\frac{\partial}{\partial m_0^2} G_2(p_1, p_2 = -p_1)$$

while

$$\Gamma_{2,1}(p_1, p_2 = -p_1, q = 0) = \frac{\partial}{\partial m_0^2} \Gamma_2(p_1, p_2 = -p_1)$$

The relation (99) expresses the multiplicative renormalization of one insertion of $\frac{1}{2}\varphi^2$

$$\frac{1}{2}\varphi^2 = \frac{1}{2}Z_2(\varphi^2)^R \quad (100)$$

Note that $(\varphi^2)^R(x)$ differs from $(\varphi^R(x))^2$ due to additional divergences appearing when the arguments of two fields coincide. This results in a divergent ratio Z_1/Z_2 . The same factor Z_2 renormalizes a single insertion of $\frac{1}{2}\varphi^2$ in *all* correlation functions.

As previously, Z_2 can be expanded according to the number of loops

$$Z_2 = 1 + Z_2^{(1)} + Z_2^{(2)} + \dots \quad (101)$$

As $Z_1^{(1)} = 0$, equations (98) and (99) lead to the expression (figure 3)

$$\begin{aligned} 1 + Z_2^{(1)} + \Gamma_{2,1}^{R(1)}(p_1, p_2, q) \\ = 1 - \frac{1}{2}m^{4-d}g \int_{\Lambda} \frac{d^d k}{(2\pi)^d} \frac{1}{[k^2 + m^2][(k+q)^2 + m^2]} \end{aligned} \quad (102)$$

$$Z_2^{(1)} = -\frac{1}{2}m^{4-d}g \int_{\Lambda} \frac{d^d k}{(2\pi)^d} \frac{1}{[k^2 + m^2]^2}$$

$$= \begin{cases} -\frac{1}{2}g(4\pi)^{-d/2}\Gamma(2-d/2) & d < 4 \\ -\frac{1}{2}g(4\pi)^{-2}(\ln \Lambda^2/m^2 - \ln 2 - 1 - \gamma) & d = 4 \end{cases} \quad (103)$$

$$\Gamma_{2,1}^{R(1)}(p_1, p_2, q) = -\frac{1}{2}m^{4-d}g \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 + m^2}$$

$$\left\{ \frac{1}{(k+q)^2 + m^2} - \frac{1}{k^2 + m^2} \right\} \quad (104)$$

The divergence of $\Gamma_{0,2}$ containing two operators $\frac{1}{2}\varphi^2$, which does not appear as subdivergence in any diagram, is not totally removed by the Z_2 renormalization. Note that it occurs even for the free field, i.e. to order zero. An additional subtraction is needed for its compensation,

$$\Gamma_{0,2}(p, -p) = Z_2^2 \left\{ \Gamma_{0,2}^R(p, -p) + A \right\} \quad (105)$$

$$\Gamma_{0,2}^R(0, 0) = 0$$

where the normalization implies that the corresponding bare term is included in the specific heat.

Finally, the quantity $\Gamma_{0,1} = \langle \frac{1}{2}\varphi^2(0) \rangle$ diverges quadratically in four dimensions. This constant may be used in the computation of the internal energy, but will not play a role in the following. Insertions of operators of higher degree will be studied later.

5.3 The renormalization group

5.3.1 Renormalization flow

For several reasons, the previous discussion might superficially look irrelevant in the study of critical phenomena. First, we were initially interested in the bare theory. Indeed, the coefficients which are directly related to the temperature are those of the original Lagrangian. Secondly, there is no reason to assume that the coupling constant g_0 is small. In dimension d smaller than four, this is rather a large quantity, due to its positive dimension $4-d$ in terms of the original scale $1/a$. A perturbative expansion is thus not suitable, except if we succeed in extracting information

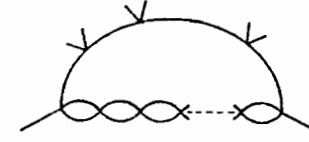


Fig. 4 Example of a diagram with a chain of "bubbles" leading to infrared divergences in a massless theory.

after resummation. Finally, we have discussed the short-range singularities, although we are interested in the infrared domain where the correlation length is infinite. For $d < 4$, the infrared divergences at $m = 0$ are more and more severe as the perturbative order increases. For instance, if p represents the momentum scale, the 4-point function behaves as $p^{-(4-d)}$. Hence its iteration n times will produce infrared divergences if $n(4-d) > d$ in diagrams which are ultraviolet convergent (figure 4).

To control infrared divergences, we will keep a nonzero mass m for $d < 4$. The critical theory appears thus in a domain $p/m \gg 1$ and $p/\Lambda \ll 1$. From the point of view of the renormalized theory, the scale Λ has been absorbed in a redefinition of T_c by the use of the correlation length m^{-1} as a fundamental parameter and by a rescaling of the field (the magnetization). The second condition $p/\Lambda \ll 1$ is then automatically fulfilled. The first condition shows that there is a relation between the critical behaviour and the short-distance analysis of the renormalized theory in a scale $1/m$. In order to obtain the scale θ of the deviations of the temperature from T_c , the relation between m and m_0 has to be elucidated. These indications suggest how the problems raised above will be solved. One will choose as an expansion parameter the dimensionless renormalized coupling constant g , which will turn out to remain finite whatever the value of $g_0 \geq 0$. For $d < 4$, one keeps a finite mass in the symmetric (high temperature) phase and discusses the corresponding ultraviolet behaviour. On the other hand, there are no infrared difficulties when $d = 4$. In this case, it is possible to study directly the massless theory, which is related to the $g \rightarrow 0$ limit. In this sense, we obtain an infrared asymptotically free theory in four dimensions. It will then be possible to proceed to a double expansion in g and $\epsilon = 4 - d$.

Consequently, the applications of renormalization theory to critical phenomena can be presented either as a study at fixed $d < 4$ (2 or 3), or as an ϵ -expansion. We begin with a survey from the first point of view.

Let us assume $d < 4$, $m > 0$. We shall see that the critical theory corresponds to the limit where the bare coupling goes to infinity, $g_0 \rightarrow \infty$. The essential hypothesis is that the correlation functions, renormalized following the scheme valid at $d = 4$ (renormalizations of the mass, the coupling constant and the wavefunction) remain finite as $g_0 \rightarrow \infty$. This hypothesis seems to be verified when d is close to 4. We know from quantum mechanics that scattering amplitudes approach the Born terms at high energy. Similarly, the super-renormalizability of the φ^4 theory suggests that the Green functions tend towards the leading terms of their perturbative expansion when the momenta increase indefinitely. In particular, the 4-point vertex function tends towards $g_0 \rightarrow \infty$. The hypothesis amounts to saying that the value at fixed momentum remains finite although the short-distance behaviour exhibits ultraviolet divergences.

To investigate this question, let us consider a change of scale in the correlation functions. This can be done by varying the mass m at fixed g_0 and Λ . For a bare vertex function, we obtain

$$\begin{aligned} m \frac{\partial}{\partial m} \Gamma_n(\mathbf{p}; m_0, g_0, \Lambda) \Big|_{g_0, \Lambda} \\ = \left(m \frac{\partial}{\partial m} m_0^2 \Big|_{g_0, \Lambda} \right) \Gamma_{n,1}(\mathbf{p}, \mathbf{q} = 0; m_0, g_0, \Lambda) \end{aligned} \quad (106)$$

The bare and renormalized functions are related through

$$\begin{aligned} \Gamma_n(\mathbf{p}; m_0, g_0, \Lambda) &= Z_1^{-n/2} \Gamma_n^R(\mathbf{p}; m, g) \\ \Gamma_{n,1}(\mathbf{p}, \mathbf{q}; m_0, g_0, \Lambda) &= Z_1^{-n/2} Z_2 \Gamma_{n,1}^R(\mathbf{p}, \mathbf{q}; m, g) \end{aligned}$$

We introduce the dimensionless quantities

$$\begin{aligned} \beta &= m \frac{\partial g}{\partial m} \Big|_{g_0, \Lambda} \\ \gamma_1 &= Z_1^{-1} m \frac{\partial Z_1}{\partial m} \Big|_{g_0, \Lambda} \\ \gamma_2 &= Z_2^{-1} m \frac{\partial Z_2}{\partial m} \Big|_{g_0, \Lambda} \end{aligned} \quad (107)$$

and express equation (106) in terms of renormalized functions as

$$\begin{aligned} \left\{ m \frac{\partial}{\partial m} + \beta \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1 \right\} \Gamma_n^R(\mathbf{p}; m, g) \\ = Z_2 \left(m \frac{\partial m_0^2}{\partial m} \Big|_{g_0, \Lambda} \right) \Gamma_{n,1}^R(\mathbf{p}, 0; m, g) \end{aligned}$$

Note that β , γ_1 and γ_2 (which does not yet occur) are *a priori* functions of g and m/Λ , but do not depend on n and on the momenta. In particular, if we choose $n = 2$ and $\mathbf{p} = 0$, we have

$$\Gamma_2^R(\mathbf{p} = 0; m, g) = m^2 \quad \text{and} \quad \Gamma_{2,1}^R(\mathbf{p} = 0, \mathbf{q} = 0; m, g) = 1$$

Therefore,

$$(2 - \gamma_1) m^2 = Z_2 m \frac{\partial m_0^2}{\partial m} \Big|_{g_0, \Lambda} \quad (108)$$

Inserting this expression in the r.h.s. of the above relation, any explicit reference to the bare theory disappears, and thus the coefficients β and γ_1 either depend only on g or are infinite. This last possibility can be excluded by expressing these coefficients in terms of the renormalized vertex functions for particular values of the momenta. Hence one obtains the set of equations

$$\begin{aligned} \left\{ m \frac{\partial}{\partial m} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) \right\} \Gamma_n^R(\mathbf{p}; m, g) \\ = (2 - \gamma_1(g)) m^2 \Gamma_{n,1}^R(\mathbf{p}, 0; m, g) \end{aligned} \quad (109)$$

Equivalent forms of such relations were first derived in the context of quantum electrodynamics, in a discussion involving the arbitrariness of renormalization conditions. The subject was brought back to life by Callan and Symanzik in an analysis of anomalies pertaining to scale invariance. The quantities β , γ_1 and γ_2 , which partly depend on the normalization conditions, have finite perturbative expansions in g , and equation (109) relates only renormalized quantities. It remains valid at $d = 4$.

The derivation of these equations is based in an essential way on the multiplicative renormalization of the fields and the fact

that the dimensionful cutoff factor Λ disappears in the renormalized theory expressed in terms of m and g . The multiplicative renormalization is responsible for the factor $\gamma_1(g)$, related to the anomalous dimension of the field φ , while $\gamma_2(g)$ is similarly related to the anomalous dimension of $\frac{1}{2}\varphi^2$. The interesting phenomenon is the appearance of the factor $\beta(g)$. It expresses the fact that a change in scale is accompanied by a flow of the coupling. This phenomenon would obviously generalize to a multidimensional case of several coupling constants. The *critical points* where $\beta(g_c) = 0$ play the crucial role of attractive or repulsive points. This is not completely unexpected for $d < 4$, since the dimensionful initial coupling constant would have to change with scale. The fact might seem more surprising in dimension four since both g and g_0 are dimensionless. However, it then appears as a relic of the existence of a cutoff Λ in the underlying regularized theory. Let us point out also that the right-hand side in equation (109) shows that these equations are the first in a hierarchy involving more insertions of $\frac{1}{2}\varphi^2$

$$\left\{ m \frac{\partial}{\partial m} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) + s \gamma_2(g) \right\} \Gamma_{n,s}^R(\{p\}, \{q\}; m, g) \\ = (2 - \gamma_1(g)) m^2 \Gamma_{n,s+1}^R(\{p\}, \{q, 0\}; m, g) \quad (110)$$

Derive equations (110).

This system might seem of little utility, except in cases where there are regions in momentum space where the right-hand side is negligible. Let us consider for instance the relations (109). Although the momentum configuration of the r.h.s. is in a sense exceptional (the insertion of φ^2 is at zero momentum), $\Gamma_{n,1}^R(p, 0)$ is of order $p^{-2} \Gamma_n^R(p)$, up to powers of $\ln p$, for large p , order by order in a perturbative expansion. This is an example of a general theorem obtained by Weinberg. If these logarithms do not sum up to compensate the factor p^{-2} , this r.h.s. can be neglected in the region of large momenta. The coherence of this procedure can be checked on the next equation of the hierarchy (110). Finally, near $d = 4$, one can expect that

the deviations from canonical dimensions are at least of order $(4 - d)$.

(1) Show that equation (108) can be re-expressed as a relation between Z_2 and the other renormalization constants, in the form

$$Z_2 = Z_1 \frac{\partial}{\partial m_0^2} \left(\frac{m^2}{Z_1} \right)_{g_0, \Lambda} \quad (111)$$

(2) In the limit $\Lambda \rightarrow \infty$, g_0/m^{4-d} is a function of g only (if g and m are finite), so that $\beta(g)$ is obtained for $d < 4$ by the relation

$$\left\{ 4 - d + \beta(g) \frac{d}{dg} \right\} \frac{g_0}{m^{4-d}} = 0 \quad (112)$$

As g_0/m^{4-d} tends towards g for $g \rightarrow 0$, it follows that

$$\beta(g) \underset{g \rightarrow 0}{\sim} -(4 - d)g + \mathcal{O}(g^2) \quad (113)$$

As an illustration, let us present the one-loop computation of the coefficients β , γ_1 , γ_2 . First, let us summarize the expressions of g_0 , Z_1 , Z_2 and m_0^2 in both cases $d < 4$ and $d = 4$. We also include the symmetry factors for an n -vector model with a $(\Phi^2)^2$ interaction.

$$d < 4 \left\{ \begin{array}{l} g_0 = m^{4-d} \left\{ g + \frac{n+8}{6} \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} g^2 + \dots \right\} \\ Z_1 = 1 + \dots \\ m_0^2 = m^2 \left\{ 1 - \frac{n+2}{6} \frac{\Gamma(1-d/2)}{(4\pi)^{d/2}} g + \dots \right\} \\ Z_2 = 1 - \frac{n+2}{6} \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} g + \dots \end{array} \right. \quad (114a)$$

$$d = 4 \left\{ \begin{array}{l} g_0 = g + \frac{n+8}{6} \frac{\ln \frac{\Lambda^2}{m^2} - \ln 2 - 1 - \gamma}{(4\pi)^2} g^2 + \dots \\ Z_1 = 1 + \dots \\ m_0^2 = m^2 \left\{ 1 - \frac{n+2}{6} \frac{\frac{\Lambda^2}{m^2} - 1 - \ln \frac{\Lambda^2}{m^2}}{(4\pi)^2} g + \dots \right\} \\ Z_2 = 1 - \frac{n+2}{6} \frac{\ln \frac{\Lambda^2}{m^2} - \ln 2 - 1 - \gamma}{(4\pi)^2} g + \dots \end{array} \right. \quad (114b)$$

The series are inverted and lead to

$$d < 4 \left\{ \begin{array}{l} g = \frac{g_0}{m^{4-d}} - \frac{n+8}{6} \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} \left(\frac{g_0}{m^{4-d}} \right)^2 + \dots \\ m \frac{\partial g}{\partial m} \Big|_{g_0, \Lambda} = -(4-d) \frac{g_0}{m^{4-d}} \\ \quad + \frac{n+8}{3} \frac{(4-d)\Gamma(2-d/2)}{(4\pi)^2} \left(\frac{g_0}{m^{4-d}} \right)^2 + \dots \\ \beta(g) = -(4-d)g + \frac{n+8}{3} \frac{\Gamma(3-d/2)}{(4\pi)^{d/2}} g^2 + \dots \end{array} \right. \quad (115a)$$

$$d = 4 \left\{ \begin{array}{l} g = g_0 - \frac{n+8}{6} \frac{\ln \frac{\Lambda^2}{m^2} - \ln 2 - 1 - \gamma}{(4\pi)^2} g_0^2 + \dots \\ m \frac{\partial g}{\partial m} \Big|_{g_0, \Lambda} = \frac{n+8}{3} \frac{g_0^2}{(4\pi)^2} + \dots \\ \beta(g) = \frac{n+8}{3} \frac{g^2}{(4\pi)^2} + \dots \end{array} \right. \quad (115b)$$

Note that the limit $d \rightarrow 4$ is regular term by term and yields a function $\beta(g)$ perturbatively finite, as a consequence of the renormalizability of the theory.

For $d \leq 4$, one gets

$$\gamma_1(g) = \mathcal{O}(g^2) \quad (116)$$

$$\begin{aligned} \gamma_2(g) &= \left(\frac{n+2}{3} \right) \frac{\Gamma(3-d/2)}{(4\pi)^{d/2}} g + \dots \\ &= \left(\frac{n+2}{3} \right) \frac{g}{(4\pi)^2} \quad (d=4) \end{aligned} \quad (117)$$

Compute $\beta(g)$ using equation (113) and check (116) and (117).

For large Λ and fixed g_0 , the variation of g is related to the last free parameter, i.e. the temperature or equivalently, the correlation length. As this length increases, m decreases and g varies according to the sign of $-\beta(g)$. For small g , we have just seen that β is negative for $d < 4$. Recall that the reasoning is valid in the symmetric phase ($T > T_c$). Therefore, as T decreases towards T_c , m decreases and g increases. If g is bounded, this growth has to stop and $\beta(g)$ should vanish for some positive value g_c , as suggested by the one-loop formula. Moreover, as the

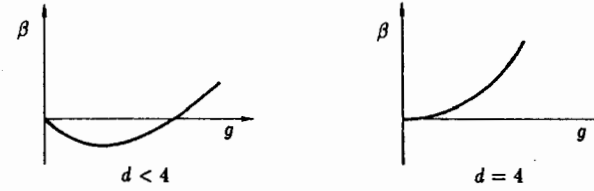


Fig. 5 Behaviour of the function $\beta(g)$.

coefficients of the expansion have a finite limit when $d \rightarrow 4$, and if the perturbative expansion is asymptotically related to a well-defined function, g_c is of order $4-d$ for $d \rightarrow 4$. This hypothetical behaviour of $\beta(g)$ is displayed in figure 5. Of course, we only consider the stable perturbative region $g > 0$, although the point $g = 0$ does not seem to be singular at this level.

Using the terms explicitly given in (115), one obtains

$$\frac{g_c}{(4\pi)^{d/2}} = \left(\frac{3}{n+8} \right) \frac{4-d}{\Gamma(3-d/2)} + \dots \quad (118)$$

behaving as $(4-d)3/(n+8)$ as $d \rightarrow 4$. In three dimensions (although this calculation might seem unrealistic in this case!), we find

$$\frac{g_c}{16\pi} = \frac{3}{n+8} + \dots \quad (119)$$

Covariance of the flow in coupling space The notion of a fixed point is independent of a regular reparametrization in coupling space. For the φ^4 model, there is only one coupling constant. Given an invertible transformation $g' = g'(g)$, the equality of vector fields

$$\beta(g) \frac{\partial}{\partial g} = \beta'(g') \frac{\partial}{\partial g'} \quad (120a)$$

yields the transformation law

$$\beta'(g') = \beta(g) \frac{dg'}{dg} \quad (120b)$$

If the transformation is monotonic – as is required if it is to be invertible – there is a one-to-one correspondence between the zeroes of β and β' as well as between the sign of the functions. Let $g_c \leftrightarrow g'_c$ be a fixed point where both functions vanish. The slope is

a reparametrization invariant, according to

$$\omega' = \frac{d\beta'(g_c)}{dg'} = \frac{dg}{dg'} \frac{d}{dg} \left(\beta(g) \frac{dg'}{dg} \right)_{g_c} = \frac{d\beta(g_c)}{dg} = \omega \quad (121)$$

In the particular case of a degenerate fixed point at the origin,

$$\begin{aligned} \beta(g) &= \beta_3 g^3 + \beta_5 g^5 + \mathcal{O}(g^7) \\ g' &= g + ag^3 + \dots \end{aligned} \quad (122a)$$

we observe that in

$$\beta'(g') = \beta_3 g'^3 + \beta_5 g'^5 + \mathcal{O}(g'^7) \quad (122b)$$

the first two coefficients are reparametrization invariants. Generalize these properties to the case of a multidimensional coupling constant space.

When $d < 4$, as the coupling varies with m according to (107), the quantities Z_1 , Z_2 and $g_0/m^{4-d}g$ can be computed by integrating the flow equations (107) and (112) using their values 1 for $g = 0$,

$$\begin{aligned} Z_1 &= \exp \int_0^g dg' \frac{\gamma_1(g')}{\beta(g')} \\ Z_2 &= \exp \int_0^g dg' \frac{\gamma_2(g')}{\beta(g')} \\ \frac{g_0}{m^{4-d}} &= g \exp - \int_0^g dg' \left(\frac{4-d}{\beta(g')} + \frac{1}{g'} \right) \end{aligned} \quad (123)$$

Looking at the last relation at fixed g_0 and $m \rightarrow 0$, one sees that g has to increase up to the first zero of $\beta(g)$ where g_0/m^{4-d} diverges. This can be reinterpreted at fixed m by saying that, as g runs from 0 to g_c , g_0 varies from 0 to infinity.

Suppose that g_c is a simple zero of $\beta(g)$ with a nonvanishing slope ω , as suggested in figure 5,

$$\omega = \beta'(g_c) \quad (124)$$

Close to this point, one gets

$$\begin{aligned} \frac{g_0}{m^{4-d}} &\cong A \left(1 - \frac{g}{g_c} \right)^{-(4-d)/\omega} \\ A &= g_c \exp - \int_0^{g_c} dg' \left(\frac{4-d}{\beta(g')} + \frac{1}{g'} - \frac{4-d}{\omega(g' - g_c)} \right) \end{aligned} \quad (125a)$$

or equivalently,

$$g = g_c \left[1 - \left(\frac{Am^{4-d}}{g_0} \right)^{\omega/(4-d)} \right] \quad (125b)$$

This relation involves fractional powers of g_0 , showing its nonperturbative nature. As g_c behaves as $4-d$, this expression becomes meaningless when $d \rightarrow 4$ where the interval in which g is allowed to vary shrinks to zero. The inescapable and rather astounding conclusion is that the renormalized ϕ^4 theory obtained from a cutoff regularized model becomes free in four dimensions! On second thought, this ties up very nicely with the idea that four is an upper critical dimension. To investigate this question in a meaningful way, one can consider the case where the cutoff factor is large, but not infinite. It leaves a small interval for the variation of g of order Λ^{-1} . Alternative methods, in particular numerical ones, seem to confirm the triviality of the four-dimensional ϕ^4 theory.

Show that, if the slope ω at the fixed point vanishes, i.e. for a degenerate double zero with $\beta''(g_c) \neq 0$, and $\beta(g) \sim -\frac{1}{2}\omega'(g - g_c)^2$, then

$$g_c - g \sim \frac{2(4-d)}{\omega' \ln(g_0/(A'm^{4-d}))} \quad (125c)$$

What is the corresponding expression for higher order zeroes? One notes that, as g_0 goes to infinity, g approaches g_c slower and slower.

The coefficients $\gamma_1(g)$ and $\gamma_2(g)$ have no reason to be singular at the critical point. In a generic situation $\omega \neq 0$, we get the following behaviours

$$\begin{aligned} Z_1 &\sim C_1 m^{\gamma_1(g_c)} \left[1 + \bar{C}_1 \left(m g_0^{1/(d-4)} \right)^\omega + \dots \right] \\ Z_2 &\sim C_2 m^{\gamma_2(g_c)} \left[1 + \bar{C}_2 \left(m g_0^{1/(d-4)} \right)^\omega + \dots \right] \end{aligned} \quad (127)$$

The critical theory is characterized by scaling laws in powers of the correlation length. We shall now relate them to the standard exponents.

5.3.2 Critical exponents

The present discussion is relative to the symmetric high temperature phase where

$$m_0^2 \sim a + b\theta \quad \theta = \frac{T - T_c}{T_c} > 0 \quad (128)$$

Recall the definition of the exponents ν and γ

$$m = \xi^{-1} \sim \theta^\nu$$

$$\chi = G_2(p=0) = \frac{Z_1}{m^2} \sim \theta^{-\gamma} \quad (129)$$

and the relation

$$Z_2 = \frac{\partial m^2}{\partial m_0^2} \Big|_{g_0, \Lambda} (1 - \tfrac{1}{2}\gamma_1(g))$$

where $1 - \frac{1}{2}\gamma_1(g_c)$ generically does not vanish.

Analyze the meaning of $\gamma(g_c) = 2$.

When $m \rightarrow 0$, the l.h.s. of this last equality behaves as $m^{\gamma_2(g_c)} \sim \theta^{\nu\gamma_2(g_c)}$ according to equation (127), while the r.h.s. is proportional to $dm^2/d\theta \sim \theta^{2\nu-1}$. Hence $\gamma_2(g_c)$ is related to ν through

$$\nu = \frac{1}{2 - \gamma_2(g_c)} \quad (130)$$

Similarly from equations (127, 129), one gets

$$\frac{\gamma}{\nu} = 2 - \gamma_1(g_c) \quad (131)$$

Using equation (4.11), this relation is equivalent to

$$\eta = \gamma_1(g_c) \quad (132)$$

which relates the exponent η to the abnormal dimension of the field φ . The two quantities $\gamma_1(g_c)$ and $\gamma_2(g_c)$ will yield all exponents, provided that the critical theory satisfies the standard scaling relations. To prove this fact, we will have to look at the behaviour of correlation functions. Observe also that, from equations (127), the exponent ω characterizes the deviations from critical behaviour. The coefficients of the terms in $m^\omega \sim \theta^{\omega\nu}$ depend on the details of microscopic parameters.

Let us apply now the results of the one-loop calculations to the computation of critical exponents. One gets

$$\beta(g) = -(4-d)g(1-g/g_c) \quad g_c = \frac{4-d}{n+8} \frac{3(4\pi)^{d/2}}{\Gamma(3-d/2)} \quad (133)$$

and therefore

$$\omega = 4 - d + \dots \quad (134b)$$

At the one-loop order, $\gamma_1(g)$ vanishes. Thus

$$\eta = 0 + \dots \quad (134c)$$

Finally from (117)

$$\gamma_2(g_c) = \frac{g_c}{(4\pi)^{d/2}} \frac{n+2}{3} \Gamma(3-d/2) = (4-d) \frac{n+2}{n+8} + \dots$$

yielding

$$\nu = \frac{1}{2 - (4-d)(n+2)/(n+8) + \dots}$$

$$\gamma = \frac{2 + \dots}{2 - (4-d)(n+2)/(n+8) + \dots} \quad (134d)$$

These one loop results yield the exact $n \rightarrow \infty$ limit (chapter 3)

$$\eta = 0 \quad \frac{1}{2}\gamma = \nu = 1/(d-2) \quad (135)$$

Explain why.

The predictions of the renormalization group are already non-trivial at the one-loop level. Deviations from the mean field predictions are of order $\varepsilon = 4 - d$. Close to four dimensions, this procedure is coherent. Indeed, g_c is of order ε and this fact justifies the renormalized perturbative expansion used to compute the coefficients $\beta(g)$, $\gamma_1(g)$, $\gamma_2(g)$, ... of the renormalization group equations. The whole scheme assumes that these equations have a domain of validity independent from the perturbative method. Many efforts are presently being made to justify this point.

It is remarkable that setting $d = 3$ in the one-loop expressions for the exponents yields values very close to the correct ones. For the Ising model ($n = 1$) for instance, one gets

$$\nu = \frac{1}{2}\gamma = 0.6 \quad \eta = 0 \quad (d = 3, n = 1) \quad (136)$$

to be compared with the values given in table I which appears later in this chapter.

5.3.3 From the Gaussian ultraviolet fixed point to the infrared critical point in dimension less than four

The proof of equations (123) implicitly assumes a trivial behaviour as $g \rightarrow 0$. More generally, in dimension less than four, we expect a trivial short-distance behaviour of the massive φ^4 model, which corresponds to a super-renormalizable theory, i.e. it has a free field behaviour, corrected by the first nonvanishing perturbative corrections. If this is the case, the bare functions tend toward their tree level expressions (Born amplitudes) for large momenta. Hence,

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \lambda^{-2} p^{-2} \Gamma_2^R(\lambda p; m, g) &= Z_1 \\ \lim_{\lambda \rightarrow \infty} \Gamma_4^R(\{\lambda p\}; m, g) &= m^{4-d} \frac{g_0}{m^{4-d}} Z_1^2 \\ \lim_{\lambda \rightarrow \infty} \Gamma_{2,1}^R(\{\lambda p\}, \lambda q; m, g) &= Z_1 Z_2^{-1} \end{aligned} \quad (137)$$

In the last two formulae, all points in configuration space should become coincident in the limit. Hence, one imposes the following technical restriction that none of the partial sums of momenta vanishes.

The limits in (137) state that fluctuations do not modify the dominant perturbative behaviour. This can be checked order by order in the renormalized theory for $d < 4$.

Verify this assertion at the one-loop order. For instance, for $n = 1$,

$$\begin{aligned} \Gamma_2^R &= m^2 + p^2 + \dots \\ \Gamma_4^R &= m^{4-d} g - \frac{1}{2} g^2 m^{4-d} \sum_{i=2}^4 \int \frac{d^d q}{(2\pi)^d} \\ &\quad \frac{1}{q^2 + m^2} \left(\frac{1}{(p_1 + p_i - q)^2 + m^2} - \frac{1}{q^2 + m^2} \right) + \dots \\ \Gamma_{2,1}^R &= 1 - \frac{1}{2} m^{4-d} g \int \frac{d^d q}{(2\pi)^d} \\ &\quad \frac{1}{q^2 + m^2} \left(\frac{1}{(p_1 + p_2 - q)^2 + m^2} - \frac{1}{q^2 + m^2} \right) + \dots \end{aligned}$$

Hence, from (137),

$$\begin{aligned} Z_1 &= 1 + \dots \\ \frac{g_0}{m^{4-d}} &= g + \frac{3}{2} g^2 \int \frac{d^d q}{(2\pi)^d} \frac{1}{(q^2 + 1)^2} + \dots = g + \frac{3}{2} g^2 \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} + \dots \\ Z_2 &= 1 - g \int \frac{d^d q}{(2\pi)^d} \frac{1}{(q^2 + 1)^2} + \dots = 1 - \frac{1}{2} g \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} + \dots \end{aligned}$$

in agreement with the expressions given in (114).

We can confirm the trivial ultraviolet behaviour by studying the variation of g with respect to the scale $m \ll \Lambda$ at fixed g_0 and Λ . Since $dg/\beta(g) = d \ln m$, one has

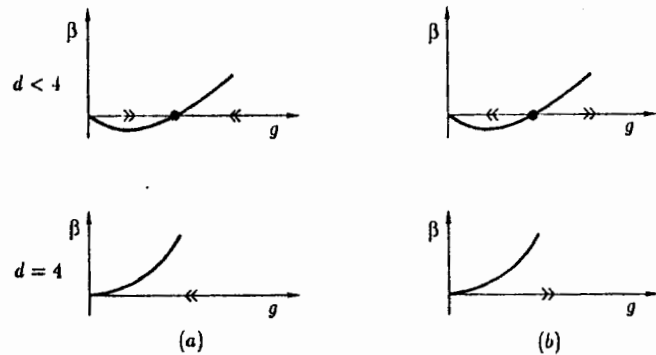
$$\frac{m_2}{m_1} = \exp \int_{g_1}^{g_2} \frac{dg'}{\beta(g')}$$

As β is negative, the effective coupling constant $g_2 \equiv g(m_2)$ vanishes exponentially in the region $\Lambda \gg m_2 \gg m_1$

$$\frac{g_2}{g_1} \sim \left(\frac{m_1}{m_2} \right)^{4-d} \quad \frac{m_2}{m_1} \gg 1 \quad (138)$$

The situation seems to be paradoxically very different in four dimensions, where $\beta(g)$ is positive and of order g^2 ($\beta(g) \sim 3g^2/(4\pi)^2 + \dots$ for $n = 1$). Apparently, the origin is a repulsive ultraviolet fixed point. As m increases, g also increases and one enters an unknown nonperturbative domain. It is conceivable that $\beta(g)$ might have a maximum, then vanish again at a nontrivial ultraviolet fixed point. This unlikely possibility seems excluded by numerical simulations, and thus we assume that it does not occur. Note that the region $g > 0$ for $d = 4$ corresponds to the domain $g > g_c$ for $d < 4$. We just saw that, at fixed m and Λ and for $d < 4$, the bare coupling constant covers the whole interval $[0, \infty]$ when the renormalized coupling varies in $[0, g_c]$. Therefore, the region $g > g_c$ is not accessible without an analytical continuation which reveals pathologies in the correlation functions, as e.g. imaginary parts. This is physically unacceptable and signals instabilities. One is tempted to make the same analysis in dimension four for the whole domain $g > 0$ (figure 6).

Formally, the terms of the perturbative renormalized expansion are nevertheless well-defined for $d = 4$. The question is to know whether these series are asymptotic to meaningful correlation

Fig. 6 Infrared (a) and ultraviolet (b) behaviours for $d < 4$ and $d = 4$.

functions. From the preceding discussion, this does not seem to be the case, but, up to now, the justification seems to be circular, since it requires that regular functions such as $\beta(g)$,... are evaluated perturbatively. The computation of such quantities within an ill-defined model seems paradoxical. However, in the framework of dimensional continuation, it is conceivable to consider these quantities as regular limits of well-defined functions for $d < 4$.

As $m \rightarrow \infty$, (after taking first the limit $\Lambda \rightarrow \infty$) the theory becomes trivial. This is not the case when m decreases. In this case, g tends towards the nontrivial attractive infrared fixed point. Let us study the expressions resulting from the renormalization group equations.

5.3.4 Correlation functions at the critical point

We fix the dimension to some value less than four. Standard dimensional analysis yields

$$\Gamma_n^R(\lambda p; \lambda m, g) = \lambda^{d - \frac{1}{2}n(d-2)} \Gamma_n^R(p; m, g) \quad (139a)$$

and allows derivatives in m at fixed p to be changed into derivatives in p at fixed m . Using the notation

$$d_n^R = d - \frac{1}{2}n(d-2) \quad (139b)$$

we can rewrite the Callan-Symanzik equations (109) as

$$\left\{ -\lambda \frac{\partial}{\partial \lambda} + \beta(g) \frac{\partial}{\partial g} + (d_n^R - \frac{1}{2}n\gamma_1(g)) \right\} \Gamma_n^R(\lambda p; m, g) = (2 - \gamma_1(g)) m^2 \Gamma_{n,1}^R(\lambda p, 0; m, g) \quad (140)$$

which makes obvious the interpretation of $\frac{1}{2}\gamma_1(g)$ as the abnormal dimension of the field φ

$$d_n^R - \frac{1}{2}n\gamma_1(g) = d - n \left(\frac{1}{2}(d-2) + \frac{1}{2}\gamma_1(g) \right) \quad (141)$$

Renormalized connected Green functions, here written without the index c for simplicity, obey similar equations. Setting

$$d_n^G = d - \frac{1}{2}n(d+2) \quad (142)$$

they read

$$\left\{ -\lambda \frac{\partial}{\partial \lambda} + \beta(g) \frac{\partial}{\partial g} + (d_n^G + \frac{1}{2}n\gamma_1(g)) \right\} G_n^R(\lambda p; m, g) = (\gamma_1(g) - 2) m^2 G_{n,1}^R(\lambda p, 0; m, g) \quad (143)$$

Prove equation (143).

Let us introduce a coupling constant $\bar{g}(\lambda, g)$ depending on the scale λ , as the solution of the first order differential equation

$$\lambda \frac{d\bar{g}(\lambda, g)}{d\lambda} = \beta[\bar{g}(\lambda, g)] \quad \bar{g}(1, g) = g \quad (144)$$

This function is given implicitly by

$$\ln \lambda = \int_g^{\bar{g}(\lambda, g)} \frac{dg'}{\beta(g')} \quad (145)$$

Recall also that, according to equations (123),

$$Z_1(g) = \exp \int_0^g dg' \frac{\gamma_1(g')}{\beta(g')} \quad (146)$$

Finally, for short, we introduce the notation

$$\bar{G}_n^R(p; m, g) = (2 - \gamma_1(g)) m^2 G_{n,1}^R(p, 0; m, g) \quad (147)$$

A solution of (143) can thus be written as

$$\begin{aligned} \lambda_2^{-d_n^G} Z_1^{n/2}(g) G_n^R(\lambda_2 \mathbf{p}; m, g) \\ - \lambda_1^{-d_n^G} Z_1^{n/2}(\bar{g}[\frac{\lambda_2}{\lambda_1}, g]) G_n^R(\lambda_1 \mathbf{p}; m, \bar{g}[\frac{\lambda_2}{\lambda_1}, g]) = \\ = \int_{\lambda_1}^{\lambda_2} \frac{d\lambda}{\lambda} \lambda^{-d_n^G} Z_1^{n/2}(\bar{g}[\frac{\lambda_2}{\lambda}, g]) \tilde{G}_n^R(\lambda \mathbf{p}; m, \bar{g}[\frac{\lambda_2}{\lambda}, g]) \end{aligned} \quad (148)$$

Using the function $\bar{g}(\lambda, g)$, show that the equations (143) transform into first order differential equations and yield (148).

In order to clarify the interpretation of equation (148), let us assume the r.h.s. negligible. Hence, with $\lambda_1 = 1$ and $\lambda_2 = \lambda$,

$$G_n^R(\lambda \mathbf{p}; m, g) \simeq \lambda^{d_n^G} \left(\frac{Z_1(\bar{g}[\lambda, g])}{Z_1(g)} \right)^{n/2} G_n^R(\mathbf{p}; m, \bar{g}[\lambda, g]) \quad (149)$$

This states that changing the scale of the momenta is equivalent to the multiplication by a factor depending on the dilatation parameter accompanied by a change in the coupling constant. At the fixed point of the transformation, the Green functions are homogeneous. Of course, this result presupposes that the r.h.s. is absent. Note that the structure of (149) is reminiscent of the relation between bare and renormalized functions.

Returning to equation (148), let us consider the ultraviolet limit where λ_1 tends to zero, $\bar{g}(\lambda_2/\lambda_1, g)$ goes to zero and $Z_1(\bar{g}(\lambda_2/\lambda_1, g))$ to 1. The Green function is finite at zero momenta since the mass does not vanish. Finally, for $n \geq 1$,

$$\lambda_1^{-d_n^G} = \lambda_1^{\frac{1}{2}n(d+2)-d}$$

also vanishes (this is the reason why we use correlation functions instead of vertex functions, which would require additional subtractions for $n = 2, 4$). Hence,

$$G_n^R(\lambda \mathbf{p}; m, g) = \lambda^{d_n^G} \int_0^\lambda \frac{d\mu}{\mu} \mu^{-d_n^G} \left(\frac{Z_1(\bar{g}[\frac{\lambda}{\mu}, g])}{Z_1(g)} \right)^{n/2} \tilde{G}_n^R(\mu \mathbf{p}; m, \bar{g}[\frac{\lambda}{\mu}, g]) \quad (150)$$

Let us now study the critical limit $m_0^2 \rightarrow m_{0c}^2$, equivalent to $m^2 \rightarrow 0$ at fixed g_0 , or also to $g \rightarrow g_c$ at fixed momenta. Moreover, this is also equivalent to letting m go to zero or the momenta to infinity, since $G_n(\lambda \mathbf{p}; m, g) = m^{d_n^G} G_n(\lambda \mathbf{p}/m, 1, g)$. We want to control the behaviour of the functions $\tilde{G}$ defined in (147) in the large momentum region. Up to a finite renormalization, the perturbative expansion of these functions looks similar to that for correlation functions, except that in turn each propagator is squared. Therefore, the high momentum behaviour is improved order by order. Let us assume that this perturbative property remains valid after resummation, and, more precisely, that

$$\lim_{\lambda \rightarrow \infty} \tilde{G}_n^R(\lambda \mathbf{p}) / G_n^R(\lambda \mathbf{p}) = 0 \quad (151)$$

Due to equation (112), one has

$$\lim_{\lambda \rightarrow \infty} \lambda^{4-d} \bar{g}[\lambda, g] = g_0 / m^{4-d} \quad (152)$$

Combining (151) and (152), one checks that, at fixed g such that $0 \leq g \leq g_c$, the limit $\lambda \rightarrow \infty$ of the integral in (150) reproduces the perturbative expansion. This was expected because the ultraviolet behaviour of the model is trivial.

This is no longer the case for $g = g_c$. In this limit,

$$\begin{aligned} \bar{g}[x, g_c] = g_c \\ \lim_{g \rightarrow g_c} \frac{Z_1(\bar{g}[x, g])}{Z_1(g)} = \lim_{g \rightarrow g_c} \exp \int_g^x dg' \frac{\gamma_1(g')}{\beta(g')} = \exp \gamma_1(g_c) \ln x \end{aligned} \quad (153)$$

Hence,

$$G_n^R(\lambda \mathbf{p}; m, g_c) = \lambda^{d_n^G + \frac{1}{2}n\gamma_1(g_c)} \int_0^\lambda \frac{d\mu}{\mu^{1+d_n^G + \frac{1}{2}n\gamma_1(g_c)}} \tilde{G}_n^R(\mu \mathbf{p}; m, g_c) \quad (154)$$

As $\lambda \rightarrow \infty$, the hypothesis (151) implies that the integral is convergent, so that the behaviour of the correlation functions at g_c in the region $|\mathbf{p}| \gg m$ is a pure power law

$$\lim_{\lambda \rightarrow \infty} G_n^R(\lambda \mathbf{p}; m, g_c) = \lambda^{d_n^G + \frac{1}{2}n\gamma_1(g_c)} \int_0^\infty \frac{d\mu}{\mu^{1+d_n^G + \frac{1}{2}n\gamma_1(g_c)}} \tilde{G}_n^R(\mu \mathbf{p}; m, g_c) \quad (155)$$

This discontinuity in the ultraviolet behaviour arises from the fact that g_c is a repulsive ultraviolet fixed point. The scaling behaviour (155) corresponds of course to the massless theory. The underlying

regularity hypothesis is that the integral is convergent at the upper limit, i.e.

$$\lim_{\lambda \rightarrow \infty} \lambda^{-d_n^G - \frac{1}{2}n\gamma_1(g_c)} \tilde{G}_n^R(\lambda p, m, g_c) = 0$$

This implies that at criticality the scaling behaviour of correlation functions can be deduced consistently from the low frequency modes of the theory. This is an aspect of universality, i.e. the independence on the details of the bare theory at short distance.

Deduce equation (155) directly from the differential equation (143), using the fact that $\beta(g_c) = 0$.

These results can now be inserted in the bare correlation functions expressed in terms of the renormalized mass m instead of the bare one m_0

$$G_n(p; \frac{m}{\lambda}, g_0)_\Lambda = \int_0^\Lambda \frac{d\mu}{\mu^{1+d_n^G}} (Z_1(\bar{g}[\mu^{-1}, g]))^{n/2} \tilde{G}_n^R(\mu p; m, \bar{g}[\mu^{-1}, g]) \quad (156)$$

where $g = g(m, g_0, \Lambda)$, $g_\lambda = g(m/\lambda, g_0, \Lambda)$, and where we used the relation $\bar{g}(\lambda/\mu, g_\lambda) = \bar{g}(1/\mu, g)$.

When $\lambda \rightarrow \infty$, $m/\lambda \rightarrow 0$, the integral may be divergent in the region of large μ . If one uses a perturbative reasoning, one encounters the difficulties discussed at the beginning of this section. Indeed, $\bar{g}$ and Z_1 are polynomials in μ^{4-d} and this generates divergences at a fixed finite order when $\mu \rightarrow \infty$. The renormalization group technique avoids this difficulty. Indeed, $\bar{g}(\mu^{-1}, g)$ tends toward g_c and one obtains finite expressions for the massless theory

$$G_n(\lambda p; 0, g_0)_\Lambda = \lambda^{d_n^G} \int_0^\infty \frac{d\mu}{\mu^{1+d_n^G}} \left(Z_1(\bar{g}[\frac{\lambda}{\mu}, g]) \right)^{n/2} \tilde{G}_n^R(\mu p; m, \bar{g}[\frac{\lambda}{\mu}, g]) \quad (157)$$

This expression illustrates the difference between the large- and short-distance behaviour. If λ becomes very large (but always such that $|\lambda p| \ll \Lambda$), one recovers the perturbative expansion in g_0 , with singularities arising at the same orders as those of the infrared perturbative divergences. Conversely, as $\lambda \rightarrow 0$, $\bar{g}(\lambda/\mu, g)$ tends towards g_c for any finite value of μ , and, from (157),

$$\lim_{\lambda \rightarrow 0} G_n(\lambda p; 0, g_0)_\Lambda$$

$$= \lambda^{d_n^G + \frac{1}{2}n\gamma_1(g_c)} \left(Z_1(g) \exp \int_g^{g_c} dg' \frac{\gamma_1(g') - \gamma_1(g_c)}{\beta(g')} \right)^{n/2} \times \int_0^\infty \frac{d\mu}{\mu^{1+d_n^G + \frac{1}{2}n\gamma_1(g_c)}} \tilde{G}_n^R(\mu p; m, g_c) \quad (158)$$

As a consequence, up to a renormalization factor, the infrared behaviour of the massless bare theory (at T_c) is identical to the ultraviolet behaviour of the massive renormalized theory at $g = g_c$.

With $n = 2$ and $d_2^G = -2$ in (158), one obtains the small momentum behaviour of the 2-point correlation function

$$G_2(p) \sim \frac{1}{p^{2-\gamma_1(g_c)}} \quad (159)$$

Equivalently, the large-distance behaviour in configuration space reads $1/|x|^{d-2+\gamma_1(g_c)}$. We recover the result (132) of the preceding subsection, $\eta = \gamma_1(g_c)$.

(1) Corrections to the dominant scaling behaviour. Check, using an expansion of (157) beyond the dominant term, that the correction of order $g_\lambda - g_c$ for $\lambda \rightarrow 0$ is (cf. (133))

$$G_n(\lambda p; 0, g_0)_\Lambda \underset{\lambda \rightarrow 0}{\sim} \text{Cst} \lambda^{d_n^G + \frac{1}{2}n\gamma_1(g_c)} (1 + \mathcal{O}(\lambda^\omega)) \quad (160)$$

where $\omega = \beta'(g_c)$ is the invariant slope of the β function at the critical point. We shall return later to the corrections to the scaling laws due to irrelevant terms in the Lagrangian.

(2) Nonanalytic behaviour of the critical theory with respect to the bare coupling constant. In dimension less than four, the mass introduced in the renormalized theory plays the role of a regulator of infrared perturbative divergences and of a scaling factor independent from the cutoff $\Lambda \gg m$. Using the homogeneity of $\tilde{G}_n^R$, equation (158) is rewritten as

$$\lim_{\lambda \rightarrow 0} G_n(\lambda p; 0, g_0)_\Lambda = \lambda^{d_n^G + \frac{1}{2}n\gamma_1(g_c)} m^{-\frac{1}{2}n\gamma_1(g_c)} \times \left\{ \left(\frac{g_c}{g} \right)^{\frac{\gamma_1(g_c)}{4-d}} e^{\gamma_1(g_c) \int_0^g dg' \left(\frac{1}{\beta(g')} + \frac{1}{(4-d)g'} \right)} e^{\int_0^{g_c} dg' \left(\frac{\gamma_1(g') - \gamma_1(g_c)}{\beta(g')} - \frac{\gamma_1(g_c)}{(4-d)g'} \right)} \right\}^{n/2} \times \int_0^\infty \frac{d\mu}{\mu^{1+d_n^G + \frac{1}{2}n\gamma_1(g_c)}} \tilde{G}_n^R(\mu p; m = 1, g_c) \quad (161)$$

The dependence on g (function of g_0) is only contained within the braces. The last integral is a constant, the second term is a regular

function as $g \rightarrow 0$. There just remains

$$\left[\left(\frac{g_c}{g} \right)^{\gamma_1(g_c)/(4-d)} m^{-\gamma_1(g_c)} \right]^{n/2} \sim \left(\frac{g_c}{g_0} \right)^{\frac{1}{2} n \gamma_1(g_c)/(4-d)}$$

which exhibits a non analytic behaviour in g_0 . This is of course once more related to the infrared singularities in the massless perturbative theory.

5.3.5 Expansion near the critical point

Critical scaling laws require the introduction of an arbitrary scale, essential to restore the standard dimensional analysis. We used previously the correlation length near T_c . However, this is not very practical when one explores the neighborhood of T_c , and in particular the ordered region $\theta < 0$. The procedure would be simplified by starting directly from the massless theory. This is of course impossible for $d < 4$ due to infrared divergences, but can be implemented in a double expansion in g and $\varepsilon = 4 - d$ following the original suggestion of Wilson and Fisher. In order to carry out this program, we introduce an arbitrary scale μ in the normalization conditions, according to

$$\begin{aligned} \Gamma_2^R(p^2; g, \mu) \Big|_{p^2=0} &= 0 \\ \frac{\partial \Gamma_2^R}{\partial p^2}(p^2; g, \mu) \Big|_{p^2=\mu^2} &= 1 \\ \Gamma_4^R(p_s; g, \mu) &= \mu^{4-d} g \end{aligned} \quad (162)$$

These conditions now replace equations (88). The symmetric point p_s satisfies

$$p_i \cdot p_j = \frac{1}{3} \mu^2 (4\delta_{ij} - 1) \quad (163)$$

The choice of nonvanishing momenta in the normalization of the four-point function and the derivative of the two-point function is such as to avoid the occurrence of infrared divergences. As we shall see later, when the massless theory is well under control, the mass can be introduced semiperturbatively, so that the small θ domain can be explored. The method is based on the fact that the functions $\beta(g, \varepsilon)$ and $\gamma_i(g, \varepsilon)$ which determine the critical behaviour can be entirely computed within the massless theory.

The renormalization group equations express now the invariance with respect to variations of the normalization scale μ , according to the original point of view of Gell-Mann and Low. Indeed the regularized bare theory does not refer to μ . Hence, differentiating the relation $\Gamma_n(p; g_0, m_0^c)|_\Lambda = Z_1^{-\frac{1}{2}n} \Gamma_n^R(p; g, \mu)$ with respect to μ , one obtains the homogeneous equation

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) \right\} \Gamma_n^R(p; g, \mu) = 0 \quad (164)$$

with

$$\begin{aligned} \beta(g) &= \mu \frac{\partial g}{\partial \mu} \Big|_{g_0, \Lambda} \\ \gamma_1(g) &= Z_1^{-1} \mu \frac{\partial Z_1}{\partial \mu} \Big|_{g_0, \Lambda} \end{aligned} \quad (165)$$

Although these quantities might differ from those defined in (107), we use the same notation. The analysis of equation (164) leads however to similar conclusions. There is a infrared stable fixed point of order $g_c \sim \varepsilon$ and the critical exponent η is equal to $\gamma_1(g_c)$. Equations (164) can also be generalized to functions with $\frac{1}{2}\varphi^2$ insertions, which are normalized according to

$$\begin{aligned} \Gamma_{n,s}(p, q; g_0, \Lambda) &= Z_1^{-n/2} Z_2^s \left[\Gamma_{n,s}^R(p, q; g, \mu) + A \delta_{n,0} \delta_{s,2} \right] \\ \Gamma_{2,1}^R(p_1, p_2, q = -p_1 - p_2; g, \mu) \Big|_{p_1^2=p_2^2=-3p_1 \cdot p_2=\mu^2} &= 1 \\ \Gamma_{0,2}^R(q; g, \mu) \Big|_{q^2=\frac{4}{3}\mu^2} &= 0 \end{aligned} \quad (166)$$

and obey

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) + s \gamma_2(g) \right\} \Gamma_{n,s}^R = \delta_{n,0} \delta_{s,2} B(g) \quad (167)$$

with

$$\gamma_2(g) = Z_2^{-1} \mu \frac{\partial Z_2}{\partial \mu} \Big|_{g_0, \Lambda} \quad (168)$$

and

$$B(g) = \left\{ \mu \frac{\partial}{\partial \mu} + 2\gamma_1(g) \right\} A \quad (169)$$

This last quantity has dimension $d - 4$. Excepting the case $n = 0$, $s = 2$ (which corresponds to the specific heat), the standard

dimensional analysis, which yields

$$\Gamma_{n,s}^R(\lambda p, \lambda q; g, \lambda \mu) = \lambda^{d-\frac{1}{2}n(d-2)-2s} \Gamma_{n,s}^R(p, q; g, \mu) \quad (170)$$

and the relation (167) taken at $g = g_c$ where $\beta(g_c) = 0$, lead to the small momentum behaviour

$$\Gamma_{n,s}^R(\lambda p, \lambda q; g, \mu) \underset{\lambda \rightarrow 0}{\sim} \lambda^{d-\frac{1}{2}n(d-2+\eta)-s/\nu} \quad (171)$$

Hence, according to (130) and (132),

$$\begin{aligned} \gamma_1(g_c) &= \eta \\ \gamma_2(g_c) &= 2 - \nu^{-1} \end{aligned} \quad (172)$$

For $n = 0, s = 2$, the inhomogeneous term yields an additional constant which may be dominant

$$\Gamma_{0,2}^R(\lambda q; g, \mu) \underset{\lambda \rightarrow 0}{\sim} a_1 + a_2 \lambda^{d-2/\nu} \quad (173)$$

The massless symmetric theory is thus constructed close to four dimensions in a $4-d$ expansion. The neighborhood of T_c will then be studied by resummation of a series in the quantity $m_0^2 - m_{0c}^2$ proportional to the deviation θ from criticality, with θ coupled to the operator $\frac{1}{2}\varphi^2$. For the bare functions, one has an expansion

$$\Gamma_n(p, m_0^2 - m_{0c}^2) = \sum_{s=0}^{\infty} \frac{(m_0^2 - m_{0c}^2)^s}{s!} \Gamma_{n,s}^R(p, q=0; m_{0c}^2)$$

Due to the relation

$$(m_0^2 - m_{0c}^2) \frac{1}{2} \varphi^2 = \theta \frac{Z_1}{Z_2} \left(\frac{1}{2} \varphi^2 \right)_R \quad (174)$$

which can be interpreted as $Z_2(m_0^2 - m_{0c}^2) = \theta$, the renormalized quantities satisfy

$$\Gamma_n^R(p; \theta) = \sum_{s=0}^{\infty} \frac{\theta^s}{s!} \Gamma_{n,s}^R(p, q=0, \theta=0) \quad (175)$$

Because of infrared divergences, this expansion is, however, only valid if we use a space dependent quantity $\theta(x)$, constant at finite distance, and vanishing sufficiently fast at infinity. It should more properly be written as

$$\Gamma_n^R(p; \theta(x)) = \sum_{s=0}^{\infty} \frac{1}{s!} \int \prod_{k=1}^s \left(\frac{d^d q_k}{(2\pi)^d} \tilde{\theta}(q_k) \right) \Gamma_{n,s}^R(p, \{q_k\}; \theta=0) \quad (176)$$

From (167), one gets

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) + \gamma_2(g) \int \frac{d^d q}{(2\pi)^d} \tilde{\theta}(q) \frac{\delta}{\delta \tilde{\theta}(q)} \right\} \times \Gamma_n^R(p; \theta(x), g, \mu) = 0$$

In the limit of uniform θ , which corresponds to the introduction of a massive propagator, one finally obtains

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) + \gamma_2(g) \theta \frac{\partial}{\partial \theta} \right\} \Gamma_n^R(p; \theta, g, \mu) = 0 \quad (177)$$

These equations are to be compared to relations (110). They differ only by finite renormalizations, but the crucial point is that they are homogeneous. The derivation shows that the coefficients $\beta, \gamma_1, \gamma_2$ for the massive theory are the same as those computed in the simpler massless theory through a double expansion in g and ϵ .

From equations (177), we can now extract all scaling relations. Let us consider first the symmetric phase $\theta > 0$ where the Green functions with an odd number of arguments identically vanish. With the notation

$$\begin{aligned} \mu(\lambda) &= \lambda \mu \\ \lambda \frac{dg(\lambda)}{d\lambda} &= \beta(g(\lambda)) & g(1) &= g \\ \lambda \frac{d \ln \theta(\lambda)}{d\lambda} &= \gamma_2(g(\lambda)) & \theta(1) &= \theta \end{aligned} \quad (178)$$

the homogeneous equations (177) are solved in the form

$$\Gamma_n^R(p; \theta, g, \mu) = \exp \left\{ -\frac{1}{2} n \int_1^\lambda \frac{d\lambda'}{\lambda'} \gamma_1(g(\lambda')) \right\} \Gamma_n^R(p; \theta(\lambda), g(\lambda), \lambda \mu) \quad (179)$$

Let us choose λ_0 such that $\theta(\lambda_0)$ has a fixed value of the form

$$\theta(\lambda_0) = \lambda_0^2 \mu^2 \quad (180)$$

which defines λ_0 implicitly through

$$\ln(\theta/\mu^2) = \int_g^{g(\lambda_0)} dg' (2 - \gamma_2(g')) / \beta(g')$$

With this choice, equation (179) can be rewritten as

$$\Gamma_n^R(p; \theta, g, \mu) = Z_1^{n/2}(m/\mu, g(\lambda_0)) \Gamma_n^R(p; m, g(\lambda_0)) \quad (181)$$

where we use the notations

$$m = \mu \exp \int_g^{g(\lambda_0)} \frac{dg'}{\beta(g')} \quad (182a)$$

$$Z_1(m/\mu, g(\lambda_0)) = \exp - \int_g^{g(\lambda_0)} dg' \gamma_1(g')/\beta(g') \quad (182b)$$

Again these quantities differ from the previous ones at most by finite renormalizations. Let us now choose the normalization scale μ in the ultraviolet domain, i.e. such that

$$\theta^{1/2}, |p| \ll \mu \ll \Lambda \quad (183)$$

In this way, μ appears as an additional arbitrary cutoff factor, but, in contradistinction with the bare theory in which we have to neglect terms of order $1/\Lambda$, this new formulation is now minimal and exact. As $\theta(\lambda_0)$ is fixed whereas μ tends to infinity, λ_0 tends to zero, and $g(\lambda_0)$ approaches the infrared attractive fixed point g_c . According to (172) and (182),

$$\begin{aligned} \frac{\theta}{\mu^2} &\rightarrow \left(\frac{m}{\mu}\right)^{1/\nu} \\ Z_1 &\rightarrow \left(\frac{m}{\mu}\right)^{-\eta} \simeq \left(\frac{\theta}{\mu^2}\right)^{-\eta\nu} \end{aligned} \quad (184)$$

In the expressions for the vertex functions at finite p/m , $g(\lambda_0)$ can be replaced in the r.h.s. of (181) by g_c , and, using the dimension

$$d_n^r = d - \frac{1}{2}n(d-2)$$

one finally gets

$$\Gamma_n^R(p; \theta, g, \mu) = \mu^{d-\frac{1}{2}n(d-2)} \left(\frac{\theta}{\mu^2}\right)^{\nu[d-\frac{1}{2}n(d-2+\eta)]} f_n \left(\frac{p}{\mu} \left(\frac{\mu^2}{\theta}\right)^\nu\right) \quad (185)$$

which is the scale invariant form in the critical domain. The correlation length ξ proportional to $\theta^{-\nu}$ appears naturally in the reduced function f_n . For $\theta > 0$, Γ_n^R has a finite limit at zero momenta behaving as $\theta^{\nu[d-\frac{1}{2}n(d-2+\eta)]}$. In particular, the inverse susceptibility corresponds to the case $n=2$ and hence the critical exponent γ satisfies

$$\boxed{\gamma = \nu(2 - \eta)} \quad (186)$$

The Γ_n^R 's have been constructed in order to be finite at $\theta = 0$. The prefactor with a fractional power of θ in (185) should thus be compensated by the behaviour of the reduced functions f_n when their arguments tend to infinity ($\theta^\nu \ll |p| \ll \mu$); hence

$$\theta^\nu \ll \lambda |p| \ll \mu \quad \Gamma_n^R(\lambda p; \theta, g, \mu) \sim \lambda^{d-\frac{1}{2}n(d-2+\eta)} \quad (187)$$

which relates the ultraviolet behaviour of the massive theory to the scaling behaviour of the massless one.

Similarly, for the functions with $\frac{1}{2}\varphi^2$ insertions, one finds (except in the cases $n=0, s=1, 2$)

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} n \gamma_1(g) + \gamma_2(g) (s + \theta \frac{\partial}{\partial \theta}) \right\} \Gamma_{n,s}^R(p, q; \theta, g, \mu) = 0 \quad (188)$$

which leads to

$$\begin{aligned} \Gamma_{n,s}^R(p, q; \theta, g, \mu) &= \mu^{d-\frac{1}{2}n(d-2)-2s} \left(\frac{\theta}{\mu^2}\right)^{-s+\nu[d-\frac{1}{2}n(d-2+\eta)]} \\ &\quad \times f_{n,s} \left(\frac{p}{\mu} \left(\frac{\mu^2}{\theta}\right)^\nu, \frac{q}{\mu} \left(\frac{\mu^2}{\theta}\right)^\nu\right) \end{aligned} \quad (189)$$

in the region $\theta \ll \mu^2$, $|p|, |q| \ll \mu$. Note that the prefactor can also be interpreted as

$$m^{-2s} \left(\frac{m}{\mu}\right)^{s\gamma_2(g_c)} \sim \mu^{-2s} \left(\frac{\mu^2}{\theta}\right)^s$$

For insertions at zero momentum $q=0$, there is an additional divergent factor $1/\theta$ which was to be expected since this insertion is equivalent to a derivative with respect to θ . As $\theta \rightarrow 0$, the same mechanism of compensations yields the scaling behaviour (171).

An additional term should be introduced in the r.h.s. of equation (188) for the function $\Gamma_{0,2}$; writing it as $\mu^{d-4}b(g)$, one obtains

$$\Gamma_{0,2}^R(q; \theta, g, \mu) \sim \mu^{d-4} \left\{ \left(\frac{\theta}{\mu^2}\right)^{-(2-\nu d)} f_{0,2} \left(\frac{q}{\mu} \left(\frac{\mu^2}{\theta}\right)^\nu\right) - \frac{\nu}{2-\nu d} b(g_c) \right\} \quad (190)$$

for $\theta \ll \mu^2$, $|q| \ll \mu$. As $q \rightarrow 0$, we obtain the specific heat singularity in $\theta^{-\alpha}$ with

$$\boxed{\alpha = 2 - \nu d} \quad (191)$$

The additive constant in (190) is the regular part of the specific heat, and ensures the coherence of the results when α is negative. Finally, normalizing the free energy to zero at $\theta = 0$ in the absence of field, and noting that $\Gamma_{0,2}^R(\theta) = d^2\Gamma^R(\theta)/d\theta^2$, one obtains its expression in the critical region for $\theta > 0$ as

$$\Gamma^R(\theta) \sim \mu^d \left[f_{0,2}(0) \frac{1}{(2-\alpha)(1-\alpha)} \left(\frac{\theta}{\mu^2} \right)^{2-\alpha} - \frac{\nu}{2\alpha} b(g_c) \left(\frac{\theta}{\mu^2} \right)^2 + C \frac{\theta}{\mu^2} \right] \quad (192)$$

The constant C is related to $\Gamma_{0,1}^R$.

5.3.6 Scaling laws below the critical temperature

If the argument φ is given a uniform value M , the functional $\Gamma(\varphi)$ becomes proportional to the volume, with a coefficient given by the effective potential $V(M)$

$$\begin{aligned} \varphi(\mathbf{x}) &\rightarrow M \\ \Gamma^R(\varphi) &\rightarrow \int d^d \mathbf{x} V(M) \\ V(M, \theta, g, \mu) &= \sum_{n=0}^{\infty} \frac{M^n}{n!} \Gamma_n^R(\mathbf{p} = 0; \theta, g, \mu) \end{aligned} \quad (193)$$

The corresponding uniform external field, denoted H , equals

$$H = \frac{\partial V}{\partial M} = \sum_{n=1}^{\infty} \frac{M^{n-1}}{n!} \Gamma_{n+1}^R(\mathbf{p} = 0; \theta, g, \mu) \quad (194)$$

The generalization to the case of multicomponent fields does not present difficulties. We restrict ourselves for the time being, to the scalar case. Let us establish the form of the equation of state in the scaling region. We do not assume a given sign for θ . The name equation of state refers to the interpretation of the model in terms of a liquid-vapor equilibrium near the critical point. Dimensional analysis yields

$$[V] = d \quad [M] = \frac{1}{2}(d-2) \quad [H] = \frac{1}{2}(d+2) = 3 - \frac{1}{2}(4-d) \quad (195)$$

Repeating the derivation of the renormalization flow equations, we obtain

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \gamma_1(g) \left[1 + M \frac{\partial}{\partial M} \right] + \gamma_2(g) \theta \frac{\partial}{\partial \theta} \right\} \times H(M, \theta, g, \mu) = 0 \quad (196)$$

Using the notation (178) and the dimensions (195), this yields

$$\begin{aligned} H(M, \theta, g, \mu) &= \mu^{\frac{1}{2}(d+2)} H\left(\frac{M}{\mu^{\frac{1}{2}(d-2)}}, \frac{\theta}{\mu^2}, g, 1\right) \\ &= (\lambda \mu)^{\frac{1}{2}(d+2)} \exp \left\{ -\frac{1}{2} \int_g^{g(\lambda)} dg' \frac{\gamma_1(g')}{\beta(g')} \right\} \\ &\quad \times H\left(\frac{M(\lambda)}{(\lambda \mu)^{\frac{1}{2}(d-2)}}, \frac{\theta(\lambda)}{\lambda^2 \mu^2}, g(\lambda), 1\right) \end{aligned} \quad (197)$$

with $M(\lambda)$ given by

$$\ln \frac{M(\lambda)}{M} = -\frac{1}{2} \int_g^{g(\lambda)} dg' \frac{\gamma_1(g')}{\beta(g')} \quad (198)$$

In order to eliminate $M(\lambda)$, we choose λ in such a way that

$$\frac{M(\lambda)}{(\lambda \mu)^{\frac{1}{2}(d-2)}} = 1 \quad (199)$$

In other words, λ is determined by the condition

$$\ln \frac{M}{\mu^{\frac{1}{2}(d-2)}} = \frac{1}{2} \int_g^{g(\lambda)} dg' \frac{d-2 + \gamma_1(g')}{\beta(g')} \quad (200)$$

In the critical domain, $M/\mu^{\frac{1}{2}(d-2)} \ll 1$. Hence (200) is satisfied if $g(\lambda) \rightarrow g_c$ and if λ tends to zero according to

$$\lambda \sim \left(\frac{M}{\mu^{\frac{1}{2}(d-2)}} \right)^{2/(d-2+\eta)} = \left(\frac{M}{\mu^{\frac{1}{2}(d-2)}} \right)^{\nu/\beta} \quad (201)$$

where the critical exponent β is defined by

$$\boxed{\beta = \frac{1}{2}\nu(d-2+\eta)} \quad (202)$$

Hence, as $g(\lambda) \rightarrow g_c$,

$$\frac{\theta(\lambda)}{\lambda^2 \mu^2} \sim \frac{\theta}{\mu^2} \lambda^{\gamma_2(g_c)-2} = \frac{\theta}{\mu^2} \lambda^{-1/\nu} = \frac{\theta}{\mu^2} \left(\frac{\mu^{\frac{1}{2}(d-2)}}{M} \right)^{2/(\nu(d-2+\eta))} \quad (203)$$

Inserting relation (199) into (197), one gets

$$H(M, \theta, g, \mu) \sim \mu^{\frac{1}{2}(d+2)} \left(\frac{M}{\mu^{\frac{1}{2}(d-2)}} \right)^{(d+2-\eta)/(d-2+\eta)} f \left[\frac{\theta}{\mu^2} \left(\frac{\mu^{\frac{1}{2}(d-2)}}{M} \right)^{2/(\nu(d-2+\eta))} \right] \quad (204)$$

Choosing now for definiteness $\mu = 1$,

$$\frac{H}{M^\delta} = f\left(\frac{\theta}{M^{1/\beta}}\right) \quad (205)$$

where

$$\delta = \frac{d+2-\eta}{d-2+\eta} \quad (206)$$

We have therefore recovered all the relations between the exponents which were discussed in the previous chapter.

We know that H is analytic and odd with respect to M when θ is positive. This forces the following large x behaviour of $f(x)$ in equation (205)

$$f(x) \underset{x \rightarrow \infty}{\sim} \sum_{p=0}^{\infty} f_p x^{\beta\delta - \beta(2p+1)} = \sum_{p=0}^{\infty} f_p x^{\gamma - 2p\beta} \quad (207)$$

since $\beta(\delta - 1) = \nu(2 - \eta) = \gamma$. For small θ and finite M , H remains analytic in θ as long as M does not vanish. Hence $f(x)$ is regular at the origin.

The spontaneous magnetization can be extracted from equation (205) by letting H approach zero, assuming that the equation $f(x) = 0$ has a negative solution $-x_0$ (θ is of course negative). Hence

$$M_{sp} = \mu^{\frac{1}{2}(d-2)} \left(-x_0 \frac{\theta}{\mu^2} \right)^\beta \quad (208)$$

This gives the usual interpretation of the critical exponent β .

In the scalar case considered so far, Goldstone's theorem does not apply. For a given magnetization, the correlation functions in

the low temperature phase

$$\begin{aligned} \Gamma_n^R(p_1, \dots, p_n; \theta, M, g, \mu) &= \\ &= \int \prod_1^n (e^{ip_k \cdot x_k} d^d x_k) \frac{\delta^n}{\delta \varphi(x_1) \dots \delta \varphi(x_n)} \Gamma^R(\varphi; \theta, g, \mu) \Big|_{\varphi(x)=M} \\ &= \sum_{k=0}^{\infty} \frac{M^k}{k!} \Gamma_{n+k}^R(p_1, \dots, p_n, 0, \dots, 0; g, \theta, \mu) \end{aligned} \quad (209)$$

satisfy

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \gamma_1(g) \left[n + M \frac{\partial}{\partial M} \right] + \gamma_2(g) \theta \frac{\partial}{\partial \theta} \right\} \times \Gamma_n^R(p; \theta, M, g, \mu) = 0 \quad (210)$$

Hence

$$\begin{aligned} \Gamma_n^R(p; \theta, M, g, \mu) &= (\lambda \mu)^{d - \frac{1}{2}n(d-2)} \left[\exp -\frac{1}{2}n \int_g^{g(\lambda)} dg' \frac{\gamma_1(g')}{\beta(g')} \right] \\ &\times \Gamma_n^R\left(\frac{p}{\lambda \mu}; \frac{\theta(\lambda)}{\lambda^2 \mu^2}, \frac{M(\lambda)}{(\lambda \mu)^{\frac{1}{2}(d-2)}}, g(\lambda), 1\right) \end{aligned} \quad (211)$$

Let us use the same condition (199) to determine λ , and pick $M \ll \mu^2$ in the critical domain (which forces $\lambda \rightarrow 0$ and $g(\lambda) \rightarrow g_c$). One gets

$$\begin{aligned} \Gamma_n^R(p; \theta, M, g, \mu) &\sim \mu^{d - \frac{1}{2}n(d-2)} \left(\frac{M}{\mu^{\frac{1}{2}(d-2)}} \right)^{[2d/(d-2+\eta)] - n} \\ &\times f_n \left(\frac{p}{\mu} \left(\frac{\mu^{\frac{1}{2}(d-2)}}{M} \right)^{\nu/\beta}, \frac{\theta}{\mu^2} \left(\frac{\mu^{\frac{1}{2}(d-2)}}{M} \right)^{1/\beta} \right) \end{aligned} \quad (212)$$

$|p| \ll \mu \qquad \theta \ll \mu^2 \qquad M \ll \mu^{\frac{1}{2}(d-2)}$

Note that Γ_n^R is regular in M for $\theta > 0$ and in θ for $M \neq 0$.

The vertex functions for $M \neq 0$ are the Legendre transforms of the connected Green functions which, in the present case, contain the subtraction of the 1-point function, i.e. the magnetization M . For $\theta < 0$, $M \rightarrow M_{sp}$ when the external field vanishes. With $\mu = 1$, the functions behave in the critical domain as

$$\Gamma_n^R(p; \theta, M_{sp}) \sim (-\theta)^{\nu[d - \frac{1}{2}n(d-2+\eta)]} f_n((-\theta)^{-\nu} p) \quad (213)$$

This is the same structure as in equation (185) for $\theta > 0$, but n is not constrained to be even. In particular, for $\theta < 0$, the correlation length diverges as

$$\xi \sim (-\theta)^{-\nu'} \quad \text{with } \nu' = \nu \quad (214)$$

The thermal exponent ν takes the same value below as well as above the critical temperature. Similarly, for $n = 2$ and $p \rightarrow 0$, the susceptibility diverges as

$$\chi \sim (-\theta)^{-\gamma'} \quad \text{with } \gamma' = \gamma = \nu(2 - \eta) \quad (215)$$

with the same exponent as above T_c . Generalizing these results to the derivatives with respect to θ , the specific heat divergence is found to be

$$\Gamma_{0,2}^R(\mathbf{q} = 0; \theta, M_{sp}) \sim (-\theta)^{-\alpha'} \quad \text{with } \alpha' = \alpha = 2 - \nu d \quad (216)$$

Again, one finds the same exponent above and below the critical temperature.

In the case of a multicomponent field, one should distinguish transverse and longitudinal components with respect to the external field or the magnetization. If there exists a symmetry group for, e.g. $O(n)$, the functionals $F(\mathbf{H})$ and $\Gamma(\mathbf{M})$ are of course invariant.

Let $H = |\mathbf{H}|$ and $M = |\mathbf{M}|$ denote the magnitudes of a uniform external field and the corresponding magnetization. The effective potential depends only on H and therefore $\mathbf{M}$ is parallel to the applied field

$$H_a = \frac{M_a}{M} \frac{dV}{dM} \quad H = \frac{dV}{dM} \quad (217)$$

The susceptibility and its inverse are now matrices.

$$\begin{aligned} \chi_{ab}^{-1} &= \frac{\partial H_a}{\partial M_b} = \frac{M_a M_b}{M^2} \chi_i^{-1} + \left(\delta_{ab} - \frac{M_a M_b}{M^2} \right) \chi_i^{-1} \\ \chi_i^{-1} &= \frac{d^2 V}{dM^2} \\ \chi_i^{-1} &= \frac{1}{M} \frac{dV}{dM} = \frac{H}{M} \end{aligned} \quad (218)$$

Below the critical point ($\theta < 0$), $M \rightarrow M_{sp}$ and χ_i^{-1} vanishes when $H \rightarrow 0$. This is Goldstone's theorem, asserting that the transverse

susceptibility is infinite in the ordered phase of a system with a continuous symmetry, due to the existence of zero modes. On the other hand, the longitudinal susceptibility has the same behaviour as in one-component systems.

5.4 Corrections to scaling laws

The discussion of correlations is not limited to the dominant scaling behaviour. In particle physics, after an analytic continuation in the Minkovskian region, the Green functions also describe bound states, scattering amplitudes, etc. We do not discuss these problems here. It is nevertheless interesting, in the framework of statistical mechanics, to compute the leading corrections to scaling. We examine successively deviations from the critical point, the role of marginal operators and the contributions of irrelevant terms.

5.4.1 Deviation from the critical point in dimension lower than four

Looking at the renormalization flow equations, such as e.g. (210), in the framework of the φ^4 model, where irrelevant operators are neglected, one sees that the correlation functions are homogeneous at $g = g_c$, whatever the auxiliary normalization parameter μ is. Let us consider a small deviation of g from g_c . The effect can be absorbed in part in a change of the normalization of correlations at g_c , the remaining part representing the essential effect.

In order to perform this splitting, let

$$\begin{aligned} M(g) &= M \exp \frac{1}{2} \int_{g_c}^g dg' \frac{\gamma_1(g') - \gamma_1(g_c)}{\beta(g')} \\ \theta(g) &= \theta \exp - \int_{g_c}^g dg' \frac{\gamma_2(g') - \gamma_2(g_c)}{\beta(g')} \end{aligned} \quad (219)$$

Hence, dropping the index for renormalized functions,

$$\begin{aligned} \Gamma_n(\mathbf{p}; M, \theta, g, \mu) &= \Gamma_n(\mathbf{p}; M(g), \theta(g), g_c, \mu) \\ &\times \exp \left\{ \frac{1}{2} n \int_{g_c}^g dg' \frac{\gamma_1(g') - \gamma_1(g_c)}{\beta(g')} \right\} D_n(\mathbf{p}; M(g), \theta(g), g, \mu) \end{aligned} \quad (220)$$

with

$$\left\{ \mu \frac{\partial}{\partial \mu} - \frac{1}{2} \gamma_1(g_c) (n + M \frac{\partial}{\partial M}) + \gamma_2(g_c) \theta \frac{\partial}{\partial \theta} \right\} \Gamma_n(\mathbf{p}; M, \theta, g_c, \mu) = 0 \quad (221)$$

$$\left\{ \begin{aligned} & \left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \frac{1}{2} \gamma_1(g_c) M \frac{\partial}{\partial M} + \gamma_2(g_c) \theta \frac{\partial}{\partial \theta} \right\} \\ & \times D_n(\mathbf{p}; M, \theta, g, \mu) = 0 \\ & D_n(\mathbf{p}; M, \theta, g_c, \mu) = 1 \end{aligned} \right. \quad (222)$$

The parametrization (220) has been chosen in such a way that, in the equation for D_n , the only reference to g appears in the term $\beta(g) \partial / \partial g$. Moreover, n does not appear explicitly in this equation. The function D_n is normalized to unity at the critical point and is dimensionless, so that

$$D_n \equiv D_n \left(\frac{\mathbf{p}}{\mu}; \frac{M}{\mu^{\frac{1}{2}(d-2)}}, \frac{\theta}{\mu^2}, g, 1 \right) \quad (223)$$

The solution of equation (222) can thus be written as

$$\begin{aligned} D_n(\mathbf{p}; M, \theta, g, \mu) &= D_n(\mathbf{p}; M \lambda^{-\frac{1}{2} \gamma_1(g_c)}, \theta \lambda^{\gamma_2(g_c)}, g(\lambda), \lambda \mu) \\ &= D_n \left(\frac{\mathbf{p}}{\lambda}; M \lambda^{-\frac{1}{2} \gamma_1(g_c) - \frac{1}{2}(d-2)}, \theta \lambda^{\gamma_2(g_c) - 2}, g(\lambda), \mu \right) \end{aligned} \quad (224)$$

where

$$\ln \lambda = \int_g^{g(\lambda)} \frac{dg'}{\beta(g')} \quad (225)$$

Let us now choose λ such that

$$1 = |\theta| \lambda^{\gamma_2(g_c) - 2} = |\theta| \lambda^{-1/\nu} \quad (226)$$

The previous relation reads

$$\begin{aligned} D_n(\mathbf{p}; \theta, M, g, \mu = 1) &= D_n \left(\frac{\mathbf{p}}{|\theta|^\nu}; \frac{M}{|\theta|^\beta}, g(|\theta|^\nu), 1 \right) \\ \nu \ln |\theta| &= \int_g^{g(|\theta|^\nu)} \frac{dg'}{\beta(g')} \end{aligned} \quad (227)$$

We approximate the function $\beta(g)$ by a linear term near g_c

$$\beta(g) \simeq \omega(g - g_c) + \mathcal{O}((g - g_c)^2) \quad (228)$$

If g is close enough to g_c and θ small, $g(|\theta|^\nu)$ is also close to g_c . Indeed, from equations (227) and (228),

$$|\theta|^\nu \omega = \frac{g_c - g(|\theta|^\nu)}{g_c - g} + \dots \quad (229)$$

and hence the correction factor D_n behaves as

$$D_n(\mathbf{p}; \theta, M, g, 1) \simeq 1 + (g - g_c) |\theta|^\nu d_n \left(\frac{\mathbf{p}}{|\theta|^\nu}, \frac{M}{|\theta|^\beta} \right) \quad (230)$$

The exponent

$$\boxed{\omega = \beta'(g_c) > 0} \quad (231)$$

characterizes therefore the deviations from $g = g_c$ independently of which of the Green functions is considered. As indicated in equation (220), this correction is accompanied by a change in the normalization of the correlation functions.

In equation (230), we have expressed the deviations from criticality in terms of the temperature. It is also possible to use the momenta or the magnetization. The corresponding corrections behave as

$$\begin{aligned} & \text{for } |\theta| \text{ as } |\theta|^\omega \\ & \text{for } |\mathbf{p}| \text{ as } |\mathbf{p}|^\omega \\ & \text{for } M \text{ as } M^{\omega\nu/\beta} \end{aligned} \quad (232)$$

To order $\epsilon = 4 - d$, $\beta = \epsilon(-1 + g/g_c)$ where g_c is of order ϵ . Hence

$$\omega = \epsilon + \mathcal{O}(\epsilon^2) \quad (233)$$

As $d \rightarrow 4$, the corrections are confluent with the dominant term.

5.4.2 Logarithmic corrections in dimension four

In the marginal dimension, the stable infrared fixed point g_c coincides with the ultraviolet fixed point at the origin, and the scale invariant theory is the Gaussian model, up to logarithmic corrections. These corrections are thus important for characterizing the

critical behaviour. The equation of state (197) now reads

$$H(M, \theta, g, \mu) = (\lambda\mu)^3 \left[\exp -\frac{1}{2} \int_g^{g(\lambda)} dg' \frac{\gamma_1(g')}{\beta(g')} \right] H \left(1, \frac{\theta(\lambda)}{\lambda^2 \mu^2}, g(\lambda), 1 \right) \quad (234)$$

with

$$\lambda = \frac{M}{\mu} \exp -\frac{1}{2} \int_g^{g(\lambda)} dg' \frac{\gamma_1(g')}{\beta(g')} \quad (235)$$

and, of course, $g(\lambda)$ and $\theta(\lambda)$ are defined through

$$\begin{aligned} \ln \lambda &= \int_g^{g(\lambda)} \frac{dg'}{\beta(g')} \\ \ln \frac{\theta(\lambda)}{\theta} &= \int_g^{g(\lambda)} dg' \frac{\gamma_2(g')}{\beta(g')} \end{aligned} \quad (236)$$

In the critical domain $M \ll \mu^2$, λ is vanishingly small. The function $\beta(g)$ is given by (cf. (115))

$$\beta(g) = \frac{n+8}{3} \frac{g^2}{(4\pi)^2} + \dots \quad (237)$$

therefore

$$\frac{g(\lambda)}{(4\pi)^2} \sim \frac{3}{n+8} \frac{1}{(-\ln \lambda)} \quad (238)$$

This expression is typical of asymptotic infrared freedom. In the exponential (235), $\gamma_1(g')$ is of order g'^2 and the integral is finite, i.e. M is of order λ . Finally, as

$$\gamma_2(g) = \frac{n+2}{8} \frac{g}{(4\pi)^2} + \dots \quad (239)$$

$$\frac{\theta(\lambda)}{\theta} \sim \text{Cst} (-\ln \lambda)^{-(n+2)/(n+8)} \sim \text{Cst} \left(\ln \frac{\mu}{M} \right)^{-(n+2)/(n+8)} \quad (240)$$

Inserting this in (234), one gets

$$\begin{aligned} H(M, \theta, g, \mu = 1) &\simeq M^3 \\ &\times H \left(1, \text{Cst} \frac{\theta}{M^2} (-\ln M)^{-(n+2)/(n+8)}, \frac{3}{n+8} \frac{(4\pi)^2}{(-\ln M)}, 1 \right) \end{aligned} \quad (241)$$

Since $g(\lambda)$ is vanishingly small, the dependence of the r.h.s. on this variable can be computed perturbatively. To lowest order, we

have

$$\begin{aligned} V(M) &= \frac{1}{2} \theta M^2 + \frac{g}{4!} M^4 + \dots \\ H &= \theta M + \frac{1}{6} g M^3 + \dots \end{aligned}$$

Therefore relation (241) becomes (with A denoting a constant)

$$\begin{aligned} H(M, \theta, g, \mu = 1) &\simeq M^3 \left\{ A \frac{\theta}{M^2} (-\ln M)^{-(n+2)/(n+8)} - \frac{(4\pi)^2}{2(n+8) \ln M} \right\} \\ &\times \left\{ 1 + \mathcal{O} \left(\frac{1}{\ln M} \right) \right\} \end{aligned} \quad (242)$$

where the correction arises from the term in $g(\lambda)$.

At the critical temperature, this relation yields

$$H \simeq \frac{M^3}{-\ln M} \quad (243)$$

The critical exponent δ takes its mean field value $\delta = 3$, but there is a logarithmic correction. This result can be obtained heuristically, by replacing g by its running value $g(\lambda) \simeq g(M) \sim -1/\ln M$ in the mean field expression $\frac{1}{6} g M^3$.

The spontaneous magnetization is obtained for vanishing H and $T < T_c$, as

$$M \simeq \sqrt{-\theta} (-\ln -\theta)^{3/(n+8)} \quad (244)$$

with logarithmic corrections (depending on n) to the mean field behaviour characterized by an exponent $\beta = \frac{1}{2}$.

In zero external field and for $\theta \neq 0$, the susceptibility (the longitudinal one for $n \geq 2$ and $\theta < 0$) deduced from (242) reads

$$\chi = |\theta|^{-1} (-\ln |\theta|)^{(n+2)/(n+8)} \quad (245)$$

with γ equal to -1 up to logarithmic corrections.

To study the specific heat, let us look at the equation for the dimensionless function $\Gamma_{0,2}^R$

$$\left\{ \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} + \gamma_2(g) \left[2 + \theta \frac{\partial}{\partial \theta} \right] \right\} \Gamma_{0,2}^R = b(g) \quad (246)$$

where $b(g)$ approaches a finite constant b_0 as g vanishes. Upon integration,

$$\Gamma_{0,2}^R(\mathbf{q}; \theta, g, \mu) = \left[\exp 2 \int_g^{g(\lambda)} dg' \frac{\gamma_2(g')}{\beta(g')} \right] \Gamma_{0,2}^R(\mathbf{q}; \theta(\lambda), g(\lambda), \lambda\mu) - \int_g^{g(\lambda)} \frac{dg'}{\beta(g')} b(g') \exp 2 \int_g^{g'} dg'' \frac{\gamma_2(g'')}{\beta(g'')} \quad (247)$$

At zero momentum, using the fact that $\Gamma_{0,2}^R$ is dimensionless, we have

$$\Gamma_{0,2}^R(0; \theta, g, \mu = 1) = \left[\exp 2 \int_g^{g(\lambda)} dg' \frac{\gamma_2(g')}{\beta(g')} \right] \Gamma_{0,2}^R\left(0; \frac{\theta(\lambda)}{\lambda^2}, g(\lambda), 1\right) - \int_g^{g(\lambda)} \frac{dg'}{\beta(g')} b(g') \exp 2 \int_g^{g'} dg'' \frac{\gamma_2(g'')}{\beta(g'')} \quad (248)$$

Let us choose λ by requiring that $|\theta(\lambda)| = \lambda^2$, thus

$$\lambda^2 \simeq |\theta| \left(\frac{g(\lambda)}{g} \right)^{(n+2)/(n+8)} \quad (249)$$

Since

$$g(\lambda) \simeq \frac{\text{Cst}}{\ln \lambda} \quad (250)$$

the dominant contribution in (248) comes from the inhomogeneous term and one obtains

$$\Gamma_{0,2}^R \simeq \frac{\text{Cst}}{(-\ln |\theta|)^{(n-4)/(n+8)} + \text{Cst}} \quad (251)$$

with again a logarithmic correction.

These results agree in the limit $n \rightarrow \infty$ with those obtained in chapter 3. Note a change in the behaviour of the specific heat for $n = 4$ with a divergence for $n < 4$. For instance, when $n = 1$, the specific heat diverges as $(-\ln |\theta|)^{1/3}$.

Let us again emphasize the paradox of the four-dimensional continuous φ^4 theory. The renormalized theory is free, since the interval $[0, g_c]$ shrinks to zero. The preceding results are meaningful only if one keeps a large, but not infinite, cutoff Λ , so that the interval is not strictly 0, but of order $1/\ln \Lambda$.

5.4.3 Irrelevant operators

We finally discuss the effect of keeping local terms with higher powers of the field or its derivatives in the Lagrangian. The presence of such terms cannot be avoided in any consistent regularization scheme. Close to four dimensions, these terms have coefficients with a negative power of Λ , and therefore were neglected in a first approach. It is crucial to the proof of universality to make sure that, through renormalization effects, they do not affect the leading scaling behaviour. We do not discuss here dimensions low enough for some of these operators to become marginal or relevant, as is a φ^6 term in dimension three. This corresponds to a richer phase diagram, with a possible occurrence of first order transitions and tricritical phenomena (appendix A).

We shall content ourselves here with estimating perturbatively the effect of irrelevant operators, taking into account renormalization effects. We examine an operator $\mathcal{O}_i$ with n_1 fields and n_2 derivatives (symbolically $\partial^{n_2} \varphi^{n_1}$). Its coefficient in the Lagrangian will be

$$g(\mathcal{O}_i) \sim \Lambda^{-\delta_i} \quad (252)$$

and its canonical degree is δ_i . For instance,

$$\begin{aligned} \delta_{\varphi^6} &= 2d - 6 \\ \delta_{\varphi^2(\partial\varphi)^2} &= d - 2 \end{aligned} \quad (253)$$

At first sight, operators with $\delta_i > 0$ are negligible. However, their insertion in Green functions lead to additional divergences, and their effect should therefore be analyzed more carefully. For instance, we have seen the important effect of the marginal φ^4 term in four dimensions. Concerning operators with negative degree, we have already analyzed the φ^2 case which generates logarithmic divergences in dimension four.

Consider now operators with nonnegative degree in dimension four. The results will then be continued using an ϵ -expansion. The insertion of φ^4 corresponds to a shift of the coupling constant from its fixed point value. Vertex functions Γ_{2,φ^4} and Γ_{4,φ^4} are respectively quadratically and logarithmically divergent. The logarithmic divergence of the 4-point function can be treated by a multiplicative renormalization, but the 2-point function needs

Starting with the field φ we can define various families of operators, among them (Wick ordered) exponentials $e^{i\alpha\varphi}$. The quantity α is analogous to a charge, since correlation functions takes the form of Boltzmann weights involving classical Coulomb potentials (i.e. logarithmic in two dimensions) between pairs. Such correlations make sense only when the total charge vanishes, as we shall shortly explain. Wick ordering implies omission of self-contractions, and one finds

$$\left\langle \prod_j e^{i\alpha_j \varphi(z_j, \bar{z}_j)} \right\rangle_{\Sigma \alpha_j = 0} = \prod_{j < k} \left(\frac{z_{jk} \bar{z}_{jk}}{R^2} \right)^{\frac{1}{2} \alpha_j \alpha_k} \quad (165)$$

If instead we were to keep self-contractions with an ultraviolet cutoff $\Lambda = a^{-1}$, the above expression would appear multiplied by a product of factors $\prod_j (a/R)^{\alpha_j^2/2}$, showing that Wick ordering amounts to an (infinite) multiplicative renormalization of $e^{i\alpha\varphi}$. However, this factor also depends on the infrared scale R . Choosing to normalize the operators at a given value R_0 , we may study the remaining dependence on R of the form $R^{-u/2}$,

$$u = \sum_j \alpha_j^2 + \sum_{j \neq k} \alpha_j \alpha_k = (\Sigma \alpha_j)^2$$

Thus correlations are R -independent if and only if we have neutrality: $\Sigma \alpha_j = 0$. Should we violate this condition, correlations would vanish in the limit $R \rightarrow \infty$. This discussion also shows that the dimension of $e^{i\alpha\varphi}$ is $\frac{1}{2}\alpha^2$.

In particular

$$\left\langle e^{i\alpha\varphi(z_1, \bar{z}_1)} e^{-i\alpha\varphi(z_2, \bar{z}_2)} \right\rangle = \frac{1}{(z_{12} \bar{z}_{12})^{\frac{1}{2}\alpha^2}} \quad (166)$$

This means that $e^{i\alpha\varphi}$, or perhaps better the real forms $\cos \alpha\varphi$ and $\sin \alpha\varphi$, have conformal weights

$$h = \bar{h} = \frac{1}{4}\alpha^2 \quad (167)$$

- i) When α is an integer, the associated character is given by equation (159) and the corresponding Verma module admits a singular vector at level $\alpha + 1$. This level is two if $\alpha = 1$, in which case $h = \bar{h} = \frac{1}{4}$ and the correlation (166) reduces to

$C_{12} = (z_{12} \bar{z}_{12})^{-\frac{1}{2}}$. This is in agreement with equation (87), where for $c = 1$, $h_+ = h_- = \frac{1}{4}$. Check that equation (95) is satisfied, as

$$\frac{\partial^2}{\partial z_1} C(z_{12}, \bar{z}_{12}) = \left(\frac{1}{4z_{12}^2} + \frac{1}{z_{12}} \frac{\partial}{\partial z_2} \right) C(z_{12}, \bar{z}_{12})$$

- ii) We have assumed that the field φ takes arbitrary real values. However, the Lagrangian is invariant under $\varphi \rightarrow \varphi + \text{cst}$ and $\varphi \rightarrow -\varphi$. This may be used to restrict to configurations, with the points φ and $\varphi + 2\pi\rho$ identified (ρ a constant) implying that φ takes its values on a circle (of radius ρ), as was the case for the XY-model. Nothing is changed in the correlations of vertex operators provided we assume that the charges α are of the form $\alpha = n/\rho$. We have also the possibility of including other operators corresponding to a frustration line, in the sense that, as $\mathbf{x}$ circles around one of the end points (a vortex), φ increases by an amount $2\pi m\rho$, with m an integer. These defects have been studied in the case of the XY-model. The correlation function of a pair $(m, -m)$ of vortices located at points 1 and 2 is a ratio of frustrated to unfrustrated partition functions. By shifting φ by φ_c , a multivalued harmonic solution (everywhere except at the sources), this ratio is equal to $\exp -S(\varphi_c)$. A similar interpretation can be given to the correlation of charges distinguished by a uniform harmonic function with δ -function sources at points 1 and 2. We can even generalize to a pair of composite $(n, m), (-n, -m)$ operators with

$$\begin{aligned} \varphi_c &= \frac{in}{\rho} \text{Re} \left(\ln \frac{z - z_2}{z - z_1} \right) - \rho m \text{Im} \left(\ln \frac{z - z_2}{z - z_1} \right) \\ &= \frac{1}{2}i \left(\frac{n}{\rho} + m\rho \right) \ln \left(\frac{z - z_2}{z - z_1} \right) - \frac{1}{2}i \left(\frac{n}{\rho} - m\rho \right) \ln \left(\frac{\bar{z} - \bar{z}_2}{\bar{z} - \bar{z}_1} \right) \end{aligned} \quad (168)$$

The occurrence of a factor i is such that in the shift $\varphi \rightarrow \varphi + \varphi_c$, the crossterm in the action yields $(in/\rho)[\varphi(1) - \varphi(2)]$. When one computes the corresponding classical action and performs the necessary subtractions, one finds

$$\begin{aligned} \exp(-S_{n,m}(1,2)) &= \frac{1}{z_{12}^{2h_{n,m}} \bar{z}_{12}^{2\bar{h}_{n,m}}} \\ &\equiv \langle O_{n,m}(z_1, \bar{z}_1) O_{-n,-m}(z_2, \bar{z}_2) \rangle \end{aligned} \quad (169)$$

with

$$\begin{aligned} h_{n,m} &= \frac{1}{4} \left(\frac{n}{\rho} + m\rho \right)^2 \\ \bar{h}_{n,m} &= h_{n,-m} = \frac{1}{4} \left(\frac{n}{\rho} - m\rho \right)^2 \end{aligned} \quad (170)$$

As a consequence, the generalized vertex operators $O_{n,m}$ have an integral spin equal to nm . The fact that the product of charge times vorticity (or magnetic charge) has to be an integer is reminiscent of the analogous property derived by Dirac for the quantum mechanical motion of a point charge in a three-dimensional monopole magnetic field. The operators $O_{n,m}$ have conformal weights depending on a continuous parameter ρ . They occur in the discussion of several models at $c = 1$, with a line of continuously varying exponents like the six vertex or Ashkin-Teller model (see section 3.6).

Similarly one can make use of the symmetry $\varphi \rightarrow -\varphi$ to restrict the values of the field to a so-called orbifold, here the quotient of the circle by a reflection with respect to a diameter.

9.2.2 Ising model

We want to cast in the present framework the critical Ising model presented in chapter 2 (volume 1). We recall that the action at criticality is effectively the one pertaining to a Majorana field. The action and two-point correlations read

$$\begin{aligned} S &= \int \frac{d^2x}{2\pi} \psi \bar{\partial} \psi - \bar{\psi} \partial \bar{\psi} \\ \langle \psi(z_1) \psi(z_2) \rangle &= 1/z_{12} \quad \langle \bar{\psi}(\bar{z}_1) \bar{\psi}(\bar{z}_2) \rangle = 1/\bar{z}_{12} \end{aligned} \quad (171)$$

These correlations are odd in the interchange $1 \leftrightarrow 2$, because of Fermi statistics, and one is analytic, the other antianalytic. Noether's theorem yields the energy-momentum tensor as

$$T = -\frac{1}{2} \psi \partial \psi \quad \bar{T} = -\frac{1}{2} \bar{\psi} \bar{\partial} \bar{\psi} \quad (172)$$

with a normal ordering prescription as usual, so that

$$\langle T(z_1) \psi(z_2) \psi(z_3) \rangle = \frac{z_{23}}{2(z_{12})^2 (z_{13})^2} \quad (173)$$

Equations (171) and (173) are consistent with the fact that ψ has weights $(\frac{1}{2}, 0)$, and $\bar{\psi}$ $(0, \frac{1}{2})$. Thus ψ and $\bar{\psi}$ have half integral

spin, as expected. The Ward identity is in agreement with these assignments, since it follows that

$$\langle T(z_1) \psi(z_2) \psi(z_3) \rangle = \left[\frac{1}{2z_{12}^2} + \frac{1}{2z_{13}^2} + \frac{1}{z_{12}} \partial_2 + \frac{1}{z_{13}} \partial_3 \right] \langle \psi(z_2) \psi(z_3) \rangle$$

The central charge is obtained by computing

$$\begin{aligned} \langle T(z_1) T(z_2) \rangle &= \frac{1}{4} \{ \langle \psi(z_1) \partial \psi(z_2) \rangle \langle \partial \psi(z_1) \psi(z_2) \rangle \\ &\quad - \langle \psi(z_1) \psi(z_2) \rangle \langle \partial \psi(z_1) \partial \psi(z_2) \rangle \} \\ &= \frac{1}{4z_{12}^4} \end{aligned} \quad (174)$$

Consequently

$$c = \frac{1}{2} \quad (175)$$

It is natural at this point to expect that the model is built out from the minimal (unitary) representations of the Virasoro algebra for $c = \frac{1}{2}$, i.e. $m = 3$. We saw that the Kac table furnishes the weights $0, \frac{1}{2}, \frac{1}{16}$, and we can tentatively identify the observables with integral spin, indeed zero, as being those with $h = \bar{h}$, through

	$(h, \bar{h})$	Δ	
identity	$(0, 0)$	0	I
energy density	$(\frac{1}{2}, \frac{1}{2})$	1	$\varepsilon = \bar{\psi} \psi$
spin	$(\frac{1}{16}, \frac{1}{16})$	$\frac{1}{8}$	σ

(176)

The energy density has a simple interpretation in terms of the Majorana field, while the spin σ has a nonlocal expression in $\psi, \bar{\psi}$. The critical exponent η is given by $\eta = 2\Delta_\sigma = \frac{1}{4}$ as expected. The logarithmic divergence of the integral over the energy-energy correlation $\langle \varepsilon \varepsilon \rangle$ implies $\alpha = 0$ (up to a logarithm) and $\nu = 1$. The other critical exponents are given by the standard scaling analysis as $\gamma = \frac{7}{4}$ and $\beta = \frac{1}{8}$. To summarize, once the value of the central charge $c = \frac{1}{2}$ is known, the minimal assumption gives the other exponents as well as a means to compute all critical correlations. Not much is gained in the present case, where integrability (and a long history) has provided us with the results beforehand. We can use this knowledge to check the formalism. But the way is now open to apply the same reasoning to other cases where much less is known, and the power of conformal field theory becomes obvious.

Since $\frac{1}{2}$ and $\frac{1}{16}$ are the two solutions $h_{\pm}$ of equations (87)–(90) implying the existence of singular vectors at level two, the corresponding correlation should satisfy a second order (partial) differential equation. According to chapter 2

$$\langle \varepsilon(1) \cdots \varepsilon(2n) \rangle = \left| \text{Pf} \frac{1}{z_{ij}} \right|^2 \quad (177)$$

$$\langle \sigma(1) \cdots \sigma(2n) \rangle^2 = \sum_{\substack{\alpha_i = \pm 1 \\ \sum \alpha_i = 1}} \prod_{i < j} (z_{ij} \bar{z}_{ij})^{\frac{1}{4} \alpha_i \alpha_j} \quad (178)$$

Take the four point functions as an example. With

$$\langle \varepsilon(1) \varepsilon(2) \varepsilon(3) \varepsilon(4) \rangle = |E(z_1, z_2, z_3, z_4)|^2$$

$$E(z_1, z_2, z_3, z_4) = \frac{1}{z_{12} z_{34}} - \frac{1}{z_{13} z_{24}} + \frac{1}{z_{14} z_{23}}$$

One checks that

$$\left\{ \frac{3}{4} \frac{\partial^2}{\partial z_1^2} - \frac{1}{2} \left(\frac{1}{z_{12}^2} + \frac{1}{z_{13}^2} + \frac{1}{z_{14}^2} \right) - \left(\frac{1}{z_{12}} \partial_2 + \frac{1}{z_{13}} \partial_3 + \frac{1}{z_{14}} \partial_4 \right) \right\} E(z_1, z_2, z_3, z_4) = 0$$

The case of the four spin correlation is more interesting, since it is not the modulus square of an analytic function, but rather a combination of the independent solutions of the corresponding equation. First rewrite explicitly the expression (178) for $n = 2$, with $r_{ij}^2 = z_{ij} \bar{z}_{ij}$, as

$$\langle \sigma(1) \sigma(2) \sigma(3) \sigma(4) \rangle^2 = \left(\frac{r_{14} r_{32}}{r_{12} r_{34} r_{13} r_{42}} \right)^{\frac{1}{2}} + \left(\frac{r_{13} r_{42}}{r_{12} r_{34} r_{14} r_{32}} \right)^{\frac{1}{2}} + \left(\frac{r_{12} r_{34}}{r_{13} r_{42} r_{14} r_{32}} \right)^{\frac{1}{2}}$$

The equation is

$$\left\{ \frac{4}{3} \frac{\partial}{\partial z_1^2} - \frac{1}{16} \left(\frac{1}{z_{12}^2} + \frac{1}{z_{13}^2} + \frac{1}{z_{14}^2} \right) - \left(\frac{1}{z_{12}} \partial_2 + \frac{1}{z_{13}} \partial_3 + \frac{1}{z_{14}} \partial_4 \right) \right\} \langle \sigma(1) \sigma(2) \sigma(3) \sigma(4) \rangle = 0$$

The correlation is real, symmetric and positive. From the asymptotic factorization condition,

$$\lim_{\mathbf{R} \rightarrow \infty} \langle \sigma(\mathbf{r}_1 + \mathbf{R}) \sigma(\mathbf{r}_2 + \mathbf{R}) \sigma(\mathbf{r}_3) \sigma(\mathbf{r}_4) \rangle = \langle \sigma(\mathbf{r}_1) \sigma(\mathbf{r}_2) \rangle \langle \sigma(\mathbf{r}_3) \sigma(\mathbf{r}_4) \rangle$$

its normalization follows from the one of the two-point function, here $\langle \sigma(1) \sigma(2) \rangle = \sqrt{2}/r_{12}^{1/4}$. Using global invariance, and the cross-

ratio

$$x = \frac{z_{12} z_{34}}{z_{14} z_{32}}$$

with $x = 0, 1, \infty$ as z_1 coincides with z_2, z_3 or z_4 , we write

$$\langle \sigma(1) \sigma(2) \sigma(3) \sigma(4) \rangle = \left(\frac{z_{14} z_{32}}{z_{12} z_{34} z_{13} z_{42}} \right)^{\frac{1}{8}} (c.c.)^{\frac{1}{8}} f(x, \bar{x})$$

The powers in the prefactors are in agreement with the weights of σ . The equation takes the hypergeometric form

$$\left\{ x(1-x) \frac{\partial^2}{\partial x^2} + \left(\frac{1}{2} - x \right) \frac{\partial}{\partial x} + \frac{1}{16} \right\} f = 0$$

A general solution is a combination of $\sqrt{1 \pm \sqrt{1-x}}$. We have the same equation for the $\bar{x}$ dependence. Consequently with $a_{\pm\pm}$ constants

$$f = \sum_{\pm\pm} a_{\pm\pm} \sqrt{1 \pm \sqrt{1-x}} \sqrt{1 \pm \sqrt{1-\bar{x}}}$$

Ignoring the overall multiplicative normalization, we are left with three constants to be determined by the symmetry (and reality) of the solution in any interchange $(z_i, \bar{z}_i) \leftrightarrow (z_j, \bar{z}_j)$. One also requires monodromy of the solution, namely that the correlation function be single valued. This can be checked by following a path surrounding any of the singularities, 0, 1, ∞ . Analyzing these constraints gives $a_{++} = a_{--}$, $a_{+-} = a_{-+} = 0$. Restoring the normalization yields

$$\langle \sigma(1) \sigma(2) \sigma(3) \sigma(4) \rangle = \frac{1}{\sqrt{2}} \left| \frac{z_{14} z_{32}}{z_{12} z_{34} z_{13} z_{42}} \right|^{\frac{1}{4}} \times \left\{ \sqrt{1 + \sqrt{1-x}} \sqrt{1 + \sqrt{1-\bar{x}}} + \sqrt{1 - \sqrt{1-x}} \sqrt{1 - \sqrt{1-\bar{x}}} \right\}$$

Squaring this expression we see that it agrees with the expected result (178).

9.2.3 Three state Potts model

The Potts model was introduced in chapter 4 (volume 1). Here we specialize to the three state case. On a square lattice the partition function is

$$Z = \sum_{\varphi_{\mathbf{x}}=0, \pm 2\pi/3} \exp \beta \sum_{(\mathbf{x}, \mathbf{x}')} \frac{2}{3} [\cos(\varphi_{\mathbf{x}} - \varphi_{\mathbf{x}'}) + \frac{1}{2}] \quad (179)$$

The “spin” $\sigma = e^{i\varphi}$ takes as values the cubic roots of the identity, and the exponent is such that the contribution of neighbouring spins is $\delta_{\sigma\sigma'}$. The model admits a global Z_3 symmetry, and obeys a duality relation with a fixed point

$$e^{\beta_c} = 1 + \sqrt{3} \quad (180)$$

More generally for a Q -state model $e^{\beta_c} = 1 + \sqrt{Q}$. The case $Q = 2$ is the Ising one (up to a rescaling of β). For $Q > 4$, it is known to have a first order transition while for $Q \leq 4$ the transition is continuous. Baxter (1980) has shown that for $Q = 3$, two critical exponents can be obtained

$$\alpha = \frac{1}{3} \quad \beta = \frac{1}{9} \quad (181)$$

from the assumption that the three state Potts model is in the same universality class as the hard hexagon model which he solved.

The latter is a lattice gas model where “molecules” in the shape of hexagons have their centers on the nodes of a triangular lattice. Hexagons covering six elementary triangles are forbidden to overlap. One computes a generating function for the number of such configurations. The problem admits a Z_3 symmetry corresponding to the three sublattices on which the gas may “crystallize”. Since both the hard hexagon and the three state Potts model admit a Z_3 symmetry and are “ferromagnetic” models with short-range interactions, it is most likely that they belong to the same universality class.

From the scaling relations $\nu d = 2 - \alpha$, and $\beta = \frac{1}{2}\nu(d - 2 + \eta)$ it follows that

$$\nu = \frac{5}{6} \quad \eta = \frac{4}{15} \quad (182)$$

The real field $2\cos\varphi = \sigma + \bar{\sigma}$ should therefore have conformal weights $(\frac{1}{15}, \frac{1}{15})$. Similarly the thermal operator (coupled to $\beta - \beta_c$), denoted ε as before, has a dimension $\Delta_\varepsilon = \frac{1}{2}(d - \alpha/\nu) = \frac{4}{5}$. The corresponding conformal field theory should therefore contain the scalar fields

$$\begin{aligned} \sigma + \bar{\sigma} &\rightarrow A_{\frac{1}{15}, \frac{1}{15}} \\ \varepsilon &\rightarrow A_{\frac{2}{5}, \frac{2}{5}} \end{aligned} \quad (183)$$

We now scan the Kac tables for the unitary minimal representations of the Virasoro algebra (Figure 4), to find that for

$$m = 5, \quad c = \frac{4}{5} \quad (184)$$

one finds indeed the two weights $\frac{1}{15}$ and $\frac{2}{5}$, with the second admitting a singular vector in the associated Verma module at level two again.

The table contains many more candidates as conformal weights, and the previous assignments are still tentative. It is however possible by an accurate numerical treatment, using the finite scaling method outlined in section 1.4, to confirm the value $c = \frac{4}{5}$. Further arguments will be presented in section 3.

Check the consistency of the above assignments using the three point function

$$\langle \varepsilon(z_1, \bar{z}_1) \cos \varphi(z_2, \bar{z}_2) \cos \varphi(z_3, \bar{z}_3) \rangle = \text{cst} \frac{1}{z_{12}^{h_\varepsilon} z_{13}^{h_\varepsilon} z_{23}^{2h_\sigma - h_\varepsilon}} \times (\text{c.c}) \quad (185)$$

which should satisfy the differential equation

$$\left[\frac{3}{2(2h_\varepsilon + 1)} \partial_1^2 - \sum_{2,3} \left(\frac{h_\sigma}{(z_1 - z_k)^2} + \frac{1}{z_1 - z_k} \partial_k \right) \right] \times \langle \varepsilon(1) \cos \varphi(2) \cos \varphi(3) \rangle = 0.$$

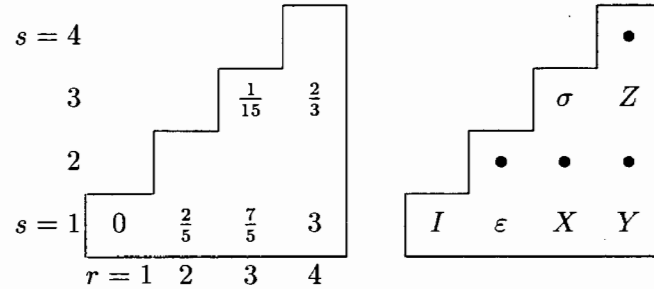
This requires that

$$h_\sigma = \frac{h_\varepsilon(1 - h_\varepsilon)}{2(2h_\varepsilon + 1)} \quad (186)$$

and is indeed satisfied for $h_\varepsilon = \frac{2}{5}$, $h_\sigma = \frac{1}{15}$ (as well as $h_\varepsilon = \frac{1}{2}$, $h_\sigma = \frac{1}{16}$ in the Ising case).

It turns out that not all of the conformal weights occurring in the Kac table for $m = 5$ do appear for the Potts model. This will be discussed in the next section, when we look at requirements of a global nature. As we shall see, the correspondence between

possible weights and scalar primary fields is established as follows



Only the rows with $s = 1, 3$ occur in the model. We list first the scalar fields including degeneracy (i.e. the number of independent such primary fields). For instance, the spin field is expected to have degeneracy two.

$A_{0,0}$	identity	I	
$A_{\frac{2}{5}, \frac{2}{5}}$	energy	ε	
$A_{\frac{7}{5}, \frac{7}{5}}$		X	
$A_{3,3}$		Y	
$A_{\frac{1}{15}, \frac{1}{15}}$	spin field	σ	2 such fields
$A_{\frac{2}{3}, \frac{2}{3}}$		Z	2 such fields

(187)

To this list should be added fields with nonzero angular momentum (respectively ± 3 and ± 1)

$A_{0,3}$	$A_{3,0}$	
$A_{\frac{2}{5}, \frac{7}{5}}$	$A_{\frac{7}{5}, \frac{2}{5}}$	

(188)

allowing for a great diversity of observables.

It is possible to define an extension of the Q -state Potts model to continuous values of Q varying between 0 and 4. The central charge is given by the standard formula $c = 1 - 6/m(m+1)$, with a relation

$$Q = 4 \cos^2 \frac{\pi}{m+1} \quad (189)$$

The cases $m = 2, 3, 5$ correspond to percolation (although when $c = 0$ one uses an infinite sequence of nontrivial representations) Ising, and three state Potts model. The case $Q = 4$ appears as a limiting one with $m \rightarrow \infty$, $c \rightarrow 1$. See section 3.6.

Taking the value $c = \frac{4}{5}$ for granted in the case $Q = 3$, let us show a nontrivial application by computing, according to Dotsenko, the correlation function of four energy operators ε , given that it must satisfy a particular case of equation (95), namely

$$\left\{ \frac{5}{6} \partial_1^2 - \sum_{k=2,3,4} \left[\frac{2}{5} \frac{1}{z_{1k}^2} + \frac{1}{z_{1k}} \partial_k \right] \right\} \langle \varepsilon(1) \varepsilon(2) \varepsilon(3) \varepsilon(4) \rangle = 0 \quad (190)$$

Set

$$\langle \varepsilon(1) \varepsilon(2) \varepsilon(3) \varepsilon(4) \rangle = \left(\frac{z_{13} z_{24}}{z_{12} z_{23} z_{34} z_{41}} \right)^{\frac{4}{5}} \times (\text{c.c.})^{\frac{4}{5}} f(x, \bar{x}) \quad (191)$$

$$x = \frac{z_{12} z_{34}}{z_{13} z_{24}}$$

Substituting, we again obtain a hypergeometric equation

$$\left[x(1-x) \frac{\partial^2}{\partial x^2} + \frac{2}{5} (2x-1) \frac{\partial}{\partial x} - \frac{8}{25} \right] f(x, \bar{x}) = 0 \quad (192)$$

Recall that the hypergeometric equation

$$\left[x(1-x) \frac{d^2}{dx^2} + (\gamma - (1+\alpha+\beta)x) \frac{d}{dx} - \alpha\beta \right] f(x) = 0 \quad (193)$$

admits the general solution

$$f(x) = A F(\alpha, \beta; \gamma; x) + B x^{1-\gamma} (1-x)^{\gamma-\alpha-\beta} F(1-\alpha, 1-\beta; 2-\gamma; x) \quad (194)$$

where F is the Gauss series

$$F(\alpha, \beta; \gamma; x) = 1 + \frac{\alpha\beta}{\gamma} \frac{x}{1!} + \dots + \frac{(\alpha)_n (\beta)_n}{(\gamma)_n} \frac{x^n}{n!} + \dots$$

$$= \frac{\Gamma(\gamma)}{\Gamma(\alpha)\Gamma(\gamma-\alpha)} \int_0^1 dt t^{\alpha-1} (1-t)^{\gamma-\alpha-1} (1-xt)^{-\beta} \quad (195)$$

using the notation $(\alpha)_n = \Gamma(\alpha+n)/\Gamma(\alpha) = \alpha(\alpha+1)\dots(\alpha+n-1)$.

Here $\alpha = -\frac{8}{5}$, $\beta = -\frac{1}{5}$, $\gamma = -\frac{2}{5}$. From the fact that $f(x, \bar{x})$ satisfies the same equation in $\bar{x}$, one derives

$$\langle \varepsilon(1) \varepsilon(2) \varepsilon(3) \varepsilon(4) \rangle = u_{11} \left| \frac{z_{13} z_{24}}{z_{12} z_{23} z_{34} z_{41}} \right|^{\frac{8}{5}} \left| F\left(-\frac{8}{5}, -\frac{1}{5}; -\frac{2}{5}; x\right) \right|^2$$

$$+ u_{12} \left(\frac{z_{13} z_{24}}{z_{12} z_{23} z_{34} z_{41}} \right)^{\frac{4}{5}} \frac{(\bar{z}_{13} \bar{z}_{24})^{-2}}{(\bar{z}_{12} \bar{z}_{23} \bar{z}_{34} \bar{z}_{41})^{-\frac{3}{5}}}$$

$$\begin{aligned}
& \times F\left(-\frac{8}{5}, -\frac{1}{5}; -\frac{2}{5}; x\right) F\left(\frac{13}{5}, \frac{6}{5}, \frac{12}{5}; \bar{x}\right) \\
& + u_{21} \frac{(z_{13}z_{24})^{-2}}{(z_{12}z_{23}z_{34}z_{41})^{-\frac{2}{5}}} \left(\frac{\bar{z}_{13}\bar{z}_{24}}{\bar{z}_{12}\bar{z}_{23}\bar{z}_{34}\bar{z}_{41}}\right)^{\frac{4}{5}} \\
& \times F\left(-\frac{8}{5}, -\frac{1}{5}; -\frac{2}{5}; \bar{x}\right) F\left(\frac{13}{5}, \frac{6}{5}, \frac{12}{5}; x\right) \\
& + u_{22} \frac{|z_{13}z_{24}|^{-4}}{|z_{12}z_{23}z_{34}z_{41}|^{-\frac{6}{5}}} |F\left(\frac{13}{5}, \frac{6}{5}, \frac{12}{5}; x\right)|^2
\end{aligned} \quad (196)$$

Again we want to find the constants u_{ij} in such a way that the correlation be uniform. Set

$$\mathbf{F} = \begin{pmatrix} F_1(x) \\ F_2(x) \end{pmatrix} \equiv \begin{pmatrix} F(\alpha, \beta; \gamma; x) \\ x^{1-\gamma}(1-x)^{\gamma-\alpha-\beta} F(1-\alpha, 1-\beta; 2-\gamma; x) \end{pmatrix} \quad (197)$$

Under analytic continuation around a closed loop avoiding the singularities at $x = 0, 1, \infty$, the vector $\mathbf{F}$ gets linearly transformed by a matrix with constant coefficients. These monodromy matrices are independent of continuous deformations of the paths avoiding the singularities. They are generated by two elementary loops around the points $x = 0$ and $x = 1$ (the loop around $x = \infty$ is equivalent to their product). Call g_0 and g_1 the respective actions on $\mathbf{F}$. One has obviously

$$g_0 \mathbf{F} = \begin{pmatrix} 1 & 0 \\ 0 & \omega_0 \end{pmatrix} \mathbf{F} \quad \omega_0 = \exp -2i\pi\gamma \quad (198)$$

To obtain g_1 , it is convenient to transform $\mathbf{F}$ to an equivalent $\mathcal{F}$ adapted to find the regular and singular solutions at $x = 1$, i.e.

$$\mathcal{F} = \begin{pmatrix} F(\alpha, \beta; \alpha + \beta + 1 - \gamma; 1 - x) \\ x^{1-\gamma}(1-x)^{\gamma-\alpha-\beta} F(1-\alpha, 1-\beta; 1 + \gamma - \alpha - \beta; 1 - x) \end{pmatrix} \quad (199)$$

From the integral representation (195) one has

$$\begin{aligned}
\mathbf{F} &= \begin{pmatrix} A & B \\ \lambda A & \lambda' B \end{pmatrix} \mathcal{F} \\
A &= \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)} & B &= \frac{\Gamma(\gamma)\Gamma(\alpha+\beta-\gamma)}{\Gamma(\alpha)\Gamma(\beta)} \\
\lambda &= \frac{\Gamma(2-\gamma)\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}{\Gamma(\gamma)\Gamma(1-\alpha)\Gamma(1-\beta)} & \lambda' &= \frac{\Gamma(2-\gamma)\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+1-\gamma)\Gamma(\beta+1-\gamma)}
\end{aligned} \quad (200)$$

Since

$$g_1 \mathcal{F} = \begin{pmatrix} 1 & 0 \\ 0 & \omega_1 \end{pmatrix} \mathcal{F} \quad \omega_1 = \exp -2i\pi(\gamma - \alpha - \beta) \quad (201)$$

it follows that

$$g_1 \mathbf{F} = \frac{1}{\lambda' - \lambda} \begin{pmatrix} \lambda' - \lambda\omega_1 & \omega_1 - 1 \\ \lambda\lambda'(1 - \omega_1) & \lambda'\omega_1 - \lambda \end{pmatrix} \mathbf{F} \quad (202)$$

The two transformations (198) and (202) become unitary for α, β, γ real and $\lambda\lambda' < 0$, which is verified in the present case, if we rescale the components of $\mathbf{F}$ as

$$\tilde{\mathbf{F}} = \begin{pmatrix} 1 & 0 \\ 0 & (-\lambda\lambda')^{-\frac{1}{2}} \end{pmatrix} \mathbf{F} \quad (203)$$

leading to

$$\begin{aligned}
g_0 \tilde{\mathbf{F}} &= \begin{pmatrix} 1 & 0 \\ 0 & \omega_0 \end{pmatrix} \tilde{\mathbf{F}} \\
g_1 \tilde{\mathbf{F}} &= \frac{1}{\lambda' - \lambda} \begin{pmatrix} \lambda' - \lambda\omega_1 & \sqrt{-\lambda\lambda'}(\omega_1 - 1) \\ \sqrt{-\lambda\lambda'}(\omega_1 - 1) & \lambda'\omega_1 - \lambda \end{pmatrix} \tilde{\mathbf{F}}
\end{aligned} \quad (204)$$

As a result, the combination $|F_1|^2 - |F_2|^2 / \lambda\lambda'$ is invariant under g_0 and g_1 , hence under the full monodromy group. Correspondingly, up to an overall factor, the final expression for the correlation is

$$\begin{aligned}
& \langle \varepsilon(1)\varepsilon(2)\varepsilon(3)\varepsilon(4) \rangle \\
&= \frac{\Gamma(\frac{6}{5})\Gamma(-\frac{1}{5})^2\Gamma(-\frac{8}{5})}{\Gamma(-\frac{2}{5})^2} \left| \frac{z_{13}z_{24}}{z_{12}z_{23}z_{34}z_{41}} \right|^{\frac{8}{5}} \left| F\left(-\frac{8}{5}, -\frac{1}{5}; -\frac{2}{5}; \frac{z_{12}z_{34}}{z_{13}z_{24}}\right) \right|^2 \\
&- \frac{\Gamma(\frac{13}{5})\Gamma(\frac{6}{5})^2\Gamma(-\frac{1}{5})}{\Gamma(\frac{12}{5})^2} \frac{|z_{13}z_{24}|^{-4}}{|z_{12}z_{23}z_{34}z_{41}|^{-\frac{6}{5}}} \left| F\left(\frac{13}{5}, \frac{6}{5}, \frac{12}{5}; \frac{z_{12}z_{34}}{z_{13}z_{24}}\right) \right|^2
\end{aligned} \quad (205)$$

Both coefficients are positive, since $\Gamma(-\frac{8}{5})$ as well as $-\Gamma(-\frac{1}{5})$ are positive. Moreover the full expression is symmetric in its arguments.

- i) Check the symmetry of the correlation.
- ii) Applying the same procedure in the case of correlations involving two spins and two energies, derive Dotsenko's expression

$$\begin{aligned}
& \langle \cos \varphi(1)\varepsilon(2) \cos \varphi(3)\varepsilon(4) \rangle \\
&= \frac{\Gamma(-\frac{4}{5})\Gamma(\frac{3}{5})\Gamma(\frac{6}{5})\Gamma(-\frac{1}{5})}{\Gamma(\frac{2}{5})^2} \frac{|z_{13}|^{\frac{4}{5}}}{|z_{12}z_{23}z_{34}z_{41}|^{\frac{4}{5}}} \left| F\left(-\frac{4}{5}, \frac{3}{5}, \frac{2}{5}; \frac{z_{12}z_{34}}{z_{13}z_{24}}\right) \right|^2 \\
&- \frac{\Gamma(\frac{9}{5})\Gamma(\frac{2}{5})\Gamma(\frac{6}{5})\Gamma(-\frac{1}{5})}{\Gamma(\frac{8}{5})^2} \frac{|z_{13}|^{-\frac{16}{5}}|z_{24}|^{-\frac{12}{5}}}{|z_{12}z_{23}z_{34}z_{41}|^{-\frac{2}{5}}} \left| F\left(\frac{2}{5}, \frac{9}{5}, \frac{8}{5}; \frac{z_{12}z_{34}}{z_{13}z_{24}}\right) \right|^2
\end{aligned} \quad (206)$$

While it is true that the method is very powerful, the explicit expressions soon become very cumbersome. Alternative integral representations are discussed in the literature. Nevertheless it is comforting to see such nontrivial results emerging from a very general approach. They allow verification of the consistency of the theory, such as tested by the implementation of short-distance expansions.

Rather than carrying the investigation in this direction or presenting other models, we turn to the study of the same consistency conditions following ideas put forward by Cardy. A new idea will emerge from the study of finite size effects, namely modular invariance.

9.3 Finite size effects and modular invariance

In this section we present further developments and in particular, a classification of universality classes of minimal models.

9.3.1 Partition functions on a torus

In section 1.4, we have shown the relation between the central charge and finite size (or Casimir) effects in a periodic strip. We want to look at the matter in greater detail. As we recall, the transfer matrix per unit length on a periodic strip, with width L , reads

$$\mathcal{T} = \exp \left\{ -\frac{2\pi}{L} (L_0 + \bar{L}_0 - \frac{1}{12}c) \right\} \quad (207)$$

with the shift $\exp(\frac{1}{6}\pi c/L)$ being the above mentioned effect. The normalization is such that the free energy per unit area vanishes in the bulk. Hence, in a box of size L, M , with periodic conditions on both sides, the partition function reads

$$Z = \text{Tr} \exp \left\{ -2\pi \frac{M}{L} (L_0 + \bar{L}_0 - \frac{1}{12}c) \right\} \quad (208)$$

A box with periodic boundary conditions has the topology of a torus, justifying the title of this section. Let us generalize slightly the construction by accompanying the “time” translation M by an additional “space” translation (figure 5). While in the

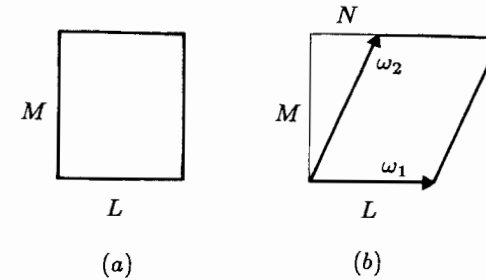


Fig. 5 (a) A rectangular periodic box. (b) More general periodic box; ω_1 and ω_2 are two generators of a corresponding lattice in complex notation, with $|\omega_1| = L$, $|\omega_2| = \sqrt{M^2 + N^2}$.

plane $L_0 + \bar{L}_0$ generates dilatations and $i(L_0 - \bar{L}_0)$ rotations, after a logarithmic map they become generators of time and space translations. Thus if we impose periodic boundary conditions according to the parallelogram of figure 5b, the partition function may be cast into the form

$$\begin{aligned} Z &= \text{Tr} \exp -2\pi \frac{M}{L} (L_0 + \bar{L}_0 - \frac{1}{12}c) + 2i\pi \frac{N}{L} (L_0 - \bar{L}_0) \\ &= \text{Tr} \left(q^{L_0 - c/24} \bar{q}^{\bar{L}_0 - c/24} \right) \end{aligned} \quad (209)$$

where we have set

$$\begin{aligned} \omega_1 &= L, & \omega_2 &= N + iM & \tau &= \omega_2/\omega_1 \\ q &= \exp 2i\pi\tau & \bar{q} &= \exp -2i\pi\bar{\tau} \end{aligned} \quad (210)$$

The conventions are such that $\text{Im} \tau > 0$, hence $|q| < 1$. The purpose of this choice of boundary conditions is to disentangle the dependence on the τ and $\bar{\tau}$ (or q and $\bar{q}$) variables. Suppose that we decompose the space of states into irreducible representations of the product of algebras $(V, \bar{V})$, each one pertaining to highest weights $(h, \bar{h})$. To these representations will be associated primary operators with the same weights. Let $N_{h, \bar{h}}$ be the number of times the representation $(h, \bar{h})$ occurs in this decomposition. If the number of distinct irreducible representations is finite (or denumerable), we can use the definition of conformal characters

according to equations (152), (155) and (158) to write

$$Z = \sum N_{h,\bar{h}} \chi_{(c,h)}(\tau) \chi_{(c,\bar{h})}(\bar{\tau}) \quad (211)$$

Not only are the integers $N_{h,\bar{h}}$ positive, but the uniqueness of the ground state, or vacuum state, invariant under global transformations (in the plane $SL(2, \mathbb{C})$, on the torus translations), requires $N_{0,0} = 1$. Since we use real representations, $\chi_{c,\bar{h}}(\bar{\tau}) = \chi_{c,h}(\tau)$. In a rectangular box, $q = \bar{q} = \exp(-2\pi M/L)$. Let us consider the strip limit where $M/L \rightarrow \infty$, $q \rightarrow 0$. If we assume that all primary operators have a positive dimension $\Delta = h + \bar{h} > 0$, we easily recover the limiting behaviour

$$Z \underset{M/L \rightarrow \infty}{\sim} \exp \frac{1}{6} \pi c M/L \quad (212)$$

as was claimed in section 1.4. Any state with $\Delta_n = h + \bar{h} + n$ will contribute an exponentially small correction, $\exp(-2\pi \Delta_n M/L)$. Thus the spectrum of the transfer matrix yields the set of all conformal dimensions in the theory, with the ratios of eigenvalues

$$\lambda_n/\lambda_0 = \exp(-2\pi \Delta_n/L) \quad (213)$$

such that $\ln(\lambda_n/\lambda_0)$ scales as L^{-1} . If however some operators carry a negative dimension, the leading behaviour (212) will be modified, as anticipated in section 1.4, and

$$Z \underset{M/L \rightarrow \infty}{\sim} \exp \left(\frac{1}{6} \pi c_{eff} M/L \right) \quad (214)$$

$$c_{eff} = c - 12\Delta_{inf} \quad \Delta_{inf} < 0$$

Similarly, the succession of eigenvalues of the transfer matrix will be shifted by $-\Delta_{inf}$, as now

$$\lambda_n/\lambda_0 = \exp(-2\pi(\Delta_n - \Delta_{inf})/L) \quad (215)$$

This shows that one should be careful in interpreting numerical results.

A typical example of this phenomenon is the Lee-Yang edge singularity with $c = -\frac{22}{5}$, $\Delta_{inf} = -\frac{2}{5}$. One predicts (and indeed observes numerically) that $c_{eff} = \frac{2}{5}$. More generally in a minimal model (equations (146) and (147)) characterized by a pair (p, p') one has in the worst case

$$\Delta_{inf} = \frac{1 - (p - p')^2}{2pp'} \quad c_{eff} = 1 - \frac{6}{pp'} \geq 0 \quad (216)$$

The partition function in a periodic box has to satisfy global consistency constraints. From Euclidean invariance, space and time play a symmetric role. We can interchange them, so that in a rectangle of size L, M , the partition function assumes the two alternative expressions $Z = \text{Tr}(\mathcal{T}_L)^M = \text{Tr}(\mathcal{T}_M)^L$. None of these expressions exhibits this symmetry explicitly, and its implementation is a strong limitation on the choice of coefficients in the expansion (211). In more general terms, a torus can be considered as a unit cell (fundamental domain) when taking the quotient of the complex plane by a lattice of translations Λ , generated by two independent periods ω_1 and ω_2 . Any two other periods ω'_1 and ω'_2 will also generate Λ , provided the relation between (ω_1, ω_2) and (ω'_1, ω'_2) is both ways linear with integral coefficients, i.e. invertible, and of unit determinant, in order to preserve the area as well as the orientation. This gives rise to the group $SL(2, \mathbb{Z})$. Moreover, from rotational and dilatation invariance, the modular ratio $\tau = \omega_2/\omega_1$ is the only significant variable. The above transformations act on τ as fractional linear transformations with integral coefficients and determinant unity, with matrices $\pm A$ are identified. This is the modular group $SL(2, \mathbb{Z})/Z_2$ (written as $PSL(2; \mathbb{Z})$) and the constraint of invariance is called modular invariance.

9.3.2 Kronecker's limit formula

Let us return to the Gaussian model to test modular invariance. The space of states is the Fock space of a free field. Eigenmodes correspond to a quantized momentum p_x taking integral values in units $2\pi/L$. Right movers have a positive p_x , left movers a negative one, the modes being associated with the propagation on a circle. We subtract the zero mode which would lead to an infinity by a multiplicative renormalization of Z . Each mode has an occupation number ranging from zero to infinity, and the central charge is $c = 1$. As a result we expect that the partition function is a product of Bose statistical factors of the form

$$Z_1 \propto (q\bar{q})^{-\frac{1}{24}} \prod_{n=1}^{\infty} \frac{1}{(1 - q^n)(1 - \bar{q}^n)} = \frac{1}{\eta(\tau)\eta(\bar{\tau})} \quad (217)$$

in agreement with equation (153) for $c = 1$, $h = 0$, using the definition (157) of Dedekind's function $\eta(\tau)$.

A direct proof of (217), with further generalizations, is known since the work of Kronecker in the 1880's in the context of number theory. The proof proceeds from a direct evaluation of the path integral and exhibits an unexpected factor, missing in (217) and essential for modular invariance.

Let k^i be the generators of the lattice dual to Λ ,

$$\operatorname{Re} k^i \bar{\omega}_j = \delta^i_j \quad k^1 = -i\omega_2/A \quad k^2 = i\omega_1/A \quad (218)$$

where $A = \operatorname{Im} \omega_2 \bar{\omega}_1$ is the area of the torus. The eigenvalues of the Laplacian are

$$E_{n_1, n_2} = (2\pi)^2 |n_1 k^1 + n_2 k^2|^2 \quad (219)$$

with $n_{1,2}$ integers. The normalized zero mode eigenfunction on $T = \mathbb{C}/\Lambda$ is $\varphi_0 = A^{-1/2}$. Omitting this zero mode, the functional integral expressing the partition function is formally

$$\begin{aligned} Z_1 &= \int \mathcal{D}\varphi A^{1/2} \delta \left(\int_T d^2x \varphi \varphi_0 \right) \exp \left(-\frac{1}{2} \int_T d^2x (\nabla \varphi)^2 \right) \\ &= A^{1/2} \prod'_{n_1, n_2} \frac{1}{E_{n_1, n_2}^{1/2}} \end{aligned} \quad (220)$$

The purpose of the factor $A^{1/2}$ is to make the final result dimensionless. Note that the area A is modular invariant. In equation (220), the product over modes is ultraviolet divergent since E_{n_1, n_2} is unbounded for large n_1, n_2 . The prime in the product stands for the omission of the contribution $n_1 = n_2 = 0$. A standard regularization (and renormalization) uses the so-called ζ -regularization. The name derives from the analytic continuation of Riemann's ζ function used in studying the distribution of primes. Introduce the analogous function

$$G(s) = \sum'_{n_1, n_2} \frac{1}{E_{n_1, n_2}^s} \quad (221)$$

absolutely convergent, hence analytic, for $\operatorname{Re} s > 1$. As we shall see below, this function has a pole at $s = 1$, but admits an analytic continuation down to $s = 0$. Comparing the expressions (220) and

(221), one defines the renormalized Z_1 as

$$Z_1 = A^{1/2} \exp \frac{1}{2} \frac{dG}{ds}(0) \quad (222)$$

As before, let τ stand for the ratio ω_2/ω_1 and $q = \exp 2i\pi\tau$. For $\operatorname{Re} s > 1$, we insert the definition of E_{n_1, n_2} in (221), resulting in

$$\left| \frac{2\pi\omega_1}{A} \right|^{2s} G(s) = \sum'_{m, n} \frac{1}{|m + n\tau|^{2s}} = 2\zeta(2s) + \sum'_n \left(\sum_m \frac{1}{|m + n\tau|^{2s}} \right) \quad (223)$$

This starting point of the calculation has explicit modular invariance, for if we substitute in $G(s)$, $\omega_1 = d\omega'_1 + c\omega'_2$, $\omega_2 = b\omega'_1 + a\omega'_2$, hence $\tau = (a\tau' + b)/(c\tau' + d)$, $ad - bc = 1$, we readily derive that $G(s)$ is invariant. Consequently, by analytic continuation, this will follow for Z_1 .

Returning to equation (223), the first term involves Riemann's function $\zeta(s) = \sum_1^\infty m^{-s}$ with a simple pole at $s = -1$ and such that $2\zeta(0) = -1$, $2\zeta'(0) = -\ln 2\pi$. In the second term, the sum over m yields a periodic function of $n\tau$, of unit period. Hence it admits a Fourier series of the form

$$\begin{aligned} \sum_m \frac{1}{|m + n\tau|^{2s}} &= \sum_p e^{2i\pi p n \operatorname{Re} \tau} \int_0^1 dy e^{-2i\pi p y} \sum_m \frac{1}{[(m + y)^2 + n^2 \operatorname{Im} \tau^2]^s} \\ &= \sum_p \int_{-\infty}^{+\infty} dy e^{2i\pi p (n \operatorname{Re} \tau - y)} \frac{1}{[y^2 + n^2 \operatorname{Im} \tau^2]^s} \\ &= \frac{1}{\Gamma(s)} \sum_p \int_{-\infty}^{+\infty} dy e^{2i\pi p (n \operatorname{Re} \tau - y)} \int_0^\infty dt t^{s-1} e^{-t(y^2 + n^2 \operatorname{Im} \tau^2)} \\ &= \frac{\pi^{1/2}}{\Gamma(s)} \sum_p \int_0^\infty dt t^{s-3/2} e^{-[tn^2 \operatorname{Im} \tau^2 + \pi^2(p^2/t) - 2i\pi p n \operatorname{Re} \tau]} \end{aligned}$$

For $p = 0$ the integral reduces to $\Gamma(s - \frac{1}{2}) |n \operatorname{Im} \tau|^{1-2s}$. We isolate this term, sum over $n \neq 0$, and use the functional equation for

Riemann's ζ function

$$2\pi^{-\frac{1}{2}s}\Gamma\left(\frac{1}{2}s\right)\zeta(s) = 2\pi^{-\frac{1}{2}(1-s)}\Gamma\left(\frac{1}{2}(1-s)\right)\zeta(1-s) \\ = -2\left(\frac{1}{s} + \frac{1}{1-s}\right) + \int_0^1 \frac{dt}{t} \left(t^{-\frac{1}{2}s} + t^{-\frac{1}{2}(1-s)}\right) \sum'_m e^{-\pi m^2/t} \quad (224)$$

Changing t into $|\pi p/n\text{Im } \tau|t$, yields

$$\Gamma(s) \left| \frac{2\pi\omega_1}{A} \right|^{2s} \left(\frac{\text{Im } \tau}{\pi} \right)^{s-\frac{1}{2}} G(s) = 2 \left(\frac{\text{Im } \tau}{\pi} \right)^{s-\frac{1}{2}} \Gamma(s)\zeta(2s) \\ + 2 \left(\frac{\text{Im } \tau}{\pi} \right)^{\frac{1}{2}-s} \Gamma(1-s)\zeta(2-2s) \\ + \sqrt{\pi} \sum'_n \sum'_p e^{2i\pi p n \text{Re } \tau} \left| \frac{p}{n} \right|^{s-\frac{1}{2}} \int_0^\infty \frac{dt}{t} t^{s-\frac{1}{2}} e^{-\pi|pn|\text{Im } \tau(t+t^{-1})} \quad (225)$$

The last double sum is an entire even function in $s - \frac{1}{2}$, hence the right-hand side is even in $s \rightarrow 1-s$, giving the required analytic continuation of $G(s)$. Since for $x > 0$

$$\int_0^\infty \frac{dt}{t} t^{\pm\frac{1}{2}} \exp[-x(t+t^{-1})] = (\pi/x)^{\frac{1}{2}} e^{-2x}$$

and $\zeta(2) = \pi^2/6$, we find in the vicinity of $s = 0$

$$G(s) = -1 - 2s \ln \left| \frac{A}{\omega_1} \right| + \frac{1}{3}s\pi \text{Im } \tau \\ + s \sum'_n \sum'_p \frac{1}{|p|} e^{2i\pi p n \text{Re } \tau - 2\pi|pn|\text{Im } \tau} + O(s^2)$$

Finally

$$G(0) = -1 \quad G'(0) = -2 \ln \left\{ \left| \frac{A}{\omega_1} \right| \eta(q)\eta(\bar{q}) \right\} \quad (226)$$

Noting that $|\omega_1|/A^{1/2} = 1/(\text{Im } \tau)^{1/2}$, we find the precise form of the partition function Z_1 , from equation (222), as

$$Z_1 = \frac{1}{(\text{Im } \tau)^{\frac{1}{2}} \eta(\tau)\eta(\bar{\tau})} \quad (227)$$

which differs from (217) by the prefactor $(\text{Im } \tau)^{-\frac{1}{2}}$. The latter is essential to insure modular invariance, which was obvious at the start of the calculation, and its origin can be traced to the zero mode subtraction.

We deduce from equation (227) that $|\eta(\tau)|(\text{Im } \tau)^{\frac{1}{2}}$ is modular invariant. To be precise, under the two generators $\tau \rightarrow \tau + 1$ and $\tau \rightarrow -\tau^{-1}$ of the modular group, the transformation law of $\eta(\tau)$ is

$$\eta(\tau + 1) = e^{2i\pi/24} \eta(\tau) \\ \eta(-1/\tau) = (\tau e^{-i\pi/2})^{\frac{1}{2}} \eta(\tau) \quad (228)$$

Under a general transformation $\tau \rightarrow \tau' = (a\tau + b)/(c\tau + d)$ we therefore have

$$\eta(\tau') = \epsilon_A (c\tau + d)^{\frac{1}{2}} \eta(\tau) \quad (229)$$

where ϵ_A is a 24th root of unity. The number theoretic determination of these phases, not required here, is due to Jordan and Dedekind.

For the free field φ , the two-point function can easily be computed on the torus. It should satisfy

$$-\Delta \langle \varphi(x)\varphi(0) \rangle = 2\pi(\delta(x) - 1/A) \quad (230)$$

The extra factor 2π on the r.h.s. is in agreement with the normalization at short distance (equation (162)), while the term $1/A$ represents the zero mode subtraction, making $-\Delta$ invertible in the orthogonal subspace. Both the correlation and the δ -function are understood to be doubly periodic. The solution to equation (230) is expressed in terms of one of the θ -functions (Appendix A) which we normalize as

$$F(z) = 2i\pi y^{-\frac{1}{2}} \frac{\theta_1(z, \tau)}{\theta'_1(0, \tau)} \\ = P(q)^{-2} (1 - y^{-1}) \prod_1^\infty (1 - yq^n)(1 - y^{-1}q^n) \quad (231) \\ = P(q)^{-3} \sum_{-\infty}^{+\infty} (-y)^n q^{\frac{1}{2}n(n+1)}$$

where, as before, $q = \exp(2i\pi\tau)$ and $y = \exp(2i\pi z/\omega_1)$, and as usual $P(q) = \prod_1^\infty (1 - q^n)$. One verifies that

$$\Gamma_{12} = \exp[-2\langle \varphi(z_1, \bar{z}_1)\varphi(z_2, \bar{z}_2) \rangle] \\ = \exp\left(-2\pi \left[\frac{(\text{Im } z_{12}/\omega_1)^2}{\text{Im } \tau} + \text{Im } z_{12}/\omega_1 \right]\right) \left| \frac{\omega_1}{2\pi} F(z_{12}) \right|^2 \quad (232)$$

One checks that Γ_{12} is doubly periodic, symmetric in the interchange $1 \leftrightarrow 2$, and behaves at short distance as

$$\Gamma_{12} \sim z_{12} \bar{z}_{12} \quad (233)$$

For later use, we look at Γ_{12} when z_{12} has rational coordinates in the ω_1, ω_2 basis. Let N, k and ℓ be integers, then

$$\exp -2 \left\langle \varphi(0) \varphi \left(\frac{k\omega_1 + \ell\omega_2}{N} \right) \right\rangle = \frac{A}{(2\pi)^2} Z_1(\tau, \bar{\tau})^2 \left| D_{\frac{k}{N}, \frac{\ell}{N}}(\tau) \right|^2 \quad (234)$$

with

$$D_{k/N, \ell/N}(\tau) = q^{-\frac{1}{12}(6\ell(N-\ell)/N^2 - 1)} \times \prod_{n=0}^{\infty} \left(1 - e^{2i\pi k/N} q^{n+\ell/N} \right) \left(1 - e^{-2i\pi k/N} q^{n+(N-\ell)/N} \right) \quad (235)$$

By construction $D_{k/N, \ell/N}(\tau)$ vanishes when both k and ℓ are multiples of N . Check that it satisfies the properties

$$\begin{aligned} \left| D_{k/N, \ell/N}(\tau) \right| &= \left| D_{(N-k)/N, (N-\ell)/N}(\tau) \right| \\ &= \left| D_{k/N, (N-k)/N}(-1/\tau) \right| \end{aligned} \quad (236)$$

$$D_{k/N, \ell/N}(\tau + 1) = e^{-\frac{1}{12}i\pi[6\ell(N-\ell)/N^2 - 1]} D_{(k+\ell)/N, \ell/N}(\tau) \quad (237)$$

$$\begin{aligned} D_{k/N, \ell/N+p}(\tau) &= (-1)^p e^{-2i\pi kp/N} D_{k/N, \ell/N}(\tau) \\ D_{k/N+p, \ell/N}(\tau) &= D_{k/N, \ell/N}(\tau) \end{aligned} \quad (238)$$

The D -functions admit of the following alternative interpretation. Let us return to the computation of the free field partition function, but, instead of taking the field doubly periodic, let us require that, after a cycle around ω_1 or ω_2 , it is multiplied by a phase (clearly the field is then complex). Precisely for λ, μ integers

$$\varphi(x + \lambda\omega_1 + \mu\omega_2) = e^{2i\pi(k\mu - \ell\lambda)/N} \varphi(x) \quad (239)$$

In complex notation, the proper modes correspond to eigenvalues of the form

$$\begin{aligned} & - \left(n_2 + \frac{\ell}{N} \right) k^1 + \left(n_1 + \frac{k}{N} \right) k^2 \\ &= \frac{i}{A} \left\{ \left(n_1 + \frac{k}{N} \right) \omega_1 + \left(n_2 + \frac{\ell}{N} \right) \omega_2 \right\} \end{aligned} \quad (240)$$

We do not require zero mode subtraction when k or ℓ (or both) are different from a multiple of N . Finally, instead of computing the square root of the inverse determinant, we simply compute the determinant. Therefore, with

$$G_{k/N, \ell/N}(s) = \left(\frac{A}{2\pi |\omega_1|} \right)^{2s} \sum_{m,n} \frac{1}{|m + n\tau + (k + \ell\tau)/N|^{2s}} \quad (241)$$

we find after analytic continuation

$$\exp -\frac{d}{ds} G_{k/N, \ell/N}(s) \Big|_{s=0} = \left| D_{k/N, \ell/N}(\tau) \right|^2 \quad (242)$$

Thus $D_{k/N, \ell/N}(\tau)$ appears as a (renormalized) determinant over "twisted" modes.

i) Verify equation (242).

ii) In the free field case, obtain the (constant) mean value of the energy-momentum on a torus. Performing a deformation $\delta x^\mu = \delta \epsilon^{\mu\nu} x^\nu$, with $\delta \epsilon$ an infinitesimal matrix, we have

$$\begin{aligned} \langle T_{\mu\nu}(x) \rangle \partial^\mu \delta x^\nu &= \langle T_{\mu\nu}(x) \rangle \delta \epsilon^{\mu\nu} \\ &= \langle T(z) \rangle [\delta \epsilon^{11} - \delta \epsilon^{22} + i(\delta \epsilon^{12} + \delta \epsilon^{21})] + \text{c.c.} \end{aligned}$$

with $\langle T(z) \rangle = \langle T \rangle$, independent of z ,

$$\delta \ln Z_1 = -\frac{A}{2\pi} \langle T \rangle [\delta \epsilon^{11} - \delta \epsilon^{22} + i(\delta \epsilon^{12} + \delta \epsilon^{21})] + \text{c.c.}$$

Inserting the expression (227) for Z_1 , we obtain

$$\langle T \rangle = -\frac{\pi}{2A} \frac{\bar{\omega}_1}{\omega_1} - \frac{2\pi i}{\omega_1^2} \frac{\eta'(\tau)}{\eta(\tau)} = -\frac{\pi}{2A} \frac{\bar{\omega}_1}{\omega_1} + \frac{1}{2} \sum'_{n,p} \frac{1}{(n\omega_1 + p\omega_2)^2} \quad (243)$$

where the last semiconvergent sum is understood as a double limit, summing first symmetrically on p , then on n .

9.3.3 Ising model

From the expression (171) for the action in the free Fermi case, and taking statistics into account, we derive that the partition function on a torus is the product of two Pfaffians $\text{Pf}(\partial)\text{Pf}(\bar{\partial})$, with $\partial, \bar{\partial}$ the Cauchy Riemann operators. This is also the square root of the determinant of the Laplacian, $(\det -\Delta)^{1/2}$. However boundary conditions have to be chosen appropriately. According to the results of chapter 2 (volume 1), the fields $\psi, \bar{\psi}$ have to be

antiperiodic along at least one of the generators ω_1, ω_2 . Denote by 0 or $\frac{1}{2}$ the periodic or antiperiodic boundary conditions, in agreement with the notation of the previous subsection. We have just obtained the form of the $(\det -\Delta)$ as $|D_{k/2, \ell/2}(\tau)|^2$, with $k, \ell = 0, 1$. From chapter 2, it follows that we should have

$$c = \frac{1}{2} \quad Z_{\frac{1}{2}} = \frac{1}{2} \left\{ |D_{\frac{1}{2}, \frac{1}{2}}(\tau)| + |D_{0, \frac{1}{2}}(\tau)| + |D_{\frac{1}{2}, 0}(\tau)| \right\} \quad (244)$$

while the term $|D_{0,0}|$ vanishes, reflecting the presence of a zero mode. This agrees with the results of Kaufmann, Ferdinand and Fisher. Moreover, should we insist in having one of the non vanishing terms in (244), then the other two follow from modular invariance. Indeed, according to equations (236) and (237), we obtain the action of $\tau \rightarrow \tau + 1$ and $\tau \rightarrow -\tau^{-1}$ as generating permutations on the three determinants

$$\begin{array}{ccc} & |D_{\frac{1}{2}, \frac{1}{2}}| & |D_{0, \frac{1}{2}}| & |D_{\frac{1}{2}, 0}| \\ \tau \rightarrow \tau + 1 & |D_{0, \frac{1}{2}}| & |D_{\frac{1}{2}, \frac{1}{2}}| & |D_{\frac{1}{2}, 0}| \\ \tau \rightarrow -\tau^{-1} & |D_{\frac{1}{2}, \frac{1}{2}}| & |D_{\frac{1}{2}, 0}| & |D_{0, \frac{1}{2}}| \end{array} \quad (245)$$

This can be also understood by following the effect of the indicated change of basis (ω_1, ω_2) on the corresponding boundary conditions. As a result, the combination (244) is modular invariant. The various boundary conditions are called "spin structures". Three of those contribute to the partition function.

It is now possible to make contact with the expansion in characters as indicated in equation (211). The functions $D_{\frac{1}{2}, k, \ell}$ are naturally squares, so that defining

$$\begin{aligned} D_{\frac{1}{2}, \frac{1}{2}}(\tau) &= q^{-\frac{1}{24}} d_{\frac{1}{2}, \frac{1}{2}}^2(\tau) & D_{0, \frac{1}{2}}(\tau) &= q^{-\frac{1}{24}} d_{0, \frac{1}{2}}^2(\tau) \\ D_{\frac{1}{2}, 0}(\tau) &= 2q^{-\frac{1}{24}} d_{\frac{1}{2}, 0}^2(\tau) \end{aligned} \quad (246)$$

we find

$$\begin{aligned} d_{\frac{1}{2}, \frac{1}{2}}(\tau) &= \prod_0^\infty (1 + q^{n+\frac{1}{2}}) & d_{0, \frac{1}{2}}(\tau) &= \prod_0^\infty (1 - q^{n+\frac{1}{2}}) \\ d_{\frac{1}{2}, 0}(\tau) &= q^{\frac{1}{48}} \prod_1^\infty (1 + q^n) \end{aligned} \quad (247)$$

satisfying

$$\begin{aligned} d_{\frac{1}{2}, \frac{1}{2}}(\tau) d_{0, \frac{1}{2}}(\tau) d_{\frac{1}{2}, 0}(\tau) &= q^{\frac{1}{16}} \\ d_{\frac{1}{2}, \frac{1}{2}}^8(\tau) &= d_{0, \frac{1}{2}}^8(\tau) + 16d_{\frac{1}{2}, 0}^8(\tau) \end{aligned} \quad (248)$$

One recognizes that the characters $\chi_{c,h}$, with $c = \frac{1}{2}$ and $h = 0, \frac{1}{16}$, are expressed as

$$\begin{aligned} q^{\frac{1}{48}} \chi_{\frac{1}{2}, 0}(\tau) &= \frac{1}{2} \left\{ d_{\frac{1}{2}, \frac{1}{2}}(\tau) + d_{0, \frac{1}{2}}(\tau) \right\} \\ &= \frac{1}{2} \left\{ \prod_0^\infty (1 + q^{n+\frac{1}{2}}) + \prod_0^\infty (1 - q^{n+\frac{1}{2}}) \right\} \end{aligned} \quad (249a)$$

$$= \frac{1}{P(q)} \sum_{k=-\infty}^{+\infty} \left(q^{[(24k+1)^2-1]/48} - q^{[(24k+7)^2-1]/48} \right) \underset{q \rightarrow 0}{\sim} 1$$

$$\begin{aligned} q^{\frac{1}{48}} \chi_{\frac{1}{2}, \frac{1}{16}}(\tau) &= \frac{1}{2} \left\{ d_{\frac{1}{2}, \frac{1}{2}}(\tau) - d_{0, \frac{1}{2}}(\tau) \right\} \\ &= \frac{1}{2} \left\{ \prod_0^\infty (1 + q^{n+\frac{1}{2}}) - \prod_0^\infty (1 - q^{n+\frac{1}{2}}) \right\} \\ &= \frac{1}{P(q)} \sum_{k=-\infty}^{+\infty} \left(q^{[(24k+5)^2-1]/48} - q^{[(24k+11)^2-1]/48} \right) \underset{q \rightarrow 0}{\sim} q^{\frac{1}{16}} \end{aligned} \quad (249b)$$

$$\begin{aligned} q^{\frac{1}{48}} \chi_{\frac{1}{2}, \frac{1}{16}}(\tau) &= d_{\frac{1}{2}, 0}(\tau) = q^{\frac{1}{48}} \prod_1^\infty (1 + q^n) \\ &= \frac{1}{P(q)} \sum_{k=-\infty}^{+\infty} \left(q^{[(24k-2)^2-1]/48} - q^{[(24k+10)^2-1]/48} \right) \underset{q \rightarrow 0}{\sim} q^{\frac{1}{16}} \end{aligned} \quad (249c)$$

Consequently for the Ising model

$$\begin{aligned} Z_{\frac{1}{2}}(q, \bar{q}) &= (q\bar{q})^{-\frac{1}{24}} \left\{ \frac{1}{2} |d_{\frac{1}{2}, \frac{1}{2}}(\tau)|^2 + \frac{1}{2} |d_{0, \frac{1}{2}}(\tau)|^2 + |d_{\frac{1}{2}, 0}(\tau)|^2 \right\} \\ &= |\chi_{\frac{1}{2}, 0}(\tau)|^2 + |\chi_{\frac{1}{2}, \frac{1}{16}}(\tau)|^2 + |\chi_{\frac{1}{2}, \frac{1}{16}}(\tau)|^2 \end{aligned} \quad (250)$$

in agreement with the general discussion. This formula exhibits the expected operator content of the model, with three primary operators namely the identity, the energy density and the spin. The direct derivation confirms the value of the central charge, $c = \frac{1}{2}$, as well as the conformal weights $(0, 0)$, $(\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{16}, \frac{1}{16})$ respectively. Modular invariance is also clear as we have seen above.

One can also compute various correlation on a torus. First let us consider fermionic fields. As we recall, in the plane

$$\langle \psi(z_1) \psi(z_2) \rangle_{\text{plane}} = \frac{1}{z_{12}} \quad (251)$$

with the complex conjugate expression for the $\bar{\psi}$ fields. Assuming ψ of weights $(\frac{1}{2}, 0)$, it follows that on a periodic strip, with $z_{\text{plane}} = \exp(2i\pi/L)z_{\text{strip}}$

$$\langle \psi(z_1) \psi(z_2) \rangle_{\text{strip}(\frac{1}{2})} = \frac{\pi/L}{\sin(\pi z_{12}/L)} \quad (252)$$

This is antisymmetric in the interchange $z_1 \leftrightarrow z_2$, and antiperiodic $\psi(z+L) = -\psi(z)$ (hence the index $\frac{1}{2}$). There exists however an other possibility, with a periodic field $\psi(z+L) = \psi(z)$ (index 0), namely

$$\langle \psi(z_1) \psi(z_2) \rangle_{\text{strip}(0)} = \frac{\pi/L}{\tan(\pi z_{12}/L)} \quad (253)$$

On a torus, similar correlations satisfying

$$\bar{\partial}_1 \langle \psi(z_1) \psi(z_2) \rangle = \pi \delta^2(1, 2) \quad (254)$$

are distinguished by a double set of periodic or antiperiodic conditions forbidding the existence of zero modes. With $k, \ell = 0, 1$ not both 0, we write

$$\langle \psi(z_1) \psi(z_2) \rangle_{\frac{1}{2}k, \frac{1}{2}\ell} = \frac{2i\pi}{\omega_1} \frac{S_{\frac{1}{2}k, \frac{1}{2}\ell}(z_{12}, \tau)}{\sigma_{\frac{1}{2}k, \frac{1}{2}\ell}(\tau)} \quad (255)$$

with the first index $\frac{1}{2}k$ referring to the behaviour in the direction ω_2 , the second $\frac{1}{2}\ell$ in the direction ω_1 . Using the notation $F(z)$ introduced in (231), and $y = \exp(2i\pi z/\omega_1)$, it is not difficult to check that the required expressions are

$$S_{\frac{1}{2}, 0}(z, \tau) = \frac{F(z - \frac{1}{2}\omega_1)}{F(z)} = \frac{y+1}{y-1} \prod_1^\infty \frac{(1+yq^n)(1+y^{-1}q^n)}{(1-yq^n)(1-y^{-1}q^n)}$$

$$\begin{aligned} S_{0, \frac{1}{2}}(z, \tau) &= e^{-i\pi z/\omega_1} \frac{F(z - \frac{1}{2}\omega_2)}{F(z)} \\ &= \frac{1}{y^{\frac{1}{2}} - y^{-\frac{1}{2}}} \prod_1^\infty \frac{(1-yq^{n-\frac{1}{2}})(1-y^{-1}q^{n-\frac{1}{2}})}{(1-yq^n)(1-y^{-1}q^n)} \end{aligned}$$

$$S_{\frac{1}{2}, \frac{1}{2}}(z, \tau) = e^{-i\pi z/\omega_1} \frac{F(z - \frac{1}{2}(\omega_1 + \omega_2))}{F(z)}$$

$$= \frac{1}{y^{\frac{1}{2}} - y^{-\frac{1}{2}}} \prod_1^\infty \frac{(1+yq^{n-\frac{1}{2}})(1+y^{-1}q^{n-\frac{1}{2}})}{(1-yq^n)(1-y^{-1}q^n)} \quad (256)$$

The normalization factors are

$$\begin{aligned} \sigma_{\frac{1}{2}, 0}(\tau) &= 2 \prod_1^\infty \left(\frac{1+q^n}{1-q^n} \right)^2 \\ \sigma_{0, \frac{1}{2}}(\tau) &= \prod_1^\infty \left(\frac{1-q^{n-\frac{1}{2}}}{1-q^n} \right)^2 \\ \sigma_{\frac{1}{2}, \frac{1}{2}}(\tau) &= \prod_1^\infty \left(\frac{1+q^{n-\frac{1}{2}}}{1-q^n} \right)^2 \end{aligned} \quad (257)$$

These expressions reduce correctly to (252) or (253) in the limit of a strip. The latter are referred to as Neveu-Schwarz and Ramond boundary conditions respectively in the context of string field theory. From equation (255) follows the expression of the energy-energy correlation on a torus, a weighted sum over three terms

$$\langle \epsilon(z_1, \bar{z}_1) \epsilon(z_2, \bar{z}_2) \rangle = \frac{1}{2} Z_{\frac{1}{2}}^{-1} \sum'_{k, \ell} \left| D_{\frac{1}{2}k, \frac{1}{2}\ell}(q) \right| \left| \langle \psi(z_1) \psi(z_2) \rangle_{\frac{1}{2}k, \frac{1}{2}\ell} \right|^2 \quad (258)$$

The prime in the sum means that the (0,0) sector is omitted. While it is true that $|D_{0,0}|$ vanishes, this sector may contribute to other quantities. For instance, we shall see in section 3.8 that $\langle \epsilon(z, \bar{z}) \rangle$ is a nonvanishing constant, which arises only from this sector

$$\langle \epsilon \rangle = \frac{\pi |\eta(\tau)|^2}{Z_{\frac{1}{2}}} \quad (259)$$

Another example where this sector contributes is afforded by the spin-spin correlations (Di Francesco, Saleur and Zuber). Using the θ -functions notation (Appendix A),

$$\begin{aligned} |D_{\frac{1}{2}, \frac{1}{2}}(\tau)| &= |\theta_3(0, \tau)/\eta(\tau)| \\ |D_{0, \frac{1}{2}}(\tau)| &= |\theta_4(0, \tau)/\eta(\tau)| \\ |D_{\frac{1}{2}, 0}(\tau)| &= |\theta_2(0, \tau)/\eta(\tau)| \end{aligned} \quad (260a)$$

we have

$$Z_{\frac{1}{2}} = \frac{1}{2} \sum_{\nu=2}^4 |\theta_\nu(0, \tau)/\eta(\tau)|$$

and, up to an overall factor,

$$\langle \sigma(z_1, \bar{z}_1) \sigma(z_2, \bar{z}_2) \rangle = \frac{\sum_{\nu=1}^4 |\theta_\nu(\frac{1}{2}z_{12}, \tau)| \left| \frac{\theta'_1(0, \tau)}{\theta_1(z_{12}, \tau)} \right|^{\frac{1}{2}}}{\sum_{\nu=2}^4 |\theta_\nu(0, \tau)|} \quad (261)$$

$$\underset{z_{12} \rightarrow 0}{\sim} \frac{1}{(z_{12} \bar{z}_{12})^{\frac{1}{2}}}$$

with $\theta_1(z, \tau) \underset{z \rightarrow 0}{\sim} \theta'_1(0, \tau)z$, $\theta'_1(0, \tau) = 2\pi\eta(\tau)^3$.

The expression (261) behaves as expected at short distance. One can also convince oneself that the correlation is doubly periodic. The interesting point is that it is a sum of contributions arising from the four sectors in the form $\sum Z_\nu \langle \sigma \sigma \rangle_\nu / \sum Z_\nu$, where, in the numerator, the doubly periodic sector ($\nu = 1$) gives a nonvanishing contribution in spite of the fact that $Z_\nu|_{\nu=1} = 0$. Each of the partial quantities $Z_\nu \langle \sigma \sigma \rangle_\nu$ relative to a given spin structure satisfies a generalization of the Belavin, Polyakov, Zamolodchikov equations obtained by Eguchi and Ooguri in the form, valid for any operator A degenerate at level 2;

$$\left\{ \frac{3}{2(2h+1)} \partial_z^2 - 2\eta_1 z \partial_z - \sum_{i=1}^n \left(h_i [p(z - z_i) + 2\eta_1] + [\zeta(z - z_i) + 2\eta_1 z_i] \frac{\partial}{\partial z_i} \right) - 2i\pi \frac{\partial}{\partial \tau} - 2h\eta_1 \right\} Z \langle A(z, \bar{z}) A_1(z_1, \bar{z}_1) \cdots A_n(z_n, \bar{z}_n) \rangle = 0 \quad (262)$$

These equations involve Weierstrass's p and ζ functions with a double and simple pole at the origin as well as the η_1 constant

$$\begin{aligned} \zeta(z, \tau) &= \frac{\theta'_1(z, \tau)}{\theta_1(z, \tau)} + 2\eta_1(\tau)z \\ p(z, \tau) &= -\zeta'(z, \tau) \\ \eta_1(\tau) &= -\frac{1}{6} \frac{\theta''_1(0, \tau)}{\theta'_1(0, \tau)} = -2i\pi \frac{\partial}{\partial \tau} \ln \eta(\tau) \end{aligned} \quad (263)$$

Beware that $\zeta(z, \tau)$ is not doubly periodic and the constant η_1 already appeared in equation (243). Using the same notations, the

$\epsilon\epsilon$ correlation (258) reads

$$\langle \epsilon(z_1, \bar{z}_1) \epsilon(z_2, \bar{z}_2) \rangle = \left| \frac{\theta'_1(0, \tau)}{\theta_1(z_{12}, \tau)} \right|^2 \frac{\sum_{\nu=2}^4 |\theta_\nu(z_{12}, \tau)|^2 / |\theta_\nu(0, \tau)|}{\sum_{\nu=2}^4 |\theta_\nu(0, \tau)|} \quad (264)$$

Higher order correlations on a torus can also be obtained in closed form.

9.3.4 The A-D-E classification of minimal models

We return to the discussion of section 3.1, in particular to equation (211), where the partition function is expanded as a sesquilinear form in Virasoro characters with integral nonnegative coefficients. Such an expansion is intimately related to a Hamiltonian formulation, with a preferred choice of time and space directions. It was claimed that the latter is in fact irrelevant, and that under a modular transformation, the partition function should be invariant, leading to constraints on the operator content of a given model. We found in the previous subsection, equation (250), a typical example involving such a sum over three terms with unit coefficients. In a sense, the Gaussian partition function (equation (227)) is also a limiting case of this situation. In this section, we report the results of imposing the constraint of modular invariance in the general case where the central charge assumes a rational value $c = 1 - 6(p - p')^2 / pp'$, with the special, and important, unitary subcase such that p and p' are consecutive integers. Remarkably, a complete classification can be described in terms related to the Cartan-Killing classification of simple Lie algebras (to be specific, a subsequence called "simply laced"). This same A-D-E classification, where A stands for the Lie algebras of unitary groups, D for orthogonal groups in even dimension and E for the three exceptional Lie algebras E_6 , E_7 , E_8 , occurs in several distinct instances, at first seemingly unrelated. An other famous case is the one of finite subgroups of three-dimensional rotations (up to conjugacy).

The fact that we can obtain a complete description of all minimal models is, to some extent, quite surprising and shows how deep is the present approach to critical models in two dimensions. Although all its implications have not yet been worked out, we know from the works on integrable models (Baxter, Andrews and Forrester, Pasquier) that one can construct lattice models which

correspond to the predicted critical behaviour. On the other hand, the relation with integrable models is not yet completely clear. At any rate, it is of some importance to be able to obtain, even in a restricted range, a complete description of critical universality classes.

To carry out the previous program, we have to describe the behaviour of Virasoro characters under the action of the modular group. The latter is generated by the two basic transformations

$$\begin{array}{ll} T & \tau \rightarrow \tau + 1 \\ S & \tau \rightarrow -\tau^{-1} \end{array} \quad (265a)$$

satisfying

$$S^2 = (ST)^3 \quad (265b)$$

In section 1.8, it was found convenient to label the representations by an integer λ defined modulo $N = 2pp'$, with a compact form of the characters given in equation (158). We also recall that under modular transformations, Dedekind's function, which appears in the denominator of the character formula, transforms as indicated in (228). Using Poisson's formula

$$\sum_{n=-\infty}^{+\infty} f(n) = \sum_{p=-\infty}^{+\infty} \int dx f(x) e^{2i\pi px} \quad (266)$$

it is readily found that

$$T \quad \chi_\lambda(\tau + 1) = e^{2i\pi(\lambda^2/2N - \frac{1}{24})} \chi_\lambda(\tau) \quad (267a)$$

$$S \quad \chi_\lambda(-1/\tau) = \frac{1}{\sqrt{N}} \sum_{\lambda' \in \mathbb{Z}/N\mathbb{Z}} e^{2i\pi\lambda\lambda'/N} \chi_{\lambda'}(\tau) \quad (267b)$$

In other words the action of T is diagonal, being, up to a phase, the multiplication by $\exp(2i\pi\lambda^2/2N)$, while the action of S is nothing but the finite Fourier transform over integers modulo N . The latter has a square equal to $\delta_{\lambda, -\lambda' \bmod N}$, but the equality $\chi_\lambda = \chi_{-\lambda}$ restores the property $S^2 = I$. Both transformations in (267) are compatible with the antisymmetry property $\chi_\lambda = -\chi_{\omega_0\lambda}$. We recall that ω_0 is given by $\omega_0 = r_0p + s_0p' \bmod N$, with r_0, s_0 such that $r_0p - s_0p' = 1$, an identity expressing that the integers p and p' are coprime. We could of course use the symmetries of χ_λ to restrict the sum on the r.h.s. of (267b). This, however, would hide the relation between the Virasoro characters and an other set

of similar functions arising in the study of the representations of the Kac-Moody algebra $A_1^{(1)}$, namely the algebra of local $SU(2)$ currents. An interesting class of highest weight representations of the corresponding infinite Lie algebra is also characterized by an anomaly (analogous to the central charge) called, in this context, the level, and denoted by k , an integer taking the values $0, 1, \dots$. These representations can be understood as a superposition of ordinary representations of $SU(2)$, the one with lowest spin being nondegenerate with a dimension $\lambda = 2j_{\min} + 1$, and $2j_{\min}$ ranging over integers from 0 to k . We present these matters in appendix C. The reason for introducing this subject here is the close relationship with the Virasoro characters and the quest for similar modular invariant partition functions.

Let us denote by $\chi_\lambda^{aff}(\tau)$ the $A_1^{(1)}$ characters, defined in appendix C. The previous Virasoro characters might be called by contrast, conformal ones. Specifically, for a level k representation of $A_1^{(1)}$, we set $N = 2(k + 2)$, again an even integer, and the character reads

$$\chi_\lambda^{aff}(\tau) = \sum_{n=-\infty}^{+\infty} \frac{(Nn + \lambda) q^{(Nn + \lambda)^2/2N}}{\eta^3(\tau)} \quad (268)$$

We note the great similarity with the case of Virasoro characters (equation (158)), and observe that χ_λ^{aff} extends as a periodic (period N) odd function of λ

$$\chi_\lambda^{aff}(\tau) = \chi_{\lambda+N}^{aff}(\tau) = -\chi_{-\lambda}^{aff}(\tau) \quad (269)$$

In the present case the role of the involution $\lambda \rightarrow \omega_0\lambda$ is simply played by the symmetry $\lambda \rightarrow -\lambda$. The similarity is further strengthened by the modular transformation properties, to be compared with (267)

$$\chi_\lambda^{aff}(\tau + 1) = e^{2i\pi(\lambda^2/2N - \frac{1}{24})} \chi_\lambda^{aff}(\tau) \quad (270a)$$

$$\chi_\lambda^{aff}(-1/\tau) = \frac{-i}{\sqrt{N}} \sum_{\lambda' \in \mathbb{Z}/N\mathbb{Z}} e^{2i\pi\lambda\lambda'/N} \chi_{\lambda'}^{aff}(\tau) \quad (270b)$$

As suggested by Gepner and Witten, one can set forth a similar problem of classifying partition functions of the type (211) with affine characters replacing Virasoro characters, and coefficients restricted by analogous conditions, provided the sum over λ runs over a fundamental domain $1 \leq \lambda \leq \frac{1}{2}N - 1$ since $\chi_0^{aff} = \chi_{\frac{1}{2}N}^{aff} = 0$.

In both cases, the modular invariant solutions can be found in two steps. In the common first step, one simply looks for $N \times N$ matrices with arbitrary entries, call them $\mathcal{N}$, which commute with the two matrices

$$\begin{aligned} T_{\lambda, \lambda'} &= e^{2i\pi\lambda^2/2N} \delta_{\lambda, \lambda'} \\ S_{\lambda, \lambda'} &= \frac{1}{\sqrt{N}} e^{2i\pi\lambda\lambda'/N} \end{aligned} \quad (271)$$

where λ and λ' as well as the labels of $\mathcal{N}$ run over integers modulo N .

It is found that such matrices $\mathcal{N}$ can be written as combinations of a linearly independent set defined as follows. Let δ be any positive divisor of $\frac{1}{2}N$, including unity, and α the greatest common divisor of δ and $\bar{\delta} = N/2\delta$, in compact notation $\alpha = (\delta, \bar{\delta})$. Clearly α^2 is a divisor of $\frac{1}{2}N$. Define the $N \times N$ matrix $(\Omega_\delta)_{\lambda, \lambda'}$ to have vanishing matrix elements unless α divides both λ and λ' , in which case

$$(\Omega_\delta)_{\lambda, \lambda'} = \sum_{\xi \in \mathbb{Z}/\alpha\mathbb{Z}} \delta_{\lambda', \omega\lambda + \xi N/\alpha} \quad (272)$$

In this definition, the integer $\omega \bmod N/\alpha^2$ is obtained as follows. Since δ/α and $\bar{\delta}/\alpha$ are coprime, one can find two integers ρ and σ such that $\rho\bar{\delta}/\alpha - \sigma\delta/\alpha = 1$, then $\omega = \rho\bar{\delta}/\alpha + \sigma\delta/\alpha \pmod{N/\alpha^2}$. One checks that $\omega^2 = 1 \pmod{2N/\alpha^2}$, and this makes sense since in (272) it is required that $\lambda'/\alpha = \omega\lambda/\alpha \pmod{N/\alpha^2}$.

It is not difficult to see that Ω_δ commutes with both T and S . It is slightly more involved to show, as did Gepner and Qiu, that any matrix $\mathcal{N}$ with this property is a linear combination of these Ω_δ 's and that the latter are linearly independent. We remark that if $\delta = \frac{1}{2}N$, $\bar{\delta} = 1$, $\alpha = 1$, $\omega = 1$, and $\Omega_{\frac{1}{2}N} = I$.

The difficult task is to find the correct superposition of Ω_δ 's such that $\sum \chi_\lambda^*(\tau) (\gamma_\delta \Omega_\delta)_{\lambda, \lambda'} \chi_{\lambda'}(\tau)$, which is modular invariant, reduces, when summed over the fundamental (Virasoro or affine) characters, to a combination with nonnegative integral coefficients. The normalization in the affine case requires that the coefficient of $\chi_1^{aff}(\tau) \chi_1^{aff}(\tau)$ be unity. In the Virasoro case, the coefficient of the invariant state, corresponding to $h = \bar{h} = 0$, should also be unity. Using the λ -notation, this corresponds to a unit coefficient for $\chi_{p-p'}^*(\tau) \chi_{p-p'}(\tau)$. The difficulty in both cases stems from an

antisymmetry of the characters: $\lambda \rightarrow \omega_0\lambda$ or $\lambda \rightarrow -\lambda$ respectively.

Let us first describe the solution for the affine partition functions, where we write the required combination as

$$Z_{affine} = \frac{1}{2} \sum_{\lambda, \lambda' \in \mathbb{Z}/N\mathbb{Z}} \chi_\lambda^{aff*}(\tau) \left(\sum \gamma_\delta \Omega_\delta \right)_{\lambda, \lambda'} \chi_{\lambda'}^{aff}(\tau) \quad (273)$$

The coefficient $\frac{1}{2}$ takes into account the fact that, due to the antisymmetry $\chi_{-\lambda} = -\chi_\lambda$, each fundamental product $\chi_\lambda^* \chi_{\lambda'}$ is counted twice. Two infinite series and three exceptional solutions, fulfil all criteria. For conciseness we write, using as a label $n = \frac{1}{2}N$, the combination $\sum \gamma_\delta \Omega_\delta$ appearing in (273), and recall that Ω_n is the unit matrix

$n \geq 2$	Ω_n	(A_{n-1})	
$n \text{ even} \geq 6$	$\Omega_n + \Omega_2$	$(D_{n/2+1})$	
$n = 12$	$\Omega_{12} + \Omega_3 + \Omega_2$	(E_6)	(274)
$n = 18$	$\Omega_{18} + \Omega_3 + \Omega_2$	(E_7)	
$n = 30$	$\Omega_{30} + \Omega_5 + \Omega_3 + \Omega_2$	(E_8)	

As a consequence when n is odd the solution is unique. When n is even, there exist in general two distinct solutions, and three when $n = 12, 18$ or 30 . The group theoretic meaning of this result is more clearly appreciated when we give the explicit form of (273)–(274) in table I.

To correlate this result with the A – D – E classification of simply laced, simple Lie algebras (appendix C), we note that the diagonal entries (0, 1, or 2) in the matrix (274), as shown in the table, are related to the degrees of the fundamental Casimir invariant polynomials pertaining to the corresponding Lie algebra. Recall that the enveloping algebra of a Lie algebra admits a subset of invariant polynomials. The latter can in turn be written as polynomials in a set of r of them, r being the rank of the Lie algebra. If we subtract 1 from the degrees of these fundamental invariant polynomials (the first one is always quadratic), then the above diagonal entries indicate how many polynomials of this degree are found in this set. For instance, corresponding to E_8 , we find in the table that the Casimir invariants have degrees 2, 8, 12, 14, 18, 20, 24 and 30. These degrees are sometimes called Coxeter

Table I. Invariant partition functions for the Kac-Moody algebra $A_1^{(1)}$.

$n \geq 2$	A_{n-1}	$\sum_{\lambda=1}^{n-1} x_\lambda ^2$
$n = 4\rho + 2$ $\rho \geq 1$	$D_{2\rho+2}$	$\sum_{\lambda_{\text{odd}}=1}^{2\rho-1} x_\lambda + x_{4\rho+2-\lambda} ^2 + 2 x_{2\rho+1} ^2$
$n = 4\rho$ $\rho \geq 2$	$D_{2\rho+1}$	$\sum_{\lambda_{\text{odd}}=1}^{4\rho-1} x_\lambda ^2 + x_{2\rho} ^2 + \sum_{\lambda_{\text{even}}=2}^{2\rho-2} (x_\lambda x_{4\rho-\lambda} + x_{4\rho-\lambda} x_\lambda)$
$n = 12$	E_6	$ x_1 + x_7 ^2 + x_4 + x_8 ^2 + x_5 + x_{11} ^2$
$n = 18$	E_7	$ x_1 + x_{17} ^2 + x_5 + x_{13} ^2 + x_7 + x_{11} ^2 + x_9 ^2 + [(x_3 + x_{15})^2 x_9 + x_9^2 (x_3 + x_{15})]$
$n = 30$	E_8	$ x_1 + x_{11} + x_{19} + x_{29} ^2 + x_7 + x_{13} + x_{17} + x_{23} ^2$

numbers. This correspondence is the reason of the labelling. Its origin is not at present fully understood.

It would seem that we have followed a side track in classifying these affine invariants, were it not for the fact that the solution of our original problem with Virasoro characters is simply related to the above one. Namely, recall that in this case $N = 2pp'$. Then, at the price of an irrelevant doubling of the original space, it turns out that we can think of Virasoro characters as carrying labels in a tensor space of dimension $2p' \times 2p$. Elements of the commutant of T and S can also be thought of as products $\Omega_{\delta'} \otimes \Omega_{\delta}$ and the corresponding integral positive partition functions are given by a pair of combinations chosen among those occurring in equation (274), with the role of n played successively by p' and p . Moreover, since p and p' are coprime, one of them at least is necessarily odd, the corresponding invariant being then of the A type. Consequently, and this is the main result of this section, one finds two infinite series and three exceptional subseries of minimal models consistent with the requirements of conformal and modular invariance. The partition functions are displayed in table II, where it is assumed for definiteness that p is odd. The particular unitary case will then mean that if m is odd $p \rightarrow m$, $p' = m + 1$ and if m is even $p \rightarrow m + 1$, $p' = m$. Thus, we find two infinite series and three pairs of exceptional models. The factor $\frac{1}{2}$ appearing in table II reflects the symmetry $\chi_{r,s} = \chi_{p'-r, p-s}$, when we return to the labelling in integers r, s ($\lambda = rp - sp'$) as used in the table, where each invariant is denoted by a pair of Lie algebras.

The analysis of each of the models occurring in the table is beyond the scope of the present discussion. We simply present a few remarks. First let us restrict our attention to the unitary case. In the $(A-A)$ series, the first candidate corresponds to $m = 3$, $p = 3$, $p' = 4$, thus in the present notation (A_3, A_2) . This of course is the Ising model. In general, in this $A-A$ series (called the principal series), the partition function reads

$$Z = \sum_{1 \leq s \leq r \leq m-1} |\chi_{r,s}(\tau)|^2 \quad (275)$$

and one obtains a generalization of the Ising model with only one scalar operator $A_{h,h}$ for each distinct entry in the Kac table. The next case, corresponding to $m = 4$, $c = \frac{7}{10}$, can be interpreted as the tricritical Ising model. In general, the unitary $A-A$ series

Table II. Invariant partition functions in terms of Virasoro characters.

$(A_{p'-1}, A_{p-1})$		$\frac{1}{2} \sum_{s=1}^{p-1} \sum_{r=1}^{p'-1} \chi_{r,s} ^2$
$(D_{2\rho+2}, A_{p-1})$	$p' = 4\rho + 2$ $\rho \geq 1$	$\frac{1}{2} \sum_{s=1}^{p-1} \left\{ \sum_{r_{\text{odd}}=1}^{2\rho-1} \chi_{r,s} + \chi_{4\rho+2-r,s} ^2 + 2 \chi_{2\rho+1,s} ^2 \right\}$
$(D_{2\rho+1}, A_{p-1})$	$p' = 4\rho$ $\rho \geq 2$	$\frac{1}{2} \sum_{s=1}^{p-1} \left\{ \sum_{r_{\text{odd}}=1}^{4\rho-1} \chi_{r,s} ^2 + \chi_{2\rho,s} ^2 + \sum_{r_{\text{even}}=2}^{2\rho-2} (\chi_{r,s}^* \chi_{4\rho-r,s} + \chi_{4\rho-r,s}^* \chi_{r,s}) \right\}$
(E_6, A_{p-1})	$p' = 12$	$\frac{1}{2} \sum_{s=1}^{p-1} \left\{ \chi_{1,s} + \chi_{7,s} ^2 + \chi_{4,s} + \chi_{8,s} ^2 + \chi_{5,s} + \chi_{11,s} ^2 \right\}$
(E_7, A_{p-1})	$p' = 18$	$\frac{1}{2} \sum_{s=1}^{p-1} \left\{ \chi_{1,s} + \chi_{17,s} ^2 + \chi_{5,s} + \chi_{13,s} ^2 + \chi_{7,s} + \chi_{11,s} ^2 + \chi_{9,s} ^2 + \left[(\chi_{3,s} + \chi_{15,s})^* \chi_{9,s} + \chi_{9,s}^* (\chi_{3,s} + \chi_{15,s}) \right] \right\}$
(E_8, A_{p-1})	$p' = 30$	$\frac{1}{2} \sum_{s=1}^{p-1} \left\{ \chi_{1,s} + \chi_{11,s} + \chi_{19,s} + \chi_{29,s} ^2 + \chi_{7,s} + \chi_{13,s} + \chi_{17,s} + \chi_{23,s} ^2 \right\}$

seems to describe the set of multicritical scalar models, of higher and higher degree (appendix A, chapter 5), with an effective Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}(\partial\varphi)^2 + \varphi^{2k}$$

The case $k = 2$ is the Ising model again ($m = 3$), and a general k corresponds to the value $m = k + 1$. The central charge $c = 1 - 6/(k+1)(k+2)$ converges to the free field case ($c = 1$) as $k \rightarrow \infty$, in agreement with intuition. According to Zamolodchikov, the primary field with weights $h = \bar{h} = h_{2,2} = \frac{3}{4}(k+1)(k+2)$ ($k \geq 2$) corresponds to the field φ . Higher powers: φ^n : have weights $h = \bar{h} = h_{n+1,n+1} = [(n+1)^2 - 1]/4(k+1)(k+2)$ for $n = 1, 2, \dots, k-1$, and $h = \bar{h} = h_{n-k+2,n-k+1} = [(n+3)^2 - 1]/4(k+1)(k+2)$ for $n = k, \dots, 2k-2$. The field: φ^{2k-1} : is related to $\partial\bar{\partial}\varphi$ through the equation of motion. As will be discussed in the next subsection, all these models are naturally related to the spontaneous breaking of a Z_2 symmetry.

When $m = 5$ and $c = \frac{4}{5}$, we find two possibilities. According to the principal series, all primary operators are scalars, while a second possibility belongs to the (A, D) series (call it the complementary series), namely $p = 5, p' = 6$, thus (A_4, D_4) . Using first the notation $\chi_{r,s}$, then the notation $\chi_{h,r}$, the corresponding partition function reads

$$\begin{aligned}
 (A_4, D_4) \quad Z &= |\chi_{1,1} + \chi_{4,1}|^2 + |\chi_{2,1} + \chi_{3,1}|^2 + 2|\chi_{3,3}|^2 + 2|\chi_{4,3}|^2 \\
 &\equiv |\chi_0 + \chi_3|^2 + |\chi_{2/5} + \chi_{7/5}|^2 + 2|\chi_{1/15}|^2 + 2|\chi_{2/3}|^2
 \end{aligned}
 \tag{276}$$

Assuming that the three state Potts model belongs to the value $\frac{4}{5}$ for the central charge, there are various ways to ascertain that (276) is the correct choice. One of them is to recognize that we need two operators with weights $\frac{1}{15}, \frac{1}{15}$, corresponding to the two $\sin\varphi$, and $\cos\varphi$ "spin operators". Another argument, to be described in the next subsection, shows that only (276) corresponds to the possibility of a Z_3 symmetry breaking, as should be the case. There is however an almost direct "microscopic" proof using a Coulomb gas description, which agrees with this choice, as further discussed in section 3.6. Assuming that equation (276) does correctly describe the content of the three state Potts model, we note the existence of operators with angular momentum ± 3 ,

±1. Except for the principal $A-A$ series, one always finds such tensorial fields.

Huse has observed that the $A-A$ principal series was realized at the critical point of the restricted solid on solid models solved by Andrews, Baxter and Forrester. We have already mentioned that Pasquier has generalized their construction to include all other unitary cases listed in table II. Of course, the iq^3 model or Lee Yang edge singularity (section 1.7) belongs also to the table. As we recall $c = -\frac{22}{5}$ and the partition function is in the (A, A) series

$$(A_4, A_1) \quad Z = |\chi_{h=0}|^2 + |\chi_{h=-\frac{1}{2}}|^2 \quad (277)$$

Proofs and comments on the results of this section will be found in the works quoted in the notes.

9.3.5 Frustrations and discrete symmetries

Assume that one of the theories described above is a critical limit of a lattice model, with nearest neighbour interactions between discrete variables $\sigma_{i,j}$ taking finitely many values. Assume furthermore that a finite group G operates globally on the variables in such a way that the action is invariant on an infinite lattice. When the latter is restricted to a finite range L, M , instead of considering periodic boundary conditions $\sigma_{L+1,j} = \sigma_{1,j}$ or $\sigma_{i,M+1} = \sigma_{i,1}$, as was implicit up to now, we can instead introduce twisted (or frustrated) boundary conditions in the form $\sigma_{L+1,j} = g_1 \sigma_{1,j}$, $\sigma_{i,M+1} = g_2 \sigma_{i,1}$, where $g_1, g_2 \in G$. Taking the continuous limit, we have therefore to envision frustrated partition functions Z_{g_1, g_2} which are now covariant under the modular group. To be specific, let us look for such circumstances where G is a cyclic group $G \cong \mathbb{Z}/k\mathbb{Z}$, and the discrete variable $\sigma_{i,j}$ gives rise to one of the primary fields with conformal weights $(h, \bar{h})$. Furthermore let $\mathbb{Z}/k\mathbb{Z}$ act through phases, in such a way that the corresponding field $u_{h, \bar{h}}$ has to fulfil, with n_1 and n_2 integers,

$$u_{h, \bar{h}}(z + n_1 \omega_1 + n_2 \omega_2) = \exp \left[\frac{2i\pi}{k} (n_1 k_1 + n_2 k_2) \right] u_{h, \bar{h}}(z) \quad (278)$$

According to the definition of h and $\bar{h}$, we have

$$u_{h, \bar{h}}(z + n_1 \omega_1) = \exp [2i\pi(h - \bar{h})n_1] u_{h, \bar{h}}(z) \quad (279)$$

meaning that $k_1/k \equiv h - \bar{h} \pmod{1}$. Taking k_1 prime to k , and using the values of h and $\bar{h}$ in the Kac table for a minimal model, we conclude that k must divide $4pp'$ or $4m(m+1)$ if we restrict ourselves to unitary models. This is a necessary condition for $\mathbb{Z}/k\mathbb{Z}$ to be a candidate symmetry group. Denote by Z_{k_1, k_2} the partition function corresponding to the boundary conditions (278). It is only invariant under a subgroup of the modular group which respect these conditions. In particular consider $Z_{k_1, 0}$. If we set $\tau' = (a\tau + b)/(c\tau + d)$, it is seen that invariance of $Z_{k_1, 0}$ requires $a = d = \pm 1 \pmod{k}$ and $b = 0 \pmod{k}$, provided we assume k_1 and k coprime and a reality condition $Z_{k_1, 0} = Z_{-k_1, 0}$. This defines a subgroup, call it $\Gamma^0(k)$, of the full modular group Γ . The twisted boundary condition in the direction ω_1 does not prevent us from thinking of $Z_{k_1, 0}$ as being the trace of the transfer matrix in a certain subspace of states which can be decomposed into irreducible parts according to the Virasoro algebras. Consequently, we must also find non-negative integers $N_{h, \bar{h}}$ such that

$$Z_{k_1, 0} = \sum_{h, \bar{h}} N_{h, \bar{h}} \chi_h \bar{\chi}_{\bar{h}} \quad (280)$$

with the requirement that it be invariant under $\Gamma^0(k)$. If this is the case, acting with an arbitrary conformal transformation $\gamma \in \Gamma$ on the argument τ (and $\bar{\tau}$) of $Z_{k_1, 0}$, must generate a function $Z_{k_1, 0}(\gamma\tau)$ invariant under the conjugate group $\gamma\Gamma^0(k)\gamma^{-1}$. The number of such functions is the number of (right) equivalence classes, $\gamma \sim \gamma'$ if $\gamma'^{-1}\gamma \in \Gamma^0(k)$, which is equal to

$$d_k = \frac{1}{2}k^2 \prod_{\substack{p \text{ prime} \\ p|k}} \left(1 - \frac{1}{p^2}\right) \quad k \geq 3 \quad (281)$$

$$d_2 = 3$$

Here, the symbol $p|k$ means that p divides k . Each of the above functions is invariant under $\Gamma(k)$, the largest subgroup common to all $\gamma\Gamma^0(k)\gamma^{-1}$. This group, called the principal congruence subgroup (of level k) is an invariant subgroup of Γ . The finite quotient $\Gamma/\Gamma(k)$ is the modular group on integers modulo k . The d_k functions $Z_{k_1, 0}(\gamma\tau)$ correspond in fact to other boundary conditions $Z_{k_1', k_2'} = Z_{-k_1', -k_2'}$, each one being left invariant by a group $\gamma\Gamma^0(k)\gamma^{-1}/\Gamma(k)$ isomorphic to $\mathbb{Z}/k\mathbb{Z}$, the cyclic symmetry

group. These other partition functions are again sesquilinear forms of the type (280), where, however, the coefficients are not restricted any longer to be positive integers. Of course, in this notation, $Z_{0,0}$ is the unfrustrated partition function derived in the previous section. The index of $\Gamma(k)$ in Γ , or the order of $\Gamma/\Gamma(k)$, is kd_k .

Let us apply this formalism to the Ising and three state Potts models, $m = 3$ and 5 in the unitary series. When $k \leq 4$, the group $\Gamma^0(k)$ is generated by the transformations $\tau \rightarrow \tau + k$ and $\tau \rightarrow \tau/\tau + 1$. For the Ising model, we expect a group of symmetry $Z/2Z$, i.e. $k = 2$. The most general form of the frustrated partition function $Z_{1,0}$ is, using the notation χ_h ,

$$Z_{1,0} = (N_1 - N_2) \left[|\chi_0|^2 + |\chi_{\frac{1}{2}}|^2 + |\chi_{\frac{1}{18}}|^2 \right] + N_2 \left[|\chi_0 + \chi_{\frac{1}{2}}|^2 + |\chi_{\frac{1}{18}}|^2 \right] \quad (282)$$

The first bracket is $Z_{0,0}$, the unfrustrated partition function, the second term involves the fermionic primary operators through $\chi_{\frac{1}{2}}\bar{\chi}_0$ and $\chi_0\bar{\chi}_{\frac{1}{2}}$. Acting with $\tau \rightarrow -\tau^{-1}$ generates $Z_{0,1}$ in the form

$$Z_{0,1} = (N_1 + N_2) \left[|\chi_0|^2 + |\chi_{\frac{1}{2}}|^2 \right] + (N_1 - N_2) |\chi_{\frac{1}{18}}|^2 \quad (283)$$

In the strip limit, the leading terms of $Z_{0,1}$ and $Z_{0,0}$ should be identical, as the effect of reversing the spins far apart should not affect the asymptotic leading behaviour. This fixes $N_1 + N_2 = 1$ and, assuming $N_2 \neq 0$ for a nontrivial $Z_{1,0}$, requires $N_1 = 0$, $N_2 = 1$.

As a byproduct, if $\tau = iM/L$, and in the limit $M/L \rightarrow \infty$, we expect that

$$Z_{1,0}/Z_{0,0} = \exp(-Mf) \quad (284)$$

with f the interface free energy per unit length. In a self-dual model such as the present one, duality relates this interface energy with the behaviour of the spin-spin correlation function at criticality and yields for a finite system with width $L = |\omega_1|$, $f = \pi\eta/L$. Comparing with the above expressions, this produces correctly the value of the exponent $\eta = \frac{1}{4}$.

We can apply the same argument for the three state Potts model denoted (A_4, D_4) with $m = 5$ (see equation (276)). One can check

that

$$Z_{1,0} = Z_{2,0} = (\chi_0 + \chi_3)^* \chi_{\frac{1}{3}} + (\chi_0 + \chi_3) \chi_{\frac{1}{3}}^* + (\chi_{\frac{2}{3}} + \chi_{\frac{1}{3}})^* \chi_{\frac{1}{15}} + (\chi_{\frac{2}{3}} + \chi_{\frac{1}{3}}) \chi_{\frac{1}{15}}^* + |\chi_{\frac{2}{3}}|^2 + |\chi_{\frac{1}{15}}|^2 \quad (285a)$$

is invariant under $\Gamma^0(3)$, in agreement with a $Z/3Z$ symmetry of the model. It exhibits the existence of a spin $\frac{1}{3}$ "pseudofermion". The ratio $Z_{1,0}/Z_{0,0}$ in the strip limit allows one to confirm the value $\eta = \frac{4}{15}$. On the other hand, the computation of $Z_{0,1}$ allows one to understand the behaviour of the primary operators with respect to the discrete symmetry groups. For instance for the Potts model,

$$Z_{0,1} = |\chi_0 + \chi_3|^2 + |\chi_{\frac{2}{3}} + \chi_{\frac{1}{3}}|^2 - |\chi_{\frac{2}{3}}|^2 - |\chi_{\frac{1}{15}}|^2 \quad (285b)$$

In the generic case of the principal series of unitary (A, A) models, one finds a generalization of (282) in the form

$$Z_{1,0} = \frac{1}{2} \sum_{r=1}^{n-1} \sum_{s=1}^m \chi_{r,s} \chi_{r,m+1-s}^* \quad (286)$$

also invariant under $\Gamma^0(2)$ indicating a $Z/2Z$ symmetry of the corresponding restricted solid on solid, or multicritical, model. The operator with lowest dimension in (286) yields the interface free energy

$$\begin{aligned} Lf &= 2\pi \min(h_{r,s} + h_{r,m+1-s}) \\ &= \pi \frac{(m-1)(m+3)}{4m(m+1)} \quad m \text{ odd} \\ &= \pi \frac{(m-2)(m+2)}{4m(m+1)} \quad m \text{ even} \end{aligned} \quad (287)$$

The relation $Lf = \pi\eta$ seems to remain valid only up to $m = 4$ where it agrees with the value $\eta = \frac{3}{20}$ of the tricritical model.

Similarly, one can show that the two unitary theories with $m = 11$ or 12 , corresponding to the Lie algebra E_8 , have also a $Z/2Z$ symmetry. For instance, when $m = 11$

$$\begin{aligned} Z_{1,0} &= \sum_{r \text{ odd}=1}^9 \left\{ (\chi_{r,1} + \chi_{r,7}) (\chi_{r,5} + \chi_{r,11})^* \right. \\ &\quad + (\chi_{r,1} + \chi_{r,7})^* (\chi_{r,5} + \chi_{r,11}) \\ &\quad \left. + |\chi_{r,4} + \chi_{r,8}|^2 \right\} \end{aligned} \quad (288)$$

$$Z_{0,1} = \sum_{r \text{ odd}=1}^9 |\chi_{r,1} + \chi_{r,7}|^2 + |\chi_{r,5} + \chi_{r,11}|^2 - |\chi_{r,4} + \chi_{r,8}|^2$$

By comparing with $Z_{0,0}$, we can figure out from this last expression which are the even or odd operators under the $Z/2Z$ group.

Find an argument to show that none of these minimal models admits a continuous symmetry, from the observation that they do not contain operators with dimensions (1,0) and (0,1) as candidates for the generating current (Friedan, Qiu, Shenker).

9.3.6 Nonminimal models

We have investigated cases with finitely many primary operators, a rational central charge smaller than unity, and conformal weights in the finite Kac table. There exists however a whole zoo of alternative theories which violate one or several of the above restrictions. We describe here some examples.

In section 2.1, when discussing the Gaussian model, we observed that instead of considering a field φ taking arbitrary real values, we could look at it as an angular variable with φ and $\varphi + 2\pi\rho$ identified. This led us to define (equations (169) and (170)) mixed operators $O_{n,m}(z, \bar{z})$ with conformal weights $h_{n,m} = \frac{1}{4}(n/\rho + \rho m)^2$, $\bar{h}_{n,m} = h_{n,-m}$, with n and m relative integers giving the intensity of a source of "electric" and "magnetic" charge located at the point $(z, \bar{z})$. We now investigate a similar situation on a torus. Since φ is an angular variable, it can wind around the unit circle a number of times as we describe a closed noncontractible path on the torus. These fields are such that

$$\varphi(z + k\omega_1 + k'\omega_2) - \varphi(z) = 2\pi\rho(km + k'm') \quad (289)$$

with the variable $\bar{z}$ omitted. Let us consider the path integral in this sector. Fields fulfilling condition (289) may be written as a superposition $\varphi = \varphi_{\text{per}} + \varphi_{\text{class}}$ of a fluctuating periodic field φ_{per} and a specific solution to (289), φ_{class} , linear in $z, \bar{z}$, hence harmonic,

$$\varphi_{\text{class}}(z, \bar{z}) = 2\pi\rho \left\{ \frac{z}{\omega_1} \frac{m\bar{\tau} - m'}{\bar{\tau} - \tau} + \text{c.c.} \right\} \quad (290)$$

The action is $S_{\text{periodic}} + \pi\rho^2 |m\tau - m'|^2 / \text{Im } \tau$ and the partition function in the m, m' sector reads

$$Z_{m',m}(\tau) = Z_1(\tau) \exp -\pi\rho^2 \frac{|m\tau - m'|^2}{\text{Im } \tau} \quad (291a)$$

with $Z_1(\tau)$ the Gaussian partition function given by equation (227). From its definition, it is clear that under a modular transformation

$$Z_{m',m} \left(\frac{a\tau + b}{c\tau + d} \right) = Z_{am'+bm, cm'+dm}(\tau) \quad (291b)$$

This shows that the modular group acts by linear transformations on the winding numbers m and m' . If we sum over all m, m' with a constant weight on the orbits of the group, we obtain a modular invariant. The simplest example is $Z_1(\tau) \equiv Z_{0,0}(\tau)$. As a second simple instance, we consider the other obvious invariant obtained by summing over all pairs m, m' . We append a prefactor ρ to normalize the large τ (small q) behaviour to obtain

$$Z(\rho, \tau) = \rho Z_1(\tau) \sum_{m',m} \exp \left(-\pi\rho^2 \frac{|m\tau - m'|^2}{\text{Im } \tau} \right) \quad (292)$$

The Poisson formula applied to the m' summation yields, with $q = \exp 2i\pi\tau$,

$$Z(\rho, \tau) = \frac{1}{|\eta(\tau)|^2} \sum_{n,m} q^{\frac{1}{4}(n/\rho + \rho m)^2} \bar{q}^{\frac{1}{4}(n/\rho - \rho m)^2} \quad (293)$$

where the weights $h_{n,m} = \frac{1}{4}(n/\rho + \rho m)^2$ and $\bar{h}_{n,m} = h_{n,-m}$ are recognized as those assigned to the operator $O_{n,m}$. This is called a Coulombic partition function, referring to logarithmic pair interactions between integral electric and magnetic charges. Comparing the expressions (292) and (293), we see that a factor $\rho/\text{Im } \tau^{\frac{1}{2}}$ has disappeared. For small q , the behaviour $(q\bar{q})^{-\frac{1}{24}}$ indicates that the central charge is unity. All dimensions $\Delta_{n,m} = \frac{1}{2}(n^2/\rho^2 + n^2\rho^2)$ are nonnegative and vary continuously with ρ , and angular momenta $S_{n,m} = nm$ are integers. For a generic ρ , $q^{h_{n,m}}/\eta(\tau)$ is a character of an irreducible representation ($c = 1, h_{n,m}$) of the Virasoro algebra. It is also clear that $Z(\rho, \tau) = Z(\rho^{-1}, \tau)$.

One application of this result is to the XY-model. Recall that the renormalization flow drives the two-point function to a

critical one with exponent $\eta = \frac{1}{4}$, up to logarithmic corrections, indicating the presence of a marginal operator of dimension 2. Let us identify the spin operator of the XY-model with $O_{\pm 1,0}$ in such a way that $\eta = \frac{1}{4} = 2\Delta_{\pm 1,0}$. This suggests the choice $\rho = 2$. The corresponding magnetic or vortex operators $O_{0,\pm 1}$ then have dimension two, which signals that vortices become relevant. At any rate $Z(2, \tau) = Z(\frac{1}{2}, \tau)$ is a consistent guess for an XY partition function at the end point of the critical line which should therefore have a central charge unity.

A further possibility is to consider φ as an angular variable of period $2\pi\rho$ with φ and $-\varphi$ identified. This is the orbifold model, and it can be understood as follows. The matrices with diagonal entries $e^{\pm i\varphi/\rho}$, $e^{\mp i\varphi/\rho}$ represent a conjugacy class within the group $SU(2)$. The corresponding partition function on a torus involves three more sectors where along any one or both generators the field changes sign. From our computation for the Ising model we deduce that in the orbifold case

$$Z_{\text{orb}} = \frac{1}{2} \left[Z(\rho; \tau) + 2 \left(|A(\tau)|^2 + |B(\tau)|^2 + |C(\tau)|^2 \right) \right] \quad (294a)$$

$$Z_{\text{Ising}} = \frac{1}{2} \left[\frac{1}{|A(\tau)|^2} + \frac{1}{|B(\tau)|^2} + \frac{1}{|C(\tau)|^2} \right] \quad (294b)$$

with

$$\begin{aligned} A(\tau) &= \frac{q^{\frac{1}{48}}}{\prod_0^\infty (1 - q^{n+\frac{1}{2}})} = \frac{1}{\eta(\tau)} \sum q^{\frac{1}{4}(2p+\frac{1}{2})^2} \\ B(\tau) &= \frac{q^{\frac{1}{48}}}{\prod_0^\infty (1 + q^{n+\frac{1}{2}})} = \frac{1}{\eta(\tau)} \sum (-)^p q^{\frac{1}{4}(2p+\frac{1}{2})^2} \\ C(\tau) &= \frac{1}{\sqrt{2}} \frac{q^{\frac{1}{48}}}{q^{\frac{1}{16}} \prod_1^\infty (1 + q^n)} = \frac{1}{\sqrt{2}\eta(\tau)} \sum (-1)^p q^{p^2} \end{aligned} \quad (294c)$$

Interestingly the combination appearing in (294a) simplifies to

$$Z_{\text{orb}}(\rho) = Z_{\text{orb}}(\rho^{-1}) = \frac{1}{2} [Z(\rho) + 2Z(2) - Z(1)] \quad (294d)$$

so that this describes a family of models with the same ρ -dependent dimensions as before, as well as some constant ones, and has been identified with the Ashkin-Teller critical line. For a description of this model of two spin systems $\sigma_i = \pm 1$, $\tau_i = \pm 1$

interacting via a four-spin term with

$$S = \sum_{(ij)} \beta_\sigma \sigma_i \sigma_j + \beta_\tau \tau_i \tau_j + \beta_{\sigma\tau} \sigma_i \sigma_j \tau_i \tau_j$$

we refer to the literature. One observes that $Z_{\text{orb}}(1) = Z(2)$ describe the same XY-model. One can picture the set of Coulombic ($O(2)$ invariant) and orbifold models at $c = 1$ as in figure 6 and point out on these two lines special points corresponding to various extra symmetries. Such is the case for $Z(1)$ which describes an $SU(2) \times SU(2)$ symmetric model, or $Z(\sqrt{2})$ which coincides with a Dirac spinor model, with

$$Z(\sqrt{2}) = \frac{1}{2} \left(\frac{1}{|A(\tau)|^4} + \frac{1}{|B(\tau)|^4} + \frac{1}{|C(\tau)|^4} \right) \quad (295)$$

while $Z_{\text{orb}}(\sqrt{2})$ is Z_{Ising}^2 . Other significant values are discussed in the references. There exist also three other interesting models constructed by Pasquier which do not lie on these two lines and correspond to exceptional affine simply laced Lie algebras

$$\begin{aligned} \hat{E}_6 &\rightarrow \frac{1}{2} (2Z(3) + Z(2) - Z(1)) \\ \hat{E}_7 &\rightarrow \frac{1}{2} (Z(4) + Z(3) + Z(2) - Z(1)) \\ \hat{E}_8 &\rightarrow \frac{1}{2} (Z(5) + Z(3) + Z(2) - Z(1)) \end{aligned} \quad (296)$$

One can also use the previous formalism to generate partition functions for nonminimal models with central charge $c < 1$, describing the Q -state Potts model for $0 < Q < 4$ and the $O(n)$ model for continuous n varying between -2 and 2 , the range where there exists a critical theory. These systems are defined for continuous n or Q through a high temperature expansion. It is possible to find a chain of plausible mappings between these models and integrable six vertex models or Coulomb gases with specific restrictions. The full derivations are outside the scope of this chapter and we shall content ourselves with a description of the results.

For ρ in the interval 1 to $\sqrt{2}$, set

$$\begin{aligned} n &= -2 \cos \pi \rho^2 \\ e_0 &= \pm (\rho^2 - 1) \pmod{2} \end{aligned} \quad (297a)$$

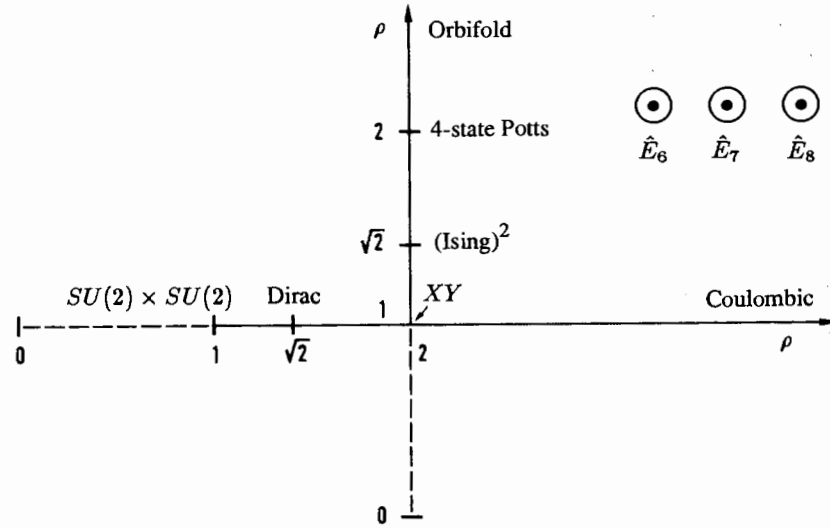


Fig. 6 Known $c = 1$ models. The horizontal line refers to $O(2)$ symmetric theories with a $2\pi\rho$ -periodic field, the vertical one to orbifold models. The dotted parts are equivalent under $\rho \rightarrow \rho^{-1}$. The extra points depict three exceptional cases.

in such away that n ranges between 2 and -2 . The corresponding partition function is given by

$$Z(n; \tau) = \frac{\rho}{2} Z_1(\tau) \sum_{m', m} \exp - \frac{\pi \rho^2}{4} \frac{|m\tau - m'|^2}{\text{Im}\tau} \cos [\pi e_0(m', m)] \quad (297b)$$

In this expression (m', m) is the largest common positive divisor of m' and m . From our previous discussion, the most general modular invariant built out of a superposition of $Z_{m', m}$ would allow an arbitrary function of (m', m) . There are two steps involved in deriving equations (297), the first is to show the origin of the relation between continuously varying n and ρ , the second to introduce the “defect charge” e_0 . In the process of using a continuously variable n through a high temperature lattice expansion, one loses the property that Z is the limit of a trace of a transfer operator. In other words, by reexpanding $Z(n, \tau)$ as a series in fractional powers of q and $\bar{q}$, we have no guarantee that the coefficients are positive integers except in the case $\rho^2 = 1$, $\frac{4}{3}$, $\frac{3}{2}$ corresponding respectively to $n = 2, 1, 0$. For

$\rho^2 = 1$, $e_0 = 0 \bmod 2$, the cosine in (297b) reduces to unity, and $Z(n = 2, \tau)$ to $Z(\rho = \frac{1}{2}; \tau) = Z(\rho = 2, \tau)$, i.e. the XY partition function, as it should. Surprisingly when $n = 1$, the above expression agrees with the Ising partition function, while for $n = 0$, its value is unity. To give a hint on the role of e_0 , we note that its purpose is to correct for weight factors attached to noncontractible loops on the torus.

In the strip limit, when $\tau \rightarrow +i\infty$, the leading term in the double sum (297b) corresponds to $m = 0$, leading to a behaviour, when summing over m' ,

$$Z(n; \tau) \sim (q\bar{q})^{(e_0^2/4\rho^2) - \frac{1}{24}}$$

for that determination of e_0 in the range $[0, 1]$, i.e. $e_0 = \rho^2 - 1$. Hence the central charge is given by

$$c = 1 - 6 \frac{e_0^2}{\rho^2} = 1 - 6 \frac{(\rho^2 - 1)^2}{\rho^2} \quad (298a)$$

Let us parametrize as usual c in the form $1 - 6/m(m+1)$. Collecting (297a) and (298a), we have

$$\rho^2 = \frac{m+1}{m} \quad n = 2 \cos \frac{\pi}{m} \quad e_0 = \frac{1}{m} \quad c = 1 - \frac{6}{m(m+1)} \quad (298b)$$

Let us now check that for $n = 1$, where $m = 3$, we do recover the known Ising result. Since $e_0 = \frac{1}{3}$, we split the sum in (297b) over values of (m', m) modulo 6, as

$$\sum_{m', m} = \sum_{(m', m) \equiv 0 \bmod 6} + \frac{1}{2} \sum_{(m', m) \equiv \pm 1 \bmod 6} - \frac{1}{2} \sum_{(m', m) \equiv \pm 2 \bmod 6} - \sum_{(m', m) \equiv 3 \bmod 6}$$

Each piece can be related to a Coulombic partition function with a different value of ρ . For instance

$$\sum_{(m', m) \equiv 0 \bmod 6} = \frac{1}{6} Z(3\rho, \tau)$$

taking into account the prefactor ρ in (297b). Similarly

$$\begin{aligned}\sum_{(m',m)=3 \bmod 6} &= \frac{1}{3}Z\left(\frac{3}{2}\rho, \tau\right) - \frac{1}{6}Z(3\rho, \tau) \\ \sum_{(m',m)=\pm 2 \bmod 6} &= \frac{1}{2}Z(\rho, \tau) - \frac{1}{6}Z(3\rho, \tau) \\ \sum_{(m',m)=\pm 1 \bmod 6} &= Z\left(\frac{1}{2}\rho, \tau\right) - \frac{1}{2}Z(\rho, \tau) - \frac{1}{3}Z\left(\frac{3}{2}\rho, \tau\right) + \frac{1}{6}Z(3\rho, \tau)\end{aligned}$$

Putting everything together, we see that for $e_0 = \frac{1}{3}$ and $\rho = \sqrt{\frac{4}{3}}$, equation (297) reduces to

$$Z(n=1, \tau) = \frac{1}{2} \left[Z(\rho = 2\sqrt{3}, \tau) - Z(2/\sqrt{3}, \tau) \right] \quad (299)$$

It is not obvious at this stage that we have reached the desired result. We leave it as an exercise for the reader to check that it is the case.

All unitary minimal partition functions of table II can be written as finite sums of Coulombic partition functions. It is left as a further nontrivial exercise to verify this point, as well as the fact that, for $n=0$, $m=2$, $c=0$, $e_0 = \frac{1}{2}$, the partition function reduces to unity. As $n \rightarrow -2$, i.e. $m \rightarrow 1$, one finds $Z(n=-2, \tau) = 0$, but

$$\left. \frac{\partial}{\partial n} Z(n, \tau) \right|_{n \rightarrow -2} = \text{Im } \tau \eta^2 \bar{\eta}^2$$

a fermionic partition function equal to the inverse square of the Gaussian free field partition function. This is in agreement with the correspondence between $O(n)$ bosonic models with $n = -2$ and (complex) fermionic ones. One can also study the limit $n \rightarrow 0$ for the polymer problem.

Similar techniques can be applied to critical Potts models for continuous Q varying between 0 and 4. One sets

$$\begin{aligned}\rho^2 &= \frac{4m}{m+1} & Q &= \left(2 \cos \frac{\pi}{m+1} \right)^2 \\ e_0 &= \frac{2}{m+1} & c &= 1 - \frac{6}{m(m+1)}\end{aligned}$$

so that everything is parametrized in terms of m . Then the partition function reads

$$\begin{aligned}Z(Q, \tau) &= \frac{1}{2}\rho Z_1(\tau) \sum_{m', n} \exp \left\{ -\frac{1}{4}\pi\rho^2 \frac{|m\tau - m'|^2}{\text{Im } \tau} \right\} \cos [\pi e_0(m, m')] \\ &\quad + \frac{1}{2}(Q-1) [Z(\rho, \tau) - Z(\frac{1}{2}\rho, \tau)]\end{aligned} \quad (300b)$$

One can look at particular cases, such as $Q = 1, 2, 3$, where, in the last case, one finds the 3-state Potts partition function given in equation (276). The limiting value for $m \rightarrow \infty$, $Q \rightarrow 4$ and $c \rightarrow 1$, is also of interest; and appears on figure 6. The first term in equation (300b) is analogous to the $O(n)$ partition function.

9.3.7 Correlations in a half plane

We have investigated in detail geometrically simple situations, the plane, the cylinder or the torus. In a sense, none of these exhibits genuine boundary effects. This is not true for a critical model in a half plane. Having in mind a discretization on a lattice as a regularization, we consider free boundary conditions. We assume that at criticality this amounts to requiring the vanishing of the magnetization at the boundary. More precisely, this is compatible with (i) the vanishing of correlations for large separations (ii) global conformal invariance in the upper half plane (i.e. under $SL(2, \mathcal{R})$), since the point at infinity belongs to the boundary.

The domain under consideration is the upper half plane $\text{Im } z > 0$. With an added point at infinity this is conformally equivalent to the interior of the unit disk, and is invariant under global real conformal transformations of $SL(2, \mathcal{R})/Z_2$, in the form of real Möbius transformations $z \rightarrow z' = (\alpha z + \beta)/(\gamma z + \delta)$, $\alpha, \beta, \gamma, \delta$ real, $\alpha\delta - \beta\gamma = 1$. Thus the same transformation applies to the variable $\bar{z}$. In particular, this involves only real translations and dilatations. Correspondingly, even the two-point function becomes nontrivial. It has a different behaviour in the directions parallel or perpendicular to the boundary. We associate to each point in the upper half plane its symmetric image with respect to the real axis. Now let two points z_1 and z_2 be given in the upper half plane, together with their images, $\bar{z}_1$ and $\bar{z}_2$ (in the lower half plane). From those four complex numbers, we can form the real

taking into account the prefactor ρ in (297b). Similarly

$$\begin{aligned}\sum_{(m',m)=3 \bmod 6} &= \frac{1}{3}Z\left(\frac{3}{2}\rho, \tau\right) - \frac{1}{6}Z(3\rho, \tau) \\ \sum_{(m',m)=\pm 2 \bmod 6} &= \frac{1}{2}Z(\rho, \tau) - \frac{1}{6}Z(3\rho, \tau) \\ \sum_{(m',m)=\pm 1 \bmod 6} &= Z\left(\frac{1}{2}\rho, \tau\right) - \frac{1}{2}Z(\rho, \tau) - \frac{1}{3}Z\left(\frac{3}{2}\rho, \tau\right) + \frac{1}{6}Z(3\rho, \tau)\end{aligned}$$

Putting everything together, we see that for $e_0 = \frac{1}{3}$ and $\rho = \sqrt{\frac{4}{3}}$, equation (297) reduces to

$$Z(n=1, \tau) = \frac{1}{2} \left[Z(\rho = 2\sqrt{3}, \tau) - Z(2/\sqrt{3}, \tau) \right] \quad (299)$$

It is not obvious at this stage that we have reached the desired result. We leave it as an exercise for the reader to check that it is the case.

All unitary minimal partition functions of table II can be written as finite sums of Coulombic partition functions. It is left as a further nontrivial exercise to verify this point, as well as the fact that, for $n=0$, $m=2$, $c=0$, $e_0 = \frac{1}{2}$, the partition function reduces to unity. As $n \rightarrow -2$, i.e. $m \rightarrow 1$, one finds $Z(n=-2, \tau) = 0$, but

$$\frac{\partial}{\partial n} Z(n, \tau) \Big|_{n \rightarrow -2} = \text{Im } \tau \eta^2 \bar{\eta}^2$$

a fermionic partition function equal to the inverse square of the Gaussian free field partition function. This is in agreement with the correspondence between $O(n)$ bosonic models with $n = -2$ and (complex) fermionic ones. One can also study the limit $n \rightarrow 0$ for the polymer problem.

Similar techniques can be applied to critical Potts models for continuous Q varying between 0 and 4. One sets

$$\begin{aligned}\rho^2 &= \frac{4m}{m+1} & Q &= \left(2 \cos \frac{\pi}{m+1} \right)^2 \\ e_0 &= \frac{2}{m+1} & c &= 1 - \frac{6}{m(m+1)}\end{aligned}$$

so that everything is parametrized in terms of m . Then the partition function reads

$$\begin{aligned}Z(Q, \tau) &= \frac{1}{2}\rho Z_1(\tau) \sum_{m', m} \exp \left\{ -\frac{1}{4}\pi\rho^2 \frac{|m\tau - m'|^2}{\text{Im}\tau} \right\} \cos [\pi e_0(m, m')] \\ &\quad + \frac{1}{2}(Q-1) [Z(\rho, \tau) - Z(\frac{1}{2}\rho, \tau)]\end{aligned} \quad (300b)$$

One can look at particular cases, such as $Q = 1, 2, 3$, where, in the last case, one finds the 3-state Potts partition function given in equation (276). The limiting value for $m \rightarrow \infty$, $Q \rightarrow 4$ and $c \rightarrow 1$, is also of interest, and appears on figure 6. The first term in equation (300b) is analogous to the $O(n)$ partition function.

9.3.7 Correlations in a half plane

We have investigated in detail geometrically simple situations, the plane, the cylinder or the torus. In a sense, none of these exhibits genuine boundary effects. This is not true for a critical model in a half plane. Having in mind a discretization on a lattice as a regularization, we consider free boundary conditions. We assume that at criticality this amounts to requiring the vanishing of the magnetization at the boundary. More precisely, this is compatible with (i) the vanishing of correlations for large separations (ii) global conformal invariance in the upper half plane (i.e. under $SL(2, \mathcal{R})$), since the point at infinity belongs to the boundary.

The domain under consideration is the upper half plane $\text{Im } z > 0$. With an added point at infinity this is conformally equivalent to the interior of the unit disk, and is invariant under global real conformal transformations of $SL(2, \mathcal{R})/Z_2$, in the form of real Möbius transformations $z \rightarrow z' = (\alpha z + \beta)/(\gamma z + \delta)$, $\alpha, \beta, \gamma, \delta$ real, $\alpha\delta - \beta\gamma = 1$. Thus the same transformation applies to the variable $\bar{z}$. In particular, this involves only real translations and dilatations. Correspondingly, even the two-point function becomes nontrivial. It has a different behaviour in the directions parallel or perpendicular to the boundary. We associate to each point in the upper half plane its symmetric image with respect to the real axis. Now let two points z_1 and z_2 be given in the upper half plane, together with their images, $\bar{z}_1$ and $\bar{z}_2$ (in the lower half plane). From those four complex numbers, we can form the real

cross ratio u , invariant under $SL(2, \mathcal{R})$,

$$u = \frac{(z_1 - \bar{z}_1)(z_2 - \bar{z}_2)}{(z_1 - \bar{z}_2)(z_2 - \bar{z}_1)} \quad 0 < u < 1 \quad (301)$$

A real global conformal transformation which maps $z_1, \bar{z}_1, z_2, \bar{z}_2$ onto $z'_1, \bar{z}'_1, z'_2, \bar{z}'_2$ preserves the cross ratio, and has the form

$$z \rightarrow z' \quad \frac{(z - \bar{z}_1)(z_2 - \bar{z}_2)}{(z - \bar{z}_2)(z_2 - \bar{z}_1)} = \frac{(z' - \bar{z}'_1)(z'_2 - \bar{z}'_2)}{(z' - \bar{z}'_2)(z'_2 - \bar{z}'_1)} \quad (302)$$

If we look at a two-point correlation $G(1, 2)$ for the primary operator $A_{h, \bar{h}}$, global invariance implies

$$\begin{aligned} G(1, 2) &\equiv \langle A_{h, \bar{h}}(z_1, \bar{z}_1) A_{h, \bar{h}}(z_2, \bar{z}_2) \rangle \\ &= \left[\frac{(z'_1 - \bar{z}'_1)(z'_2 - \bar{z}'_2)}{(z_1 - \bar{z}_1)(z_2 - \bar{z}_2)} \right]^{h+h'} \left[\frac{(\bar{z}_1 - \bar{z}_2)(z'_1 - z'_2)}{(z_1 - z_2)(\bar{z}'_1 - \bar{z}'_2)} \right]^{h-\bar{h}} G(1', 2') \end{aligned} \quad (303)$$

The group $SL(2, \mathcal{R})$ depends on three real parameters. One can therefore impose on z'_1 and z'_2 three conditions, for instance to be pure imaginary and such that $z'_1 + z'_2 = 2i$. Thus we may choose

$$z'_1 = i(1 + \sqrt{1-u}) \quad z'_2 = i(1 - \sqrt{1-u}) \quad (304)$$

For simplicity assume a scalar field, i.e. such that $h = \bar{h}$. We have then, with $z = x + iy$,

$$\begin{aligned} G(1, 2) &= g(u)/(y_1 y_2)^{2h} \quad u = \frac{4y_1 y_2}{(x_1 - x_2)^2 + (y_1 + y_2)^2} \\ g(u) &= u^{2h} \langle A_{h, h} \left(i \left[1 + \sqrt{1-u} \right] \right) A_{h, h} \left(i \left[1 - \sqrt{1-u} \right] \right) \rangle \end{aligned} \quad (305)$$

Thus global conformal invariance alone leaves us with an unknown function of a real variable $g(u)$.

In dimension larger than two, the subgroup of the global conformal group which leaves a half space invariant leads to a similar result, with y_1 and y_2 the distances to the boundary and $|x_1 - x_2|$ replaced by the length of the parallel component of the vector $\mathbf{r}_1 - \mathbf{r}_2$.

The two-point function in the present case is analogous to the four-point function in the plane with z_1, z_2, z_3, z_4 replaced by $z_1, \bar{z}_1, z_2, \bar{z}_2$. There is a similar doubling for higher correlations. To

see this relation more precisely, we look at the energy-momentum tensor and require that in Cartesian coordinates, there be no energy-momentum flow across the boundary. Using complex coordinates, this is equivalent to saying that on the boundary $T_{zz}(x) = T_{\bar{z}\bar{z}}(x)$. This enables one to extend $T(z) \equiv T_{zz}(z)$ analytically in the lower half plane through

$$\text{Im } z < 0 \quad T(z) = \bar{T}(\bar{z}) \quad (306)$$

where obviously $\bar{z}$ is in the upper half plane, and as a result the right-hand side is well-defined. Thus we now use only the diagonal part of the product of two Virasoro algebras. Repeating the reasoning of section 1.3, one arrives at a formula similar to (40) (Cardy)

$$\begin{aligned} \langle T(z) A_1(z_1, \bar{z}_1) \dots \rangle &= \sum_p \left(\frac{h_p}{(z - z_p)^2} + \frac{\bar{h}_p}{(z - \bar{z}_p)^2} \right. \\ &\quad \left. + \frac{1}{(z - z_p)} \frac{\partial}{\partial z_p} + \frac{1}{(z - \bar{z}_p)} \frac{\partial}{\partial \bar{z}_p} \right) \langle A_1(z_1, \bar{z}_1) \dots \rangle \end{aligned} \quad (307)$$

One can derive analogous consequences as we did in the plane, when we have primary fields corresponding to degenerate representations, in which case we obtain differential equations for the correlations.

Let us illustrate this on our perennial example of the Ising model for the spin-spin correlation. Adapting the calculations done in section 2.2, one finds a hypergeometric equation admitting a solution through radicals, such that

$$\begin{aligned} \langle \sigma(z_1, \bar{z}_1) \sigma(z_2, \bar{z}_2) \rangle &= \frac{1}{(y_1 y_2)^{\frac{1}{4}}} \frac{1}{\lambda^{\frac{1}{4}}} (a_+ \sqrt{1+\lambda} + a_- \sqrt{1-\lambda}) \\ \lambda = \sqrt{1-u} &= \sqrt{\frac{(x_1 - x_2)^2 + (y_1 - y_2)^2}{(x_1 - x_2)^2 + (y_1 + y_2)^2}} = \left| \frac{z_1 - z_2}{z_1 - \bar{z}_2} \right| \end{aligned} \quad (308)$$

Provided $a_+ + a_-$ is non zero, the short-distance behaviour for finite y_1 and y_2 is the expected one

$$\langle \sigma_1 \sigma_2 \rangle \underset{z_{12} \rightarrow 0}{\sim} \text{cst} \frac{(y_1 + y_2)^{\frac{1}{4}}}{(y_1 y_2)^{\frac{1}{4}}} \frac{1}{|z_{12}|^{\frac{1}{4}}} \quad (309)$$

If however, again for y_1 and y_2 finite, we let $|x_1 - x_2|$ tend to infinity, we have

$$\lambda \rightarrow 1 - 2y_1y_2/(x_1 - x_2)^2$$

meaning that if $a_+ \neq 0$ the correlation tends to a constant. If we require that under such circumstances $\langle \sigma_1 \sigma_2 \rangle$ tends to zero, we are forced to the conclusion that $a_+ = 0$, which means that up to a multiplicative constant

$$\langle \sigma_1 \sigma_2 \rangle = \frac{\text{cst}}{(y_1 y_2)^{\frac{1}{2}}} \left(\lambda^{-\frac{1}{2}} - \lambda^{\frac{1}{2}} \right)^{\frac{1}{2}} \quad (310)$$

As a consequence, when $|x_1 - x_2| \rightarrow \infty$ and y_1 and y_2 are finite, one finds

$$\langle \sigma_1 \sigma_2 \rangle \underset{|x_1 - x_2| \rightarrow \infty}{\sim} \text{cst} \frac{(y_1 y_2)^{\frac{1}{2}}}{|x_1 - x_2|} \quad (311a)$$

If the power of $|x_{12}|$ in the denominator is denoted $\eta_{\parallel}$, then

$$\eta_{\parallel} = 1 \quad (311b)$$

a new nontrivial result from local conformal invariance.

Since the point at infinity is to be considered as part of the boundary, equation (311a) implies by global conformal invariance that the field vanishes on the boundary. For instance

$$\langle \sigma(iy_1) \sigma(iy_2) \rangle \underset{y_1 \rightarrow 0}{\sim} \text{cst} \frac{1}{y_2^{\frac{1}{2}}} \left(\frac{y_1}{y_2} \right)^{\frac{1}{2}} \quad (312)$$

Thus the r.h.s. vanishes on the boundary, albeit with a fractional power (hence the function has an infinite derivative).

One can put the complete expression (310) in a suggestive form, which can be compared with the four-point function in the plane. Squaring $\langle \sigma_1 \sigma_2 \rangle$ and setting $z_{ij} = z_i - z_j$, $\bar{z}_{ij} = \bar{z}_i - \bar{z}_j$ etc, one has

$$\langle \sigma_1 \sigma_2 \rangle^2 = \left(\frac{z_{1\bar{2}} z_{\bar{1}2}}{z_{1\bar{1}} z_{2\bar{2}} z_{12} z_{\bar{1}\bar{2}}} \right)^{\frac{1}{2}} - \left(\frac{z_{12} z_{\bar{1}\bar{2}}}{z_{1\bar{1}} z_{2\bar{2}} z_{12} z_{\bar{1}\bar{2}}} \right)^{\frac{1}{2}} \quad (313)$$

which shows the symmetry in the simultaneous exchange $1 \leftrightarrow 2$, $\bar{1} \leftrightarrow \bar{2}$ and the antisymmetry in $1 \leftrightarrow \bar{1}$ or $2 \leftrightarrow \bar{2}$, the latter being defined after analytic continuation.

Another application of this same geometric situation can be obtained through a conformal mapping as follows. The transfor-

mation $z \rightarrow t$

$$t = \frac{L}{i\pi} \ln z \quad 0 < \text{Arg} z < \pi \quad (314)$$

maps the upper half z -plane onto a strip $0 < \text{Re} t < \pi$ with free (as opposed to periodic) boundaries. If we now write $t = x + iy$, the above two-point correlation (squared) becomes

$$\begin{aligned} \langle \sigma_1 \sigma_2 \rangle^2 &= \text{cst} \frac{(\pi/L)^{\frac{1}{2}}}{(\sin(\pi x_1/L) \sin(\pi x_2/L))^{\frac{1}{2}}} \\ &\times \left[\frac{\sin(\pi/2L)(x_1 + x_2 + iy_1 - iy_2)}{\sin(\pi/2L)(x_1 - x_2 + iy_1 - iy_2)} \right]^{\frac{1}{2}} - (x_2 \rightarrow -x_2) \end{aligned} \quad (315)$$

Note the translational invariance in the y direction, and the vanishing when x_1 or $x_2 = 0$ or L .

Using the mapping (314), we can also find, according to equation (50), the mean value $\langle T \rangle$ in the strip, assuming that it vanishes in the half plane. Thus for any central charge c

$$\langle T \rangle = \frac{1}{24} c \pi / L^2 \quad (316)$$

Correspondingly, the free energy per unit length behaves as

$$F(L) = F_0 L + \frac{1}{24} c \pi / L \quad (317)$$

The coefficient of the $1/L$ term is a quarter of the corresponding one on a periodic strip, where, instead of (316), we found $\langle T \rangle_{\text{periodic}} = \frac{1}{6} c \pi / L^2$, suggesting that (316) follows by replacing $L \rightarrow 2L$ in this last expression. Observe nevertheless that the correlation functions are not recovered as easily, as testified by equation (315).

9.3.8 The vicinity of the critical point

The knowledge of critical properties should enable one to understand the whole critical domain, characterized by a finite correlation length, or in the particle language, a finite mass spectrum. Two directions are open for investigation, but neither of them seems to be fully elaborated. The first is to obtain a better understanding of the relationship between integrable models and their conformal invariant limit. There are many signs that these two subjects have much in common, judging from the mathematical background. While this looks perhaps more like a formal question,

the second aspect, namely the study of massive continuous quantum field theories, is of primary importance. In this final section, we present some remarks on the subject.

Let us return to the computations performed in section 3.2 and examine the partition function of a complex massive scalar free field on a torus T , satisfying boundary conditions as indicated in equation (239). Denote the corresponding partition function over complex fields as

$$D_{k/N, \ell/N}^{-2}(T, m) = \int \mathcal{D}(\varphi \bar{\varphi}) \exp - \int_T \frac{d^2x}{2\pi} \bar{\varphi}(-\Delta + m) \varphi$$

$$\varphi(x + \lambda \omega_1 + \nu \omega_2) = \exp 2i\pi \frac{(k\mu - \ell\lambda)}{N} \varphi(x) \quad (318)$$

Retracing the steps of section 3.2, we find after some laborious computations

$$D_{k/N, \ell/N}(T, m) = \exp -\pi \text{Im} \tau \gamma_{\ell/N}$$

$$\left| \prod_{n=-\infty}^{+\infty} \left(1 - e^{-2i\pi \left[\frac{k}{N} + (n + \frac{\ell}{N}) \text{Re} \tau \right] - 2\pi \text{Im} \tau \sqrt{(n + \frac{\ell}{N})^2 + t^2}} \right) \right| \quad (319)$$

where t stands for

$$t = \frac{m |\omega_1|}{2\pi} > 0 \quad (320)$$

With $A = |\omega_1|^2 \text{Im} \tau$ the area of the torus, $\psi(z) = \Gamma'(z)/\Gamma(z)$ the logarithmic derivative of Euler's Γ -function, and $\gamma = -\psi(1)$ Euler's constant, the quantities $\gamma_{\ell/N}$ are given by

$$\frac{\ell}{N} = 0, \quad \gamma_0 = \frac{1}{6} - t + t^2 \ln \left[4\pi e^{-\gamma} \left(\frac{\text{Im} \tau}{A} \right)^{\frac{1}{2}} \right]$$

$$+ \frac{1}{2} t^4 \int_0^1 d\lambda (1 - \lambda) \sum_{n=1}^{\infty} \frac{1}{(n^2 + \lambda t^2)^{\frac{3}{2}}}$$

$$0 < \frac{\ell}{N} < 1, \quad \gamma_{\ell/N} = \frac{1}{6} \left[1 - \frac{6\ell(N - \ell)}{N^2} \right] \quad (321)$$

$$+ t^2 \left\{ \ln \left[4\pi \left(\frac{\text{Im} \tau}{A} \right)^{\frac{1}{2}} \right] + \frac{1}{2} \psi(\ell/N) + \frac{1}{2} \psi(1 - \ell/N) \right\}$$

$$+ \frac{1}{4} t^4 \int_0^1 d\lambda (1 - \lambda) \sum_{n=-\infty}^{+\infty} \frac{1}{[(n + \ell/N)^2 + \lambda t^2]^{\frac{3}{2}}}$$

We observe the occurrence of a logarithmic term which disappears in the difference $\gamma_{\ell/N} - \gamma_0$, a function of t only. Let us specialize the above calculation to a periodic scalar real field, in which case

$$Z_{\text{scalar}} = \frac{1}{D_{00}} = \exp \pi \text{Im} \tau \left\{ \frac{1}{6} - t + t^2 \ln \left[4\pi e^{-\gamma} (\text{Im} \tau / A)^{\frac{1}{2}} \right] \right.$$

$$+ \frac{1}{2} t^4 \int_0^1 d\lambda (1 - \lambda) \sum_1^{\infty} (n^2 + \lambda t^2)^{-\frac{3}{2}} \left. \right\}$$

$$\times \left\{ \prod_{n=-\infty}^{+\infty} \left(1 - \exp(2i\pi n \text{Re} \tau - 2\pi \text{Im} \tau \sqrt{n^2 + t^2}) \right) \right\}^{-1} \quad (322)$$

The infinite product is real and positive. The derivation shows that Z_{scalar} is modular invariant, even though it is expressed in terms of variables adapted to a choice of basis. The quantity t measures in dimensionless units the departure from criticality. One can interpret the prefactor as $\exp \frac{1}{6} \pi \text{Im} \tau c(t, \text{Im} \tau)$, where $c(0, \text{Im} \tau) = 1$, as defining a "scale dependent central charge". The appearance of a $\ln(\text{Im} \tau / A)^{1/2}$ dependence in the t^2 coefficient reflects the necessity of an ultraviolet renormalization of the specific heat, the second derivative with respect to m . Finally, the denominator arises naturally if one uses a transfer matrix, by decomposing the fields in proper modes, with the discrete momenta $2\pi n / |\omega_1|$ associated to eigenfrequencies $[(2\pi n / |\omega_1|)^2 + m^2]^{1/2}$. By extracting the zero mode responsible for a potential divergence, we recover the critical result

$$\lim_{m \rightarrow 0} m \sqrt{A} Z(m) = \frac{1}{\sqrt{\text{Im} \tau} |\eta(\tau)|^2} \quad (323)$$

More interesting is the application of the previous formulas to obtain the partition function of the Ising model with m proportional to $T - T_c$. We know from chapter 2 that we have to sum over the four periodic or antiperiodic boundary conditions, with a sign appended, taken care of here by an analytic continuation in m which we take at first to be positive. Then

$$Z_{\text{Ising}} = \frac{1}{2} \sum_{\text{boundary conditions}} \int \mathcal{D}(\psi \bar{\psi}) \exp \int_T \frac{d^2x}{2\pi} \{ \psi \bar{\partial} \psi - \bar{\psi} \partial \bar{\psi} + m \bar{\psi} \psi \}$$

$$= \frac{1}{2} \{ D_{\frac{1}{2}, \frac{1}{2}}(m) + D_{0, \frac{1}{2}}(m) + D_{\frac{1}{2}, 0}(m) + D_{00}(m) \} \quad (324)$$

i.e.

$$\begin{aligned}
 Z_{\text{Ising}} = & \frac{1}{2} \exp \pi \text{Im} \tau \left\{ \frac{1}{12} - t^2 \ln \left[\pi e^{-\gamma} \left(\frac{\text{Im} \tau}{A} \right)^{\frac{1}{2}} \right] \right. \\
 & \left. - \frac{1}{2} t^4 \int_0^1 d\lambda (1-\lambda) \sum_{n=1}^{\infty} \frac{1}{[(n-\frac{1}{2})^2 + \lambda t^2]^{\frac{3}{2}}} \right\} \\
 & \times \left[\sum_{\pm} \prod_{n=-\infty}^{+\infty} \left(1 \pm e^{2i\pi(n+\frac{1}{2})\text{Re}\tau - 2\pi\text{Im}\tau\sqrt{(n+\frac{1}{2})^2 + t^2}} \right) + \exp -\pi\text{Im}\tau \left\{ \frac{1}{4} - t \right. \right. \\
 & \left. \left. + t^2 \ln 4 - \frac{1}{2} t^4 \int_0^1 d\lambda (1-\lambda) \sum_{n=1}^{\infty} \left(\frac{1}{[(n-\frac{1}{2})^2 + \lambda t^2]^{\frac{3}{2}}} - \frac{1}{[n^2 + \lambda t^2]^{\frac{3}{2}}} \right) \right\} \right. \\
 & \left. \times \sum_{\pm} \prod_{n=-\infty}^{+\infty} \left(1 \pm e^{2i\pi n\text{Re}\tau - 2\pi\text{Im}\tau\sqrt{n^2 + t^2}} \right) \right] \quad (325)
 \end{aligned}$$

Absolute value signs have been suppressed since all infinite products are real, and the expression reduces to $Z_{\frac{1}{2}}$ at criticality ($m = 0$). All terms sensitive to the sign of t are in the exponential factors of $D_{\frac{1}{2},0}$ and $D_{0,0}$ and their zero mode factors. Combining them we get a term $(\exp(\pi\text{Im}\tau t) \pm \exp(-\pi\text{Im}\tau t))$.

The relativistic spectrum of states is the expected one, very much as in the scalar case. Not only does the central charge become scale dependent, but so also do the dimensions of the various operators. For instance, the spin and disorder operator which were indistinguishable at criticality split into two operators with scale dependent dimensions related by duality, i.e. $t \leftrightarrow -t$. As was pointed out, the occurrence of a term in $t^2 \ln(\text{Im}\tau/A)^{\frac{1}{2}}$ in the central charge is a manifestation of the renormalization properties of the dimension one operator $\bar{\psi}\psi$. In a mass perturbation, an ultraviolet subtraction is required to second order, related to a logarithmic divergence in the specific heat.

The modular invariance of equation (325) is not explicit, but an expansion in powers of m exhibits this very clearly. Denote as above by $Z_{\frac{1}{2}}$ and Z_1 the critical partition functions of the Ising and Gaussian models respectively, and let $D_{ij} \equiv D_{ij}(0)$ (with $i, j = 0, \frac{1}{2}$, omitting the $(0,0)$ combination since $D_{00} = 0$). Then

expanding (325) we get

$$\begin{aligned}
 \ln Z_{\text{Ising}}(m) = & \ln Z_{\frac{1}{2}} + \frac{mA^{\frac{1}{2}}}{2Z_1 Z_{\frac{1}{2}}} + m^2 A \left[\frac{1}{4\pi} \ln \left(\frac{Z_1 \sqrt{A} e^{\gamma}}{\pi} \right) \right. \\
 & \left. - \frac{1}{4\pi} \frac{\sum D_{ij} \ln D_{ij}}{Z_{\frac{1}{2}}} - \frac{1}{8(Z_1 Z_{\frac{1}{2}})^2} \right] + O((mA^{\frac{1}{2}})^3) \quad (326)
 \end{aligned}$$

The term linear in m corresponds to the nonvanishing expectation value of the energy operator in the spin sector. This effect was mentioned earlier. The arbitrary length scale involved in the term $m^2 \ln A^{\frac{1}{2}}$ appears as an additive renormalization of the specific heat as expected. Finally, the existence of odd as well as even terms in m shows that on a torus the maximum of the specific heat (replacing the divergence in the infinite plane) is not at $m = 0$ ($T = T_c$) but depends on the modular ratio. It is noticeable that the specific heat contains an entropic contribution due to the various fermionic spin structures.

These exact results are specific to the Ising case, and only apply to one of the relevant couplings, namely to the mass. In general one can only treat perturbatively the vicinity of the critical point. However, such a perturbation theory is not trivial as it incorporates the knowledge of the critical quantities. As an indication of the procedure, let us look at a system on a torus, assuming one of the periods ω_1 to be real. The time evolution is dictated by a Hamiltonian

$$\begin{aligned}
 H &= H_0 + V \\
 H_0 &= \frac{2\pi}{\omega_1} (L_0 + \bar{L}_0 - \frac{1}{12}c) \\
 V &= - \sum_i G_i \int_0^{\omega_1} \frac{du_1}{2\pi} \varphi_i(u_1, 0) \quad (327)
 \end{aligned}$$

The coordinate on the torus is written $u = u_1 + iu_2$, and the potential is a sum over local perturbations with couplings G to relevant fields of dimension smaller than two. For simplicity consider the case where the sum is limited to a unique term, with a primary scalar field of weights (h, h) . Let us further restrict our attention to the computation of the running central charge, although the formalism is by no means limited to this. Let P

be the momentum operator, and $|0\rangle$ the ground state. Instead of computing the full partition function, consider the following matrix element, where T denotes time ordering according to u_2

$$\begin{aligned} Z_0 &= \langle 0 | \exp \{-\text{Im} \omega_2 H + i \text{PRe} \omega_2\} | 0 \rangle \\ &= \exp \pi \text{Im} \tau \frac{1}{6} c \langle 0 | T \exp \frac{G}{2\pi} \int_0^{\text{Im} \omega_2} du_2 \int_0^{\omega_1} du_1 \varphi(u_1, u_2) | 0 \rangle \end{aligned} \quad (328)$$

It is convenient to return to the punctured plane using

$$\begin{aligned} x &= \exp 2i\pi u / \omega_1 & d^2 u &= \left(\frac{\omega_1}{2\pi} \right)^2 \frac{d^2 x}{x\bar{x}} \\ d^2 u \varphi_{\text{torus}}(u, \bar{u}) &= \left(\frac{\omega_1}{2\pi} \right)^{2-2h} \frac{d^2 x}{(x\bar{x})^{1-h}} \varphi_{\text{plane}}(x, \bar{x}) \end{aligned} \quad (329)$$

The integral is then in an annulus, $\rho = e^{-2\pi \text{Im} \tau} \leq |x| \leq 1$, and ordering is in the radial direction. It is also natural to use a dimensionless coupling

$$g = G \left(\frac{|\omega_1|}{2\pi} \right)^{2-2h} \quad (330)$$

In the previous discussion we had $G = m$, $h = \frac{1}{2}$, and t was used instead of g to recall that it played the role of a reduced temperature. On the other hand we can think of a magnetic perturbation with G the magnetic field and φ the spin field ($h = \frac{1}{16}$). Relevant perturbations will have $2 - 2h > 0$, in which case for fixed G , g grows as $|\omega_1| \rightarrow \infty$, and perturbation theory is dangerous unless $G \rightarrow 0$ to maintain a finite g . This is the most interesting case, as it may produce a flow to a different universality class. One can look at two typical possibilities. In the first, the conformal dimension $2h$ of φ is close to 2, and one may attempt a perturbation theory close to marginality, in the small parameter $2 - 2h$. Alternatively, one may perform a straightforward expansion in g . This expansion is *a priori* a safe one. For $2h$ a fractional number between 0 and 2, the ultraviolet behaviour is under control (except for a logarithmic divergence at second order when $2h = 1$, the case encountered above). The infrared singularities are suppressed by the finite geometry, and the series is likely to be convergent, as the physical quantities are analytic in g under those circumstances (this is exhibited in the previous example). We require only the correlations of the φ field

at criticality since

$$\begin{aligned} Z_0(g, \text{Im} \tau) &= e^{\frac{1}{6} \pi \text{Im} \tau c} \sum_{n=0}^{\infty} \left(\frac{g}{2\pi} \right)^n \int_{\rho \leq |x_1| \leq \dots \leq |x_n| \leq 1} \\ &\quad \prod_1^n \frac{d^2 x}{(x_1 \bar{x}_1)^{1-h}} \langle 0 | \varphi(x_1, \bar{x}_1) \dots \varphi(x_n, \bar{x}_n) | 0 \rangle \end{aligned} \quad (331)$$

If the Hamiltonian H has a ground state for $g \neq 0$, we presume that $\ln Z / \text{Im} \tau$ will have a finite limit as $\text{Im} \tau \rightarrow +\infty$. This justifies the definition of a running central charge according to

$$C(g) = 12 \lim_{\rho \rightarrow 0} \left\{ \frac{\ln Z_0(g, \rho)}{\ln 1/\rho} \right\} \quad (332)$$

If we assume that odd correlations vanish in the expansion (331), this will mean

$$C(g) = c + g^2 C_2 + g^4 C_4 + \dots \quad (333)$$

with the zeroth order term being the usual central charge. Inserting the two-point correlation

$$\langle 0 | \varphi(x_1, \bar{x}_1) \varphi(x_2, \bar{x}_2) | 0 \rangle = \frac{1}{|x_{12}|^{4h}} \quad (334)$$

which normalizes the coupling g , we derive that

$$C_2 = 6 \sum_{n=0}^{\infty} \frac{1}{n+h} \left(\frac{\Gamma(n+2h)}{n! \Gamma(2h)} \right)^2 = 6 \int_0^1 d\lambda \lambda^{h-1} F(2h, 2h; 1; \lambda) \quad (335)$$

meaningful for $2h < 1$. It is possible to recover in this way the running central charge of the Ising case for a thermal perturbation as computed previously (with due care paid to the logarithmic singularity in C_2) by performing similar calculations to all orders. One can also compute the operator dimensions along the flow. We leave it to the reader to pursue the matter further. It would of course be quite valuable to find a compact form for these expressions instead of power series.

Appendix 9.A Jacobian θ -series and products

We collect here some formulas on elliptic θ -functions. We use a lattice of translations Λ generated by 1 and τ such that $\text{Im} \tau > 0$.

We have constantly set $q = \exp 2i\pi\tau$, $|q| < 1$. With z a variable in the complex plane, we also write $y = \exp 2i\pi z$. The four Jacobi θ -functions defined through

$$\begin{aligned}\theta_1(z, \tau) &= -i \sum_{-\infty}^{+\infty} (-1)^n y^{n+\frac{1}{2}} q^{(n+\frac{1}{2})^2/2} \\ &= -iy^{\frac{1}{2}} q^{\frac{1}{8}} \prod_1^{\infty} (1-q^n) \prod_0^{\infty} (1-yq^{n+1})(1-y^{-1}q^n) \\ \theta_2(z, \tau) &= \sum_{-\infty}^{+\infty} y^{n+\frac{1}{2}} q^{(n+\frac{1}{2})^2/2} \\ &= y^{\frac{1}{2}} q^{\frac{1}{8}} \prod_1^{\infty} (1-q^n) \prod_0^{\infty} (1+yq^{n+1})(1+y^{-1}q^n) \quad (A.1) \\ \theta_3(z, \tau) &= \sum_{-\infty}^{+\infty} y^n q^{\frac{1}{2}n^2} \\ &= \prod_1^{\infty} (1-q^n) \prod_0^{\infty} (1+yq^{n+\frac{1}{2}})(1+y^{-1}q^{n+\frac{1}{2}}) \\ \theta_4(z, \tau) &= \sum_{-\infty}^{+\infty} (-1)^n y^n q^{\frac{1}{2}n^2} \\ &= \prod_1^{\infty} (1-q^n) \prod_0^{\infty} (1-yq^{n+\frac{1}{2}})(1-y^{-1}q^{n+\frac{1}{2}})\end{aligned}$$

are all entire functions of z . Their behaviour under translations in Λ is readily ascertained, as well as the location of their zeroes. The equality of the sums with the products is the content of Jacobi's triple product identity. The latter is a source of a large number of identities among which Euler's pentagonal identity

$$P(q) \equiv \prod_1^{\infty} (1-q^n) = \sum_{-\infty}^{+\infty} (-1)^k q^{\frac{1}{2}k(3k+1)} \quad (A.2)$$

Other useful identities include (Euler)

$$\prod_1^{\infty} (1-q^{2n-1})(1+q^n) = 1 \quad (A.3)$$

as well as Watson's quintuple product

$$\begin{aligned}\prod_1^{\infty} (1-q^n)(1-yq^n)(1-y^{-1}q^{n-1})(1-y^2q^{2n-1})(1-y^{-2}q^{2n-1}) \\ = \sum_{-\infty}^{+\infty} (y^{3m} - y^{-3m-1}) q^{\frac{1}{2}m(3m+1)}\end{aligned} \quad (A.4)$$

Dedekind's function is

$$\eta(\tau) = q^{\frac{1}{24}} P(q) \quad (A.5)$$

The function θ_1 vanishes at the origin. With prime meaning derivative with respect to z ,

$$\frac{1}{2\pi} \theta_1'(0, \tau) = \frac{1}{2} \theta_2(0, \tau) \theta_3(0, \tau) \theta_4(0, \tau) = \eta^3(\tau) \quad (A.6)$$

The quantities $\theta_i(0, \tau)$ ($2 \leq i \leq 4$) are sometimes called the θ -constants. One has

$$\begin{aligned}\frac{\theta_2(0, \tau)}{\eta(\tau)} &= 2 \left[\frac{\eta(2\tau)}{\eta(\tau)} \right]^2 = 2 \left[q^{\frac{1}{16} - \frac{1}{48}} \prod_1^{\infty} (1+q^n) \right]^2 \\ \frac{\theta_3(0, \tau)}{\eta(\tau)} &= \left[e^{-\frac{2i\pi}{48}} \frac{\eta(\frac{1}{2}(\tau+1))}{\eta(\tau)} \right]^2 = \left[q^{-\frac{1}{48}} \prod_1^{\infty} (1+q^{n-\frac{1}{2}}) \right]^2 \\ \frac{\theta_4(0, \tau)}{\eta(\tau)} &= \left[\frac{\eta(\frac{1}{2}\tau)}{\eta(\tau)} \right]^2 = \left[q^{-\frac{1}{48}} \prod_1^{\infty} (1-q^{n-\frac{1}{2}}) \right]^2\end{aligned} \quad (A.7)$$

They satisfy

$$[\theta_3(0, \tau)]^4 = [\theta_2(0, \tau)]^4 + [\theta_4(0, \tau)]^4 \quad (A.8a)$$

This is equivalent to

$$\begin{aligned}2 \sum_{\substack{n_1, n_2, n_3, n_4 \\ \sum n_i \text{ odd}}} q^{(n_1^2 + n_2^2 + n_3^2 + n_4^2)} \\ = \sum_{n_1, n_2, n_3, n_4} q^{(n_1 + \frac{1}{2})^2 + (n_2 + \frac{1}{2})^2 + (n_3 + \frac{1}{2})^2 + (n_4 + \frac{1}{2})^2}\end{aligned} \quad (A.8b)$$

expressing that on a hypercubic four-dimensional lattice $\mathcal{L}$ the centers of hypercubes can be divided in two integral lattices $\mathcal{L}^{\pm}$ generated by the vectors $(\frac{1}{2}\epsilon_1, \frac{1}{2}\epsilon_2, \frac{1}{2}\epsilon_3, \frac{1}{2}\epsilon_4)$ with $\epsilon_i = \pm 1$ and $\prod_i \epsilon_i = \pm 1$. These two lattices are related by a symmetry with respect to any hyperplane perpendicular to an axis of coordinates

and each one is isometric to the sublattice $\mathcal{L}_{\text{odd}} \subset \mathcal{L}$ ($\mathcal{L}_{\text{odd}}$ with corresponding to an odd sum of coordinates).

At some point, the following Gauss sum is useful, with ω standing for the primitive N th root of unity $\omega = e^{2i\pi/N}$

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \omega^{n^2} = \frac{1}{2}(1+i) + (-i)^{N-1} \frac{1}{2}(1-i) \quad (\text{A.9})$$

Equivalently, when $N \equiv 0, 1, 2, 3 \pmod{4}$ the sum is $1+i, 1, 0, i$ respectively.

The modular transformation properties follow from Poisson's formula. Set

$$\theta(\zeta, \tau) \equiv \theta_2(\zeta - \frac{1}{2}, \tau) = \sum e^{2i\pi(n+\frac{1}{2})(\zeta-\frac{1}{2})} q^{\frac{1}{2}(n+\frac{1}{2})^2} \quad (\text{A.10})$$

Then, for a unimodular matrix $A = \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$\eta\left(\frac{a\tau+b}{c\tau+d}\right) = \varepsilon_A (c\tau+d)^{\frac{1}{2}} \eta(\tau) \quad (\text{A.11})$$

$$\theta(\zeta + \mu\tau + \nu, \tau) = (-1)^{\mu+\nu} \exp -2i\pi\left(\frac{1}{2}\mu^2\tau + \mu\zeta\right) \theta(\zeta, \tau) \quad (\text{A.12})$$

$$\theta\left(\frac{\zeta}{c\tau+d}; \frac{a\tau+b}{c\tau+d}\right) = \varepsilon_A^3 (c\tau+d)^{\frac{1}{2}} \exp\left(i\pi c \frac{\zeta^2}{c\tau+d}\right) \theta(\zeta, \tau) \quad (\text{A.13})$$

with ε_A a 24th root of unity.

The θ -functions enable one to construct doubly periodic meromorphic functions by ratios or derivatives. The prototype is the Weierstrass function with

$$\begin{aligned} p(z, \tau) &= -\frac{\partial^2}{\partial z^2} \ln \theta_1(z, \tau) \\ &= \frac{1}{z^2} + \sum'_{n,m} \left[\frac{1}{(z+n+m\tau)^2} - \frac{1}{(n+m\tau)^2} \right] \\ &= \frac{1}{z^2} + 3z^2 \sum'_{n,m} \frac{1}{(n+m\tau)^4} + 5z^4 \sum'_{n,m} \frac{1}{(n+m\tau)^6} + \dots \end{aligned} \quad (\text{A.14})$$

an even meromorphic function with a double pole at the origin (as well as at lattice translates) and two zeroes in the period parallelogram. It fulfils the nonlinear first order differential equation, with prime denoting derivative with respect to z , and τ being fixed

$$\frac{1}{4}p'^2 = p^3 - 15e_4p - 35e_6 \quad (\text{A.15})$$

$$e_{2k}(\tau) = \sum'_{n,m} \frac{1}{(n+m\tau)^{2k}} \quad (\text{A.16})$$

The roots of the cubic polynomial on the r.h.s. in (A.15) are, by symmetry, the values $p(\frac{1}{2})$, $p(\frac{1}{2}\tau)$, and $p(\frac{1}{2}(1+\tau))$. Its discriminant, conventionally multiplied by 2^4 , is related to the η -function through

$$\begin{aligned} 2^4 [(p(\frac{1}{2}) - p(\frac{1}{2}\tau))(p(\frac{1}{2}) - p(\frac{1}{2}(1+\tau)))(p(\frac{1}{2}\tau) - p(\frac{1}{2}(1+\tau)))]^2 \\ = 2^4 3^5 5^2 (20e_4^3 - 49e_6^2) = (2\pi)^{12} \eta(\tau)^{24} \end{aligned} \quad (\text{A.17})$$

Appendix 9.B Superconformal algebra

We have seen in various circumstances how commuting and anticommuting variables appear on equal footing. Here we describe a natural generalization of conformal transformations to the supercomplex plane parametrized by a pair of variables $\xi \equiv (z, \theta)$ and $\bar{\xi} \equiv (\bar{z}, \bar{\theta})$, with θ and $\bar{\theta}$ anticommuting.

An analytic superfield is a function of z and θ only, hence of the form

$$\phi(\xi) = \varphi_0(z) + \theta \varphi_1(z) \quad (\text{B.1})$$

A natural square root of the Cauchy-Riemann operator is defined as

$$D\phi(\xi) \equiv \left(\frac{\partial}{\partial \theta} + \theta \frac{\partial}{\partial z} \right) \phi(\xi) = \varphi_1(z) + \theta \frac{\partial}{\partial z} \varphi_0(z) \quad (\text{B.2})$$

with

$$D^2\phi(\xi) = \frac{\partial}{\partial z} \phi(\xi) \quad (\text{B.3})$$

Henceforth we write ∂ for $\partial/\partial z$. A supercontour integral means

$$\oint_C d\xi \omega(\xi) \equiv \oint_C dz \int d\theta [\omega_0(z) + \theta \omega_1(z)] = \oint_C dz \omega_1(z) \quad (\text{B.4})$$

while a definite integral

$$f(\xi_1, \xi_2) = \int_{\xi_2}^{\xi_1} d\xi \omega(\xi) \quad (\text{B.5})$$

is such that

$$D_1 f(\xi_1, \xi_2) = \omega(\xi_1) \quad f(\xi_1, \xi_1) = 0 \quad (B.6)$$

Thus

$$\int_{\xi_2}^{\xi_1} d\xi \equiv \theta_{12} = \theta_1 - \theta_2 \int_{\xi_2}^{\xi_1} d\xi \int_{\xi_2}^{\xi} d\xi' \equiv \xi_{12} = z_1 - z_2 - \theta_1 \theta_2 \quad (B.7)$$

A superanalytic function admits a Taylor expansion

$$f(\xi_1) = \sum_0^{\infty} \frac{1}{n!} \xi_{12}^n \partial_2^n [1 + \theta_{12} D_2] f(\xi_2) \quad (B.8)$$

Since

$$\oint_{C_2} \frac{d\xi_1}{2i\pi} \xi_{12}^{-n-1} = 0 \quad \oint_{C_2} \frac{d\xi_1}{2i\pi} \theta_{12} \xi_{12}^{-n-1} = \delta_{n,0} \quad (B.9)$$

the residue theorems are

$$\begin{aligned} \oint_{C_2} \frac{d\xi_1}{2i\pi} f(\xi_1) \theta_{12} \xi_{12}^{-n-1} &= \frac{1}{n!} \partial_2^n f(\xi_2) \\ \oint_{C_2} \frac{d\xi_1}{2i\pi} f(\xi_1) \xi_{12}^{-n-1} &= \frac{1}{n!} \partial_2^n D_2 f(\xi_2) \end{aligned} \quad (B.10)$$

Superconformal transformations are such that the map $\xi \rightarrow \xi'$ satisfies

$$D = D(\theta') D' \quad (B.11)$$

meaning that D transforms homogeneously. This is not the case of the general superanalytic transformations where

$$D = (D\theta') D' + [D(z') - \theta' D(\theta')] D'^2$$

Global superconformal transformations depend on three complex parameters (z_0, z_1, w_0) and two anticommuting ones θ_0, θ_1 . They can be written

$$\xi \rightarrow \xi' \quad z' = z_0 - \frac{w_0 - \theta_0 \theta_1}{z - z_1 - \theta \theta_1} \quad \theta' = \theta_0 + \frac{\theta - \theta_1}{z - z_1 - \theta \theta_1} \quad (B.12)$$

This defines a supersymmetric extension of the group $SL(2, \mathbb{C})$. Local superconformal transformations satisfying (B.11) corresponding to a supervector field $v(\xi)$ are generated by the super-energy-momentum tensor

$$T(\xi) = T_F(z) + \theta T_B(z) \quad (B.13)$$

in the form

$$\delta_v \phi(\xi_2) = \oint_{C_2} d\xi_1 v(\xi_1) T(\xi_1) \phi(\xi_2) \quad (B.14)$$

In formula (B.13), $T_B(z)$ is the ordinary bosonic tensor of dimension 2 and $T_F(z)$ is its fermionic counterpart of dimension $\frac{3}{2}$, with the understanding that θ counts dimensionally as $z^{\frac{1}{2}}$. If ϕ is a primary superfield of dimension h , we have

$$\delta_v \phi = [v\partial + \frac{1}{2} D(v) D + h(\partial v)] \phi \quad (B.15)$$

so that (B.14) is equivalent to the short-distance expansion

$$T(\xi_1) \phi(\xi_2) \sim h \frac{\theta_{12}}{\xi_{12}^2} \phi(\xi_2) + \frac{1}{2\xi_{12}} D_2 \phi(\xi_2) + \frac{\theta_{12}}{\xi_{12}} \partial_2 \phi(\xi_2) + \dots \quad (B.16)$$

The anomalous central charge appears in the corresponding formula for the product

$$T(\xi_1) T(\xi_2) \sim \frac{1}{6} \frac{c}{\xi_{12}^3} + \left[\frac{3}{2} \frac{\theta_{12}}{\xi_{12}^2} + \frac{1}{2\xi_{12}} D_2 + \frac{\theta_{12}}{\xi_{12}} \partial_2 \right] T(\xi_2) + \dots \quad (B.17)$$

or

$$\delta_v T = [v\partial + \frac{1}{2} D(v) D + \frac{3}{2} \partial v] T + \frac{1}{12} c \partial^2 D v \quad (B.18)$$

In finite form (B.15) and (B.18) read

$$\begin{aligned} \phi(\xi) &= \phi'(\xi') (D\theta')^{2h} \\ T(\xi) &= T'(\xi') (D\theta')^3 + \frac{1}{12} c \{\xi', \xi\} \end{aligned} \quad (B.19)$$

with the super-Schwarzian derivative written as

$$\{\xi', \xi\} = 2 \left(\frac{D^4 \theta'}{D\theta'} - 2 \frac{D^3 \theta'}{D\theta'} \frac{D^2 \theta'}{D\theta} \right) \quad (B.20)$$

For a particular superconformal transformation

$$z' = f(z), \quad \theta' = (\partial f(z)/\partial z)^{\frac{1}{2}} \theta$$

$\{\xi', \xi\}$ reduces to $\theta \{z', z\}$ with $\{z', z\}$ the ordinary Schwarzian derivative. The Laurent expansion of T

$$T_F(z) + \theta T_B(z) = \sum \left(\frac{1}{2} \frac{G_n}{z^{n+\frac{1}{2}}} + \frac{\theta L_n}{z^{n+2}} \right) \quad (B.21)$$

leads in the radial quantization scheme to the following set of commutation and anticommutation relations (with the notation

$$[A, B]_+ = AB + BA$$

$$\begin{aligned} [L_m, L_n] &= (m-n)L_{m+n} + \frac{1}{12}cm(m^2-1)\delta_{m+n,0} \\ [L_m, G_n] &= (\tfrac{1}{2}m-n)G_{m+n} \\ [G_m, G_n]_+ &= 2L_{m+n} + \tfrac{1}{3}c(m^2-\tfrac{1}{4})\delta_{m+n,0} \end{aligned} \quad (B.22)$$

In (B.21) the sum over indices for the Virasoro generators runs on integers as usual. But for the G_n partners, there are two choices defining what are called the Neveu-Schwarz (N-S) sector, with $n \in \mathbb{Z} + \frac{1}{2}$ leading to single valued spinor fields (as well as T_F) in the z -plane, or the Ramond (R-) sector, with $n \in \mathbb{Z}$ and double valued fermionic fields (as well as T_F) in the z -plane. These properties are reversed in a periodic strip, according to (B.19) where N-S fermionic fields are antiperiodic, R-fermionic fields periodic. Thus we have in fact two distinct superconformal algebras.

We have only discussed the ξ -dependence of fields. Of course similar properties hold with respect to $\bar{\xi}$. For short and with some abuse of language, we continue to omit this parallel dependence.

The operator L_0 acts as a Hamiltonian with the vacuum $|0\rangle$ being in the N-S sector, annihilated by the five generators $L_0, L_{\pm 1}, G_{\pm 1/2}$ of global superconformal transformations. The highest weight vectors $|h\rangle$ of the N-S sector are then in one-to-one correspondence with the superfields $\phi_h(\xi)$ through $|h\rangle = \phi_h(0)|0\rangle$.

The R-sector has the additional feature that G_0 acts as a supersymmetric charge. It commutes with L_0 and $G_0^2 = L_0 - \frac{1}{24}c$. Thus, in general, highest weight R-states come in pairs $|h^\pm\rangle$ with $|h^-\rangle = G_0|h^+\rangle$. One calls spin fields, those ordinary conformal fields $\theta^\pm(z)$ which intertwine the N-S and R sectors by generating the states $|\dot{h}^\pm\rangle$ out of the vacuum, $\theta^\pm(0)|0\rangle = |\dot{h}^\pm\rangle$. The condition for the existence of an unbroken supersymmetry generated by G_0 in the R-sector is that $G_0|h^+\rangle = 0$ (in other words the Verma module generated by $|h^+\rangle$ decouples from the theory) and this is possible if and only if $h^+ = \frac{1}{24}c$.

Let F count the number of fermionic fields, then $\Gamma = (-1)^F$ is called the chirality operator. Paired R-states have opposite chirality as G_0 anticommutes with Γ . Spin fields of equal chirality are relatively local, while the opposite is true for spin fields of unequal chirality.

We restrict ourselves to the description of unitary representations of the superconformal algebras. If $c \geq \frac{3}{2}$ and $h \geq 0$ all rep-

resentations afforded by Verma modules are equivalent to unitary ones. However when $c < \frac{3}{2}$ there only exists a discrete sequence of (degenerate) unitary representations. They are similar to the ones found in the unitary case for the ordinary conformal algebra. Let M denote an integer larger or equal to 2, then for this unitary series, in either the N-S or R-case the central charge is given by

$$c = \frac{3}{2} \left(1 - \frac{8}{M(M+2)} \right) \quad (B.23)$$

while the weights belong to the finite table

$$h_{r,s} = \frac{[(M+2)r - Ms]^2 - 4}{8M(M+2)} + \frac{1}{32}(1 - (-1)^{r-s}) \quad (B.24)$$

$$1 \leq r < M \quad 1 \leq s < M+2$$

Here $r-s$ is even for N-S representations and odd for R-representations. For $M=2$ we have the trivial representation. When M is odd, $h = \frac{1}{24}c$ is not in the table, so that supersymmetry is broken in the R-sector, while for M even, $h_{\frac{1}{2}M, \frac{1}{2}M+1} = \frac{1}{24}c$ and supersymmetry can be preserved.

More generally one has formulas for the determinants of the contragredient forms as well as for characters. In particular for the unitary representations (B.24), there are three types of characters. With the central charge understood, $q = e^{2i\pi\tau}$, and r and s labelling the weights, we denote them $\chi^{NS}, \bar{\chi}^{NS}, \chi^R$

$$\begin{aligned} \chi_{r,s}^{NS}(\tau) &= \text{Tr}_{r,s} q^{L_0} \quad r-s \equiv 0 \pmod{2} \\ &= \prod_{p=1}^{\infty} \left(\frac{1+q^{p-\frac{1}{2}}}{1-q^p} \right) \sum_{n=-\infty}^{+\infty} \left\{ q^{\{[2nM(M+2)+r(M+2)-sM]^2-4\}/8M(M+2)} - (s \leftrightarrow -s) \right\} \\ \chi_{r,s}^{NS}(\tau) &= \text{Tr}(-1)^F q^{L_0} \quad r-s \equiv 0 \pmod{2} \\ &= \prod_{p=1}^{\infty} \left(\frac{1-q^{p-\frac{1}{2}}}{1-q^p} \right) \sum_{n=-\infty}^{+\infty} (-1)^{Mn} \left\{ q^{\{[2nM(M+2)+r(M+2)-sM]^2-4\}/8M(M+2)} - (-1)^{rs}(s \leftrightarrow -s) \right\} \\ \chi_{r,s}^R(\tau) &= \text{Tr} q^{L_0} \quad r-s \equiv 1 \pmod{2} \\ &= q^{\frac{1}{16}} \prod_{p=1}^{\infty} \left(\frac{1+q^p}{1-q^p} \right) \sum_{n=-\infty}^{+\infty} \end{aligned}$$

$$\left\{ q^{\{2nM(M+2)+r(M+2)-sM\}^2-4}/8M(M+2) - (s \leftrightarrow -s) \right\} \quad (B.25)$$

A classification of consistent minimal models (in particular unitary ones) can be achieved along the lines of section 3.4 using modular invariance on a torus. Let us just show here how the formalism relates to statistical physics by recording the interpretation of the first nontrivial model in the list (B.23) corresponding to $M = 3$ (Friedan, Qiu and Shenker). It turns out to be unique, and to coincide with the case $m = 4$ for the ordinary conformal classification. In both cases $c = \frac{7}{10}$ and the possible weights in the Virasoro case were (figure 4) $h = 0, \frac{7}{16}, \frac{3}{2}, \frac{3}{80}, \frac{3}{5}$ and $\frac{1}{10}$, while, according to the stronger superconformal invariance, we have $h_{1,1} = 0, h_{2,2} = \frac{1}{10}$ in the N-S sector and $h_{1,2} = \frac{3}{80}, h_{2,1} = \frac{7}{16}$ in the R-sector.

This can be identified with the tricritical Ising model. Indexing by V the ordinary Virasoro representations, the Z_2 even sector is spanned by the two N-S representations

$$\begin{aligned} (0)_{N-S} &= (0)_V \oplus \left(\frac{3}{2}\right)_V \\ \left(\frac{1}{10}\right)_{N-S} &= \left(\frac{1}{10}\right)_V \oplus \left(\frac{3}{5}\right)_V \end{aligned} \quad (B.26)$$

while the Z_2 odd sector is the R -one, containing magnetic like operators, with

$$\begin{aligned} \left(\frac{3}{80}\right)_R &= \left(\frac{3}{80}\right)_V \\ \left(\frac{7}{16}\right)_R &= \left(\frac{7}{16}\right)_V \end{aligned} \quad (B.27)$$

Appending the usual prefactor $q^{-c/24}$ in the characters, one has

$$\begin{aligned} \chi_0^V \pm \chi_{\frac{3}{2}}^V &= \left\{ \begin{array}{l} \chi_0^{N-S} \\ \chi_{\frac{3}{2}}^{N-S} \end{array} \right. = q^{-\frac{7}{240}} \prod_{n=0}^{\infty} \frac{(1 \pm q^{5n+\frac{1}{2}})(1 \pm q^{5n+\frac{3}{2}})(1 \pm q^{5n+\frac{5}{2}})}{(1 - q^{5n+3})(1 - q^{5n+4})} \\ \chi_{\frac{1}{10}}^V \pm \chi_{\frac{3}{5}}^V &= \left\{ \begin{array}{l} \chi_{\frac{1}{10}}^{N-S} \\ \chi_{\frac{3}{5}}^{N-S} \end{array} \right. = q^{-\frac{7}{240} + \frac{1}{10}} \prod_{n=0}^{\infty} \frac{(1 \pm q^{5n+\frac{1}{2}})(1 \pm q^{5n+\frac{3}{2}})(1 \pm q^{5n+\frac{5}{2}})}{(1 - q^{5n+1})(1 - q^{5n+4})} \\ \chi_{\frac{3}{80}}^V &= \chi_{\frac{3}{80}}^R = q^{-\frac{7}{240} + \frac{3}{80}} \prod_{n=0}^{\infty} \frac{(1 + q^{5n+2})(1 + q^{5n+3})(1 + q^{5n+5})}{(1 - q^{5n+1})(1 - q^{5n+4})} \\ \chi_{\frac{7}{16}}^V &= \chi_{\frac{7}{16}}^R = q^{-\frac{7}{240} + \frac{7}{16}} \prod_{n=0}^{\infty} \frac{(1 + q^{5n+1})(1 + q^{5n+4})(1 + q^{5n+5})}{(1 - q^{5n+2})(1 - q^{5n+3})} \end{aligned} \quad (B.28)$$

On a torus with modular ratio τ , we have the corresponding modular invariant partition function

$$\begin{aligned} Z(\tau, \bar{\tau}) &= \left\{ \frac{1}{2} \left(|\chi_0^{N-S}(\tau)|^2 + |\bar{\chi}_0^{N-S}(\tau)|^2 + |\chi_{\frac{1}{10}}^{N-S}(\tau)|^2 + |\bar{\chi}_{\frac{1}{10}}^{N-S}(\tau)|^2 \right) \right. \\ &\quad \left. + |\chi_{\frac{3}{80}}^R(\tau)|^2 + |\bar{\chi}_{\frac{3}{80}}^R(\tau)|^2 \right\} \\ &= \left\{ |\chi_0^V(\tau)|^2 + |\chi_{\frac{1}{10}}^V(\tau)|^2 + |\chi_{\frac{3}{5}}^V(\tau)|^2 + |\chi_{\frac{3}{2}}^V(\tau)|^2 \right. \\ &\quad \left. + |\chi_{\frac{3}{80}}^V(\tau)|^2 + |\chi_{\frac{7}{16}}^V(\tau)|^2 \right\} \end{aligned} \quad (B.29)$$

Appendix 9.C Current algebra

Conformal field theory extends to systems with continuous symmetries. An example is afforded by generalized σ -models, with fields taking their values in a Lie group G , when the kinetic Lagrangian is supplemented by a topological term to implement a local $G_x \times G_x$ invariance generated by a pair of conserved currents $J(z), \bar{J}(\bar{z})$ ($\partial J = \partial \bar{J} = 0$). The energy-momentum tensor is given by a quadratic expression in terms of the currents, with conformal weights (1,0) and (0,1) respectively. The coefficients of J and $\bar{J}$, in a Laurent expansion, understood as operators, generate a pair of commuting Kac-Moody (or affine) infinite Lie algebras. In the text we encountered as an auxiliary device the characters of the affine $SU(2)$ algebra and came across the $A-D-E$ classification of minimal models. All this justifies that we succinctly describe here some aspects of the interplay between group theory and conformal invariance.

C.1. Simple Lie algebras

The reader is presumably familiar with the Cartan-Killing classification of simple Lie algebras over complex numbers and their finite dimensional representation. The following description can be complemented by the vast literature on the subject and is only intended to present some notation used in the text.

Simple Lie algebras come in four infinite families and five exceptional cases. The four families are realized as infinitesimal

transformations corresponding to the classical complex matrix groups according to the table

$\ell \geq 1$	A_ℓ	special linear group	$SL(\ell+1)$
$\ell \geq 2$	B_ℓ	odd dimensional orthogonal group	$SO(2\ell+1)$
$\ell \geq 3$	C_ℓ	symplectic group	$Sp(2\ell)$
$\ell \geq 4$	D_ℓ	even dimensional orthogonal group	$SO(2\ell)$

The restriction on the rank ℓ , the maximal number of linearly independent commuting elements, which span a Cartan subalgebra, is to avoid isomorphisms in low dimensions, like $A_1 \sim B_1 \sim C_1$, $C_2 \sim B_2$, $D_2 \sim A_1 \times A_1$, $D_3 \sim A_3$. The five exceptional cases are less familiar; they are G_2 , F_4 , E_6 , E_7 , E_8 , where the index gives the rank.

All Cartan, i.e. maximal abelian subalgebras are conjugate in the group. The elements in one of them, denoted K , can be simultaneously diagonalized in any representation, in particular in the adjoint representation. The set of these eigenvalues are elements in the dual space K^* and are called roots in the case of the adjoint representation. Let e_α (an element in the Lie algebra) be the common eigenvector, this means that for any k belonging to K one has

$$[k, e_\alpha] = \langle \alpha | k \rangle e_\alpha \quad (C.1)$$

Simplicity of the Lie algebra implies that the roots span K^* . Moreover there exists an invariant symmetric nonsingular bilinear form on the algebra, denoted $(\ , \)$, i.e. such that

$$\begin{aligned} (a, b) &= (b, a) \\ (a, b) &= 0 \text{ for any } a \Rightarrow b = 0 \\ ([a, b], c) &= (a, [b, c]) \end{aligned} \quad (C.2)$$

Roots come in pairs $\{\alpha, -\alpha\}$, those being the only ones proportional to α . From Jacobi's identity, if two roots α and β are such that $\alpha + \beta \neq 0$ (i) if $\alpha + \beta$ is a root, $[e_\alpha, e_\beta]$ is proportional to $e_{\alpha+\beta}$, (ii) if $\alpha + \beta$ is not a root, $[e_\alpha, e_\beta] = 0$. If $\alpha + \beta = 0$, the elements

$$k_\alpha = [e_\alpha, e_{-\alpha}] \leftrightarrow (e_\alpha, e_{-\alpha}) \langle \alpha | \quad (C.3)$$

span K (the correspondence means that $(k_\alpha, k) = (e_\alpha, e_{-\alpha}) \langle \alpha | k \rangle$) with the e_α 's normalized in such a way that

$$(k_\alpha, k_\alpha) = 2(e_\alpha, e_{-\alpha}) \neq 0 \quad (C.4)$$

and the restriction of the bilinear form to K remaining nonsingular. The triplet $\{e_\alpha, e_{-\alpha}, k_\alpha\}$ generates an $SL(2)$ algebra

$$[k_\alpha, e_{\pm\alpha}] = \pm 2e_{\pm\alpha} \quad [e_\alpha, e_{-\alpha}] = k_\alpha \quad (C.5)$$

with a familiar realization as two by two matrices

$$k_\alpha \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad e_\alpha \rightarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad e_{-\alpha} \rightarrow \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (C.6)$$

and an induced bilinear form

$$\begin{aligned} (ak_\alpha + be_\alpha + ce_{-\alpha}, ak_\alpha + be_\alpha + ce_{-\alpha}) &= (k_\alpha, k_\alpha) \{a^2 + bc\} \\ &= \frac{1}{2} (k_\alpha, k_\alpha) \text{Tr} \begin{pmatrix} a & b \\ c & -a \end{pmatrix}^2 \end{aligned} \quad (C.7)$$

The element e_β is an eigenstate of k_α

$$\begin{aligned} [k_\alpha, e_\beta] &= A_{\alpha, \beta} e_\beta \\ A_{\alpha, \beta} &= \langle \beta | k_\alpha \rangle = \frac{2(k_\alpha, k_\beta)}{(k_\beta, k_\beta)} \end{aligned} \quad (C.8)$$

where $A_{\alpha, \beta}$ must be an integer as follows from studying the finite representation of the $SL_\alpha(2)$ algebra (C.5) acting on the states $e_{\beta+s\alpha}$, s integer (k_α is the analog of twice the z component of angular momentum and has therefore integral eigenvalues). Since under duality it follows from (C.3) and (C.4) that

$$\tilde{k}_\alpha = \frac{2k_\alpha}{(k_\alpha, k_\alpha)} \leftrightarrow \langle \alpha | \quad (C.9)$$

it is natural to transfer the bilinear form on the dual root space K^* through

$$(\alpha, \beta) = (\tilde{k}_\alpha, \tilde{k}_\beta) \quad (C.10)$$

in which case, we can also write

$$A_{\alpha, \beta} = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \quad (C.11)$$

Assuming $(\alpha, \beta) \neq 0$, the elements $e_{\beta+s\alpha}$ span a representation of $SL_\alpha(2)$ which must contain an element with a k_α eigenvalue opposite to the one of e_β . This is for

$$s = -\frac{2(\alpha, \beta)}{(\alpha, \alpha)}$$

with the result that the set of roots contains with every root β the one reflected in the hyperplane perpendicular to α

$$w_\alpha(\beta) = \beta + s\alpha = \beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \quad (C.12)$$

These reflections generate a point group called the Weyl group and relate the theory of simple Lie algebras with some aspects of crystallography. The Weyl group acts transitively on sets of basis roots $\{\alpha_1, \dots, \alpha_\ell\}$, also called simple roots, in terms of which all roots can be written $\alpha = \sum n_i \alpha_i$ with all n_i 's integers of the same sign (or zero). Accordingly the even number of roots is subdivided into positive or negative roots (denoted $\alpha > 0$ or $\alpha < 0$).

The 3ℓ elements $k_i \equiv k_{\alpha_i}$, $e_i^{(\pm)} \equiv e_{\pm\alpha_i}$, are generators for the Lie algebra and satisfy the following relations (Chevalley-Serre)

$$\begin{aligned} [k_i, k_j] &= 0 & [k_i, e_j^{(\pm)}] &= \pm A_{ij} e_j^{(\pm)} \\ [e_i^{(+)}, e_j^{(-)}] &= \delta_{ij} k_j & \\ (\text{ade}_i^{(+)})^{1-A_{ij}} e_j^{(+)} &= (\text{ade}_i^{(-)})^{1-A_{ij}} e_j^{(-)} = 0 & i \neq j \end{aligned} \quad (C.13)$$

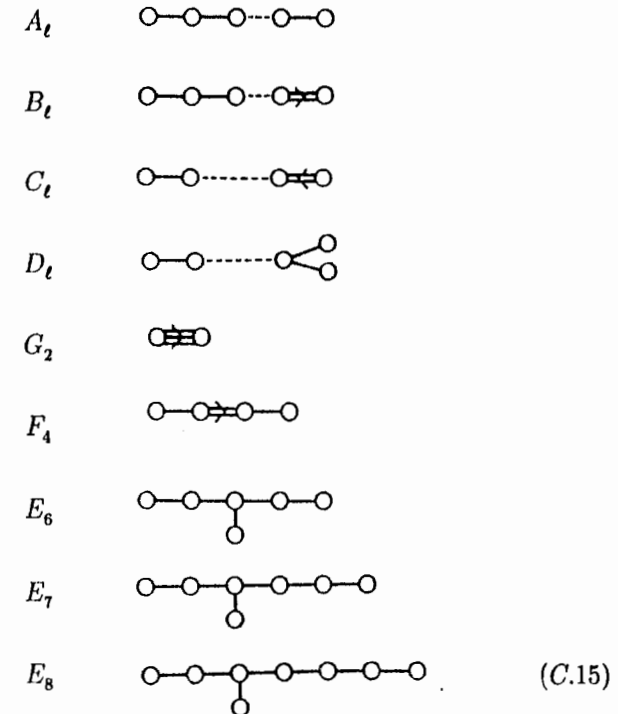
The $\ell \times \ell$ integral Cartan matrix

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} \quad (C.14)$$

therefore characterizes entirely the algebra up to simultaneous permutations of lines and columns. Its diagonal entries are equal to 2, and for $i \neq j$, $A_{ij} \neq 0$ implies $A_{ji} \neq 0$ (both being negative). The simply laced algebras A , D , E , are such that all roots are of equal length and as a consequence the Cartan matrix is symmetric.

The nonvanishing off-diagonal elements of A_{ij} are coded in a Coxeter-Dynkin diagram, where ℓ points are associated to the ℓ simple roots. Pairs of points $\{i, j\}$ are joined by $A_{ij}A_{ji} = 0, 1, 2, 3$ lines. If $A_{ij} = 0$ the corresponding roots are orthogonal and w_i and w_j commute. If the roots are not orthogonal $(w_i w_j)^{m_{ij}} = 1$, where corresponding to the possibilities of 1, 2 or 3 lines between

i and j , the integer m_{ij} equals 3, 4 or 6. The diagram is connected (otherwise the algebra would break in a direct sum of commuting subalgebras). The above description only characterizes the Weyl group, irrespective of the lengths of the roots. One shows that either the simple roots are of equal length, in which case there are no multiple lines in the diagram, or are subdivided into two groups, the long and short ones. Then if $A_{ij}A_{ji} = 1$, $(\alpha_i, \alpha_i) = (\alpha_j, \alpha_j)$ while if $(\alpha_i, \alpha_i) > (\alpha_j, \alpha_j)$ and $A_{ij} \neq 0$, then $A_{ji} = -1$ and A_{ij} equals -2 or -3 . Thus it is sufficient to put an arrow on multiple lines pointing from the long to the short root. The possible sets of such diagrams are then depicted as follows



The roots $\{\alpha\}$ generate a lattice over integers, called the root lattice. Similarly the co-roots $2\alpha/(\alpha, \alpha)$ generate another lattice isomorphic up to scale with the preceding one, except in the case of B_ℓ and C_ℓ where this duality exchanges the root lattices. These two algebras have therefore the same Weyl group.

The root lattice is invariant under the semidirect product of the Weyl group and a discrete translation group. An equivalent description is that this inhomogeneous group is generated by $n+1$ reflections in the hyperplanes bounding a simplex (which is a fundamental domain). With the origin at one of the vertices, ℓ of the hyperplanes are orthogonal to ℓ simple roots and the $(\ell+1)$ th hyperplane is orthogonal to a root α_0 expressed by an equation $\sum_0^\ell y_i \alpha_i = 0$, $y_0 = 1$, with integral positive coefficients. These characteristic integers y_i obey several interesting relations. Their sum $h = \sum_0^\ell y_i$, called the Coxeter number, is related to the dimension of the Lie algebra, denoted $\dim$, through

$$\dim = \ell(1 + h) \quad (C.16)$$

so that ℓh is even and equal to the number of roots. The notation h is standard and should not be confused with conformal weights. On the other hand, the product of y 's is related to dimension of the Weyl group $|W|$ through

$$|W| = \ell! \left(\prod_0^\ell y_j \right) \det A_{ij} \quad (C.17)$$

One defines analogous concepts for the co-roots and in particular a dual Coxeter number, denoted g .

The significance of the determinant of the Cartan matrix appears when one investigates the irreducible, finite-dimensional, representations of the Lie algebra. Those are associated to points of the lattice dual (over the integers $\mathcal{Z}$) of the co-root lattice, and are called weights. Thus for any weight q and any root α , $2(q, \alpha)/(\alpha, \alpha)$ is an integer. The weight lattice is also invariant under the Weyl group and contains as a sublattice the root lattice with an index equal to the determinant of the Cartan matrix. It is generated by ℓ fundamental weights q_i , $1 \leq i \leq \ell$, satisfying

$$\frac{2(q_i, \alpha_j)}{(\alpha_j, \alpha_j)} = \delta_{ij} \quad (C.18)$$

Dominant weights have nonnegative scalar products with the simple roots and any weight can be mapped on a dominant one by an element in the Weyl group. The sum q_m of fundamental

weights satisfies

$$q_m = \sum_1^\ell q_i = \frac{1}{2} \sum_{\alpha > 0} \alpha \quad (C.19)$$

Finite-dimensional irreducible representations have a highest weight state (with a highest weight, a dominant one, denoted q) annihilated by all e_i^{+} . The weights q_μ associated to the eigenvalues of the generators of the Cartan subalgebra (i.e. the eigenvalue of k_α is $2(\alpha, q_\mu)/(\alpha, \alpha)$) are of the form $q - \sum \mu_i \alpha_i$, where the μ_i 's are nonnegative integers. The weights of the adjoint representation are the roots, and the highest root is $\psi = -\alpha_0$, and is a long root.

The corresponding characters χ_q are given by a formula due to Weyl, as a function of a ℓ -dimensional vector x

$$\begin{aligned} \chi_q(x) &= \sum_{q'} \text{mult}(q') e^{2i\pi(q', x)} \\ &= \frac{\psi_{q_m + q}(x)}{\psi_{q_m}(x)} \\ \psi_q(x) &= \sum_{w \in W} \varepsilon(w) e^{2i\pi(w(q), x)} \end{aligned} \quad (C.20)$$

Here $\text{mult}(q')$ is the multiplicity of the weight q' in the representation, and, in the sum over the Weyl group, $\varepsilon(w)$ is plus or minus one according to the parity of its expression in terms of generating reflections. The denominator admits a factorized form

$$\psi_{q_m}(x) = \prod_{\alpha > 0} (e^{i\pi(\alpha, x)} - e^{-i\pi(\alpha, x)}) \quad (C.21)$$

Taking the limit $x \rightarrow 0$, yields the dimension of the representation

$$\dim_q = \prod_{\alpha > 0} \frac{(\alpha, q + q_m)}{(\alpha, q_m)} \quad (C.22)$$

The quadratic Casimir invariant (up to an arbitrary scale)

$$(q + q_m, q + q_m) - (q_m, q_m) = (q, q + 2q_m) \quad (C.23)$$

is insufficient to characterize the representation when $\ell > 1$. Among polynomials in the components of a weight, it is possible to single out those which are invariant under the Weyl group as being themselves expressible as polynomials in ℓ -fundamental homogeneous ones, which generalize the quadratic form (q, q) .

The degrees of these polynomials are recorded in the last column of table III. The highest degree is the Coxeter number h , the product over all degrees is the order of the Weyl group, and twice their sum is $\dim(G) + \ell$, where $\dim(G)$ is the dimension of the Lie algebra. It will be recognized that, if one subtracts 1 from these degrees (obtaining what are called exponents), one recovers the nonzero diagonal entries (with their multiplicity) appearing in the A - D - E invariant partition functions discussed in the text. The table also includes a column for the dual Coxeter number g which is the value of the quadratic Casimir invariant (C.23) in the adjoint representation when one normalizes the long roots to unity. Of course $h = g$ for simply laced groups. One has also the following formula (Freudenthal)

$$\dim(G) = \frac{24}{g} \frac{(\alpha_m, \alpha_m)}{(\alpha_\ell, \alpha_\ell)} \quad (\text{C.24})$$

where $(\alpha_\ell, \alpha_\ell)$ is the square length of a long root. The table includes a column for the dimension of the algebra, one for the order of the Weyl group, as well as one for the structure of the quotient of the weight to the root lattice considered as additive groups (the order being $\det A_{ij}$).

One could proceed to infinite Kac-Moody algebras by extending the previous description to degenerate Cartan matrices, or in a global fashion to loop groups (i.e. mappings from the circle into a Lie group) and their central extensions. We shall motivate this generalization by presenting the Wess-Zumino-Witten model. In parallel with the implementation of local scale invariance, the latter promotes a global symmetry to a local one in a tight conformally invariant structure.

C.2. The Wess-Zumino-Witten model

The asymptotically free two-dimensional nonlinear σ -model is a massive field theory with a global continuous invariance group. In an attempt to find a nontrivial infrared fixed point, Witten was led in the case where the field variable u takes its values in a compact Lie group G , to add a topological term which suggests perturbatively the existence of such a critical theory. To simplify matters, we shall assume that the compact Lie group has a simple Lie algebra (which will then be extended over complex numbers).

Table III. Data for simple Lie algebras.

Algebra	dim	$ W $	Weight/Root	g	degrees of invariants
A_ℓ	$\ell(\ell+2)$	$(\ell+1)!$	$Z/(\ell+1)Z$	$\ell+1$	$2, 3, \dots, \ell+1$
B_ℓ	$\ell(2\ell+1)$	$2^\ell \ell!$	$Z/2Z$	$2\ell-1$	$2, 4, \dots, 2\ell$
C_ℓ	$\ell(2\ell+1)$	$2^\ell \ell!$	$Z/2Z$	$\ell+1$	$2, 4, \dots, 2\ell$
D_ℓ	$\ell(2\ell-1)$	$2^{\ell-1} \ell!$	ℓ even: $Z/2Z \times Z/2Z$ ℓ odd: $Z/4Z$	$2(\ell-1)$	$2, 4, \dots, 2(\ell-1), 2\ell$
G_2	14	12	id.	4	$2, 6, 8, 12$
F_4	52	1152	id.	9	$2, 5, 6, 8, 9, 12$
E_6	78	51840	$Z/3Z$	12	$2, 5, 6, 8, 9, 12$
E_7	133	2903040	$Z/2Z$	18	$2, 6, 8, 10, 12, 14, 18$
E_8	248	696729600	id.	30	$2, 8, 12, 14, 18, 20, 24, 30$

The topological term has a structure typical of an anomaly arising when studying fermions coupled to gauge fields. We shall not emphasize this aspect here, although part of the motivation for this bosonic construction is to find an equivalence with a fermionic one.

The starting point is the standard action

$$S^{(0)} = \frac{1}{2\lambda} \int \frac{d^2x}{2\pi} \widetilde{\text{Tr}} \partial_a u^{-1} \partial_a u \quad \lambda > 0 \quad (C.25)$$

To give a meaning to this expression, we assume that u takes its values in a faithful unitary representation of G and $\widetilde{\text{Tr}}$ is the trace in this representation up to an appropriate factor (hence the tilde). Since

$$\widetilde{\text{Tr}} \partial_a u^{-1} \partial_a u = -\widetilde{\text{Tr}}(u^{-1} \partial_a u)(u^{-1} \partial_a u) \quad (C.26)$$

and for u unitary, $u^{-1} \partial_a u$ is an antihermitian matrix in the representation of the Lie algebra, it is seen that $S^{(0)}$ is positive (the Boltzmann weight is understood as $\exp -S$). The same formula shows that all what is really needed is an invariant bilinear form on the Lie algebra, identified with the unique one (hence the restriction to a simple Lie algebra) again up to a factor. Thus

$$\widetilde{\text{Tr}}(u^{-1} \partial_a u)(u^{-1} \partial_a u) \equiv (u^{-1} \partial_a u, u^{-1} \partial_a u) \quad (C.27)$$

The convenient normalization, justified below, is that this bilinear form is such that the length of long roots is equal to one. This convention should be kept in mind when comparing with expressions appearing in the works of other authors.

In arbitrary coordinates or in curved space, with a metric g_{ab} , the Lagrangian in $S^{(0)}$ should be replaced by the conformal invariant expression $\sqrt{g} g^{ab} \widetilde{\text{Tr}}(\partial_a u^{-1} \partial_b u)$. In any case, it is invariant under the global group $G \times G$, under $u \rightarrow u^{-1}$ as well as any reflection in the flat case.

To justify the structure of the added term $S^{(1)}$, one compactifies flat space to a sphere denoted here Σ_2 (due care has to be paid to a possible change of coordinate system). As a result, $x \rightarrow u(x)$ gives an image of Σ_2 into the group, which can be smoothly deformed to a point (the precise statement is that the second homotopy group $\pi_2(G)$ vanishes) so that we can extend $u(x)$ to a map $u(y)$ from the ball B (the interior of the sphere in three-space) to the group.

This enables one to write a second term

$$\begin{aligned} S^{(1)} &= -\frac{ik}{24\pi} \int d^3y \epsilon^{abc} \widetilde{\text{Tr}}(u^{-1} \partial_a u)(u^{-1} \partial_b u)(u^{-1} \partial_c u) \\ &\equiv -\frac{ik}{24\pi} \int_B (\omega_\wedge, \omega_\wedge \omega) \end{aligned} \quad (C.28)$$

where ω is the one-form $\omega = u^{-1} du$, and in the second expression, the products are understood as exterior products on forms and interior products on Lie algebra elements (clearly $\omega_\wedge \omega = \frac{1}{2} [\omega_\wedge, \omega]$ belongs to the Lie algebra).

The extension from Σ_2 to B is not unique. By glueing two such extensions on Σ_2 , one sees that the difference between two possible values of $S^{(1)}$ is given by a similar integral extending over a sphere Σ_3 , giving an element in a class of the third homotopy group $\pi_3(G)$. For simple Lie groups, the latter admits a unique generator and the two values of $S^{(1)}$ differ at most by a multiple of the integral over this generator. The normalization of $\widetilde{\text{Tr}}$ or $(,)$ is chosen in such a way that this arbitrariness is an entire multiple of $2\pi i$, leaving $\exp -S^{(1)}$ invariant, and requiring k to be an integer. Thus we must have

$$\frac{k}{48\pi^2} \int_{\Sigma_3} (\omega_\wedge, \omega_\wedge \omega) \in \mathbb{Z} \quad (C.29)$$

To see what is implied, consider an $SU(2)$ subgroup in G and represent faithfully u as $\exp i\psi \sigma \cdot \mathbf{n}$ where $\mathbf{n}$ is a unit vector on Σ_2 parametrized as $(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$, with $0 \leq \theta < \pi$, $0 \leq \varphi < 2\pi$, σ stands for the Pauli matrices and ψ ranges between 0 and π . This uses of course the fact that $SU(2)$ is topologically equivalent to Σ_3 . A simple calculation yields, with $\omega = u^{-1} du$,

$$(\omega_\wedge, \omega_\wedge \omega) = 6(\sigma_3, \sigma_3) \sin^2 \psi d\psi_\wedge \sin \theta d\theta_\wedge d\varphi$$

The homotopy generator is given by the identity map $\Sigma_3 \rightarrow SU(2)$. Since

$$\int_{\Sigma_3} \sin^2 \psi d\psi_\wedge \sin \theta d\theta_\wedge d\varphi = 2\pi^2$$

we have in this case

$$\frac{k}{48\pi^2} \int_{\Sigma_3} (\omega_\wedge, \omega_\wedge \omega) = \frac{1}{4} k(\sigma_3, \sigma_3) \quad (C.30)$$

The generator σ_3 is equivalent to k_α occurring in equations (C.5), (C.6). Therefore the quantization condition for the consistency of

a path integral is that for any root α , $\frac{1}{4}k(k_\alpha, k_\alpha)$ be an integer. But $(\alpha, \alpha) = (k_\alpha, k_\alpha) = 4/(k_\alpha, k_\alpha)$. Thus we have for any root α ,

$$k \in (\alpha, \alpha)\mathbb{Z} \quad (C.31)$$

The most stringent condition is for long roots. We have agreed that the quadratic form is normalized in such a way that $(\alpha_i, \alpha_i) = 1$, in which case k has to be an integer. The same computation shows that k might be further restricted if G is not simply connected. For instance if G is $SO(3)$ instead of $SU(2)$, then k must be even.

We also see that, even though the three-form $(\omega_\lambda, \omega_\lambda \omega)$ is closed but not exact, one can parametrize it in order that it appears locally as a differential of a (nonuniform) two-form. For $SU(2)$ for instance, we could write

$$S^{(1)} = \frac{ik}{8\pi} (\sigma_3, \sigma_3) \int_{\Sigma_2} \varphi \sin^2 \psi d\psi_\lambda \sin \theta d\theta$$

as a nonuniform integral over two-space with singularities at $\theta = 0$ and π .

We set the total action to be the sum

$$S = S^{(0)} + S^{(1)} \quad (C.32)$$

and compute its variation to obtain the classical equations of motion

$$\delta S = \int \frac{d^2x}{2\pi} \left(u^{-1} \delta y, \partial_a \left[\frac{1}{\lambda} u^{-1} \partial_a u + \frac{1}{4} k \epsilon_{ab} u^{-1} \partial_b u \right] \right) \quad (C.33)$$

which only involves a two-dimensional integral. Using complex coordinates $\{z, \bar{z}\}$, this yields

$$(\lambda^{-1} - \frac{1}{4}k) \bar{\partial} (u^{-1} \partial u) + (\lambda^{-1} + \frac{1}{4}k) \partial (u^{-1} \bar{\partial} u) = 0 \quad (C.34)$$

When

$$\lambda = \frac{4}{|k|} \quad (C.35)$$

we have either $\partial (u^{-1} \bar{\partial} u) = 0$ if k is positive, or $\bar{\partial} (u^{-1} \partial u) = 0$ if k is negative. Taking $k > 0$ for definiteness, a classical solution factorizes as

$$u_{\text{class}} = u_1(z) \bar{u}_2^{-1}(\bar{z}) \quad (C.36)$$

With this value for λ , the action takes the form

$$S(u) = \frac{k}{8\pi} \left[\int d^2x \frac{1}{2} \text{Tr}(\partial_a u^{-1} \partial_a u) - \frac{1}{3} i \int_B d^2y \epsilon^{abc} \text{Tr}(u^{-1} \partial_a u)(u^{-1} \partial_b u)(u^{-1} \partial_c u) \right] \quad (C.37a)$$

while

$$S(u + \delta u) = S(u) + k \int \frac{d^2x}{2\pi} (u^{-1} \delta u, \partial(u^{-1} \bar{\partial} u)) \quad (C.37b)$$

Set $u_t = u e^{-tH}$ with H an element in the Lie algebra. Using the above equation, one can integrate $\dot{S}(u_t)$ from 0 to 1 to get the identity

$$S(uv^{-1}) = S(u) + S(v^{-1}) + k \int \frac{d^2x}{2\pi} (v^{-1} \partial v, u^{-1} \bar{\partial} u) \quad (C.38)$$

Furthermore if we deform a solution $u_1(t, z) \bar{u}_2^{-1}(t, \bar{z})$ from $t = 0$ to 1 in such a way that at $t = 0$ it is the identity and at $t = 1$ it takes the value (C.36), it is readily seen from equation (C.37b) that the action remains constant and in fact equal to zero. Then equation (C.38) shows that, for the particular choice $\lambda = 4/|k|$, the action is invariant under the gauge group $G_z \otimes G_{\bar{z}}$

$$u(z, \bar{z}) \rightarrow u_1(z) u(z, \bar{z}) \bar{u}_2^{-1}(\bar{z}) \quad (C.39)$$

The question is now to find how this invariance reflects itself at the quantum level. Keeping λ and k as parameters, a one-loop computation shows that to leading order, the Callan-Symanzik β -function still vanishes at $\lambda = 4/|k|$, while it reveals a quantum anomaly (a Schwinger term) in the current commutators, the generators of the symmetry (C.39). Rather than trying to derive from the path integral the precise nonperturbative answers, one can attempt to find a consistent set of correlations using the Ward identities of conformal invariance together with those of gauge invariance.

The conserved currents

$$\begin{aligned} J(z) &= \frac{1}{2} k u \partial u^{-1} & \bar{\partial} J &= 0 \\ \bar{J}(\bar{z}) &= \frac{1}{2} k u^{-1} \bar{\partial} u & \partial \bar{J} &= 0 \end{aligned} \quad (C.40)$$

take their values in the Lie algebra and the conservation laws which express analyticity for J (antianalyticity for $\bar{J}$) are just a restatement of the classical equations of motion (if we note

that $\bar{\partial}J = -u(\partial\bar{J})u^{-1}$. As in the case of the energy-momentum tensor, these conservation laws are to be understood for insertions into correlation functions, up to contact terms.

Under an infinitesimal variation of the type (C.39), with $\Omega(z)$ and $\bar{\Omega}(\bar{z})$ infinitesimal elements in the Lie algebra

$$\begin{aligned} u(z, \bar{z}) &\rightarrow u(z, \bar{z}) - \Omega(z)u + u\bar{\Omega}(\bar{z}) \\ u^{-1}(z, \bar{z}) &\rightarrow u^{-1}(z, \bar{z}) - \bar{\Omega}(\bar{z})u^{-1} + u^{-1}\Omega(z) \end{aligned} \quad (C.41)$$

all this understood in a faithful matrix representation, the currents transform as

$$\begin{aligned} J &\rightarrow J + [J, \Omega] + \frac{1}{2}k\partial\Omega = J + \delta_\Omega J \\ \bar{J} &\rightarrow \bar{J} + [\bar{J}, \bar{\Omega}] + \frac{1}{2}k\bar{\partial}\bar{\Omega} = \bar{J} + \delta_{\bar{\Omega}} \bar{J} \end{aligned} \quad (C.42)$$

If we take into account the fact that the currents are themselves the generators for infinitesimal gauge transformations, these equations are a concise statement equivalent to the existence of two infinite Kac-Moody algebras, very much as equation (47) encapsulated the Virasoro algebras. It is understood up to here that J and $\bar{J}$ are classical fields appearing as arguments in correlation functions. In other words, their coefficients in an expansion over a Lie algebra basis do commute. If a field A has a local transformation law according to a representation of $G_z \otimes G_{\bar{z}}$, then equation (C.26) implies that

$$\delta_\Omega A(z', \bar{z}') = \oint \frac{dz}{2i\pi} (\Omega(z), J(z)) A(z', \bar{z}') \quad (C.43)$$

where the contour encircles the point z' , with a similar $\bar{z}$ -formula for $\delta_{\bar{\Omega}} A$. In deriving the normalization in (C.43), it is to be remembered that $d^2x = d\bar{z} \wedge dz/2i$. This is entirely parallel to equation (41) for a coordinate transformation. We apply this formula to $J(z)$ instead of A , and expand J on a Lie algebra basis t_a , with

$$\Omega = \sum \omega^a t_a, \quad J = \sum J^a t_a \quad \text{and} \quad (t_a, t_b) = \delta_{ab} \quad (C.44)$$

We define the totally antisymmetric structure constants through

$$[t_a, t_b] = if_{abc} t_c \quad (C.45)$$

Thus, in a unitary representation, the t 's would be Hermitian and coincide in the case of $SU(2)$ with $\frac{1}{2}\sigma_a$. Equation (C.43) translates

into the short-distance expansion

$$J^a(z)J^b(z') = \frac{1}{2}k \frac{\delta_{ab}}{(z-z')^2} + i \frac{f_{abc}}{(z-z')} J^c(z') + \dots \quad (C.46)$$

where the dots represent regular terms and use has been made of the total antisymmetry of the structure constants. The first term on the r.h.s. represents the quantum anomaly, analogous to the central charge for the Virasoro algebra. The expansion is in agreement with the commutative character of the coefficients $J^a(z)$ (i.e. it is invariant under a simultaneous exchange $(a, z) \leftrightarrow (b, z')$). We omit for short the $\bar{z}$ counterpart. Formula (C.46) is complemented by the standard ones giving the relation with the energy-momentum tensor and expressing that $J(z)$ has conformal weights (1,0)

$$\begin{aligned} T(z)T(z') &= \frac{c}{2(z-z')^4} + \frac{2}{(z-z')^2} T(z') \\ &\quad + \frac{1}{(z-z')} \partial T(z') + \dots \end{aligned} \quad (C.47)$$

$$T(z)J^a(z') = \frac{1}{(z-z')^2} J^a(z') + \frac{1}{(z-z')} \partial J^a(z') + \dots \quad (C.48)$$

Using radial quantization, one can define corresponding operators $\hat{T}(z)$, $\hat{J}(z)$, acting on conformal fields hence also on a linear space of states. We expand these operators as Laurent series in powers of z and omit the hats on the operator coefficients for short

$$\hat{T}(z) = \sum_{-\infty}^{+\infty} \frac{L_n}{z^{n+2}}, \quad \hat{J}^a(z) = \sum_{-\infty}^{+\infty} \frac{J_n^a}{z^{n+1}} \quad (C.49)$$

The short distance expansions are then equivalent to the set of commutation rules

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{1}{12}cn(n^2-1)\delta_{n+m,0} \quad (C.50a)$$

$$[L_n, J_m^a] = -mJ_{n+m}^a \quad (C.50b)$$

$$[J_n^a, J_m^b] = if_{abc}J_{n+m}^c + \frac{1}{2}kn\delta_{ab}\delta_{n+m,0} \quad (C.50c)$$

The last equation defines the Kac-Moody algebra, the only possible central extension of the loop group algebra

$$[j_n^a, j_m^b] = if_{abc}j_{n+m}^c$$

Again $\bar{z}$ counterparts are understood with the same values for k and c . Unitarity of a representation requires

$$(J_n^a)^\dagger = J_{-n}^a \quad (C.51)$$

A primary field, creating a highest weight state, will be characterized by a pair of conformal weights $(h, \bar{h})$ with respect to the Virasoro algebra, as well as a pair of finite-dimensional representations $(\mathcal{R}, \bar{\mathcal{R}})$ for the Lie algebras $\{J_0^a\}, \{\bar{J}_0^a\}$. Those can also be indexed by their dominant weight. Such a field evaluated at the origin (and with the barred counterparts omitted) when acting on the vacuum state generates a "ground state" $|h, \mathcal{R}\rangle$ annihilated by L_n and J_n^a for n positive.

From the specificity of the model, we expect that the energy-momentum tensor is expressible in terms of J and hence that the central charge is specified by the group G as well as the level k . Indeed at the classical level, the energy-momentum tensor is quadratic in the currents, corresponding to a structure suggested in the late sixties by Sugawara (in a fermionic context) and elaborated by many authors. Expressed in terms of operators this means that

$$\hat{T}(z) = \frac{1}{2x} \sum_a : \hat{J}^a(z) \hat{J}^a(z) : \quad (C.52)$$

for an appropriate value of the coefficient x . Normal ordering requires that operators with an index $n < 0$ be pushed to the left of operators with an index $n \geq 0$. Obviously from the diagonal form of the scalar product on the Lie algebra, the expression (C.52) is the only quadratic candidate for a group invariant (as $\hat{T}$ should be). We note the analogy with the scalar field case, where T was proportional to $(\partial\varphi)^2$ and is here proportional to its equivalent $\text{Tr} \partial u^{-1} \partial u$.

The constant x occurring in (C.52) is fixed by the consistency of the algebra. To see this, let us apply both sides of the equation to a highest weight state $|h, \mathcal{R}\rangle$ where h denotes the lowest value of L_0 , and retain only the term in z^{-1} . This reads

$$\left(x L_{-1} - \sum_b J_{-1}^b J_0^b \right) |h, \mathcal{R}\rangle = 0 \quad (C.53)$$

giving what might be termed a "null state" if we think of the module generated by $\{L_{-n}\} \{J_{-m}\} |h, \mathcal{R}\rangle$. Acting with L_1 and J_1^a

in turn, we find

$$\begin{aligned} \left(x [L_1, L_{-1}] - \sum_b [L_1, J_{-1}^b] J_0^b \right) |h, \mathcal{R}\rangle &= 0 \\ \left(x [J_1^a, L_{-1}] - \sum_b [J_1^a, J_{-1}^b] J_0^b \right) |h, \mathcal{R}\rangle &= 0 \end{aligned} \quad (C.54)$$

To use the commutation rules (C.50), we observe that acting in the irreducible representation $\mathcal{R}$ with dominant weight $q_{\mathcal{R}}$, the operator $\sum_b J_0^b J_0^b$ yields the corresponding value of the Casimir invariant

$$C_{\mathcal{R}} = (q_{\mathcal{R}}, q_{\mathcal{R}} + 2q_m) \quad (C.55)$$

while from (C.45), the similar quantity in the adjoint representation is

$$\begin{aligned} C_{\text{adj}} \delta_{c_1, c_2} &= \sum_{a, b} f_{abc_1} f_{abc_2} \\ &= (-\alpha_0, -\alpha_0 + 2q_m) \delta_{c_1, c_2} \end{aligned} \quad (C.56)$$

It can be checked that C_{adj} is nothing but the dual Coxeter number (recall that $(\alpha_i, \alpha_i) = 1$). Returning to equations (C.54), we find

$$\begin{aligned} 2hx - C_{\mathcal{R}} &= 0 \\ 2x - k - C_{\text{adj}} &\equiv 2x - k - g = 0 \end{aligned}$$

and therefore the normalization factor $2x$ is given by

$$2x = k + g \quad (C.57)$$

while the conformal weight is related to the representation $\mathcal{R}$ through

$$h = \frac{C_{\mathcal{R}}}{k + g} \quad (C.58)$$

For $SU(2)$, if we characterize the representation by the angular momentum j , knowing that g takes the value 2 (as follows from (C.55) for $f_{abc} \rightarrow \epsilon_{abc}$ or from $C_{\mathcal{R}} = j(j+1)$ for $j=1$), we have $2x = k+2$ and $h = j(j+1)/(k+2)$.

The value of the central charge can be obtained by requiring that the vacuum state be invariant under the group G and

therefore has weight $h = 0$. Thus computing

$$\begin{aligned}\langle 0 | L_2 L_{-2} | 0 \rangle &= \langle 0 | [L_2, L_{-2}] | 0 \rangle = \frac{1}{2}c \\ &= \frac{1}{4x^2} \left\langle 0 \left| \sum_{a,b} J_1^a J_1^a J_{-1}^b J_{-1}^b \right| 0 \right\rangle\end{aligned}$$

and using the Kac-Moody commutation rules, one finds upon using the above value x that

$$c = \dim(G) \frac{k}{k+g} \quad (C.59)$$

where $\dim(G)$ is the dimension of the Lie algebra of G . Thus for $SU(2)$, $c = 3k/(k+2)$. In general, for k integral, c ranges between the rank of the group (corresponding to $k = 1$ for simply laced groups) and its dimension as k goes to infinity.

By direct product, the construction extends to semisimple Lie algebras and can even include abelian factors (with corresponding value $g = 0$), where it coincides with the free field theory compactified on a circle.

There are various generalizations of the above construction. We shall limit ourselves to describing a remarkable observation made by Goddard, Kent and Olive. Suppose the simple Lie algebra of G contains a simple subalgebra corresponding to a subgroup H . If the level of G is k (the integer characterizing the anomaly in equation (C.51c)), when restricted to the subalgebra it might differ from k in the ratio of long roots, since, with $(\alpha_\ell, \alpha_\ell) = 1$ for G , the restriction of the quadratic form to the Lie algebra of H may induce a fractional value for the normalization of its long root. Assume that the first $\dim(H)$ indices refer to H , then

$$\begin{aligned}\hat{T}_G(z) &= \frac{1}{k+g_G} \sum_1^{\dim(G)} : \hat{J}^a(z) J^a(z) : \\ \hat{T}_H(z) &= \frac{1}{k + (\alpha_\ell, \alpha_\ell)_H g_H} \sum_1^{\dim(H)} : \hat{J}^a(z) \hat{J}^a(z) :\end{aligned} \quad (C.60)$$

Therefore the corresponding central charges are given by

$$c_G = \dim(G) \frac{k}{k+g_G} \quad c_H = \dim(H) \frac{k'}{k'+g_H} \quad (C.61)$$

where

$$k' = k \frac{(\alpha_\ell, \alpha_\ell)_G}{(\alpha_\ell, \alpha_\ell)_H} \geq k \quad (C.62)$$

is greater or equal to k since $(\alpha_\ell, \alpha_\ell)_G / (\alpha_\ell, \alpha_\ell)_H$ is an integer. From the commutation rules (C.50b), it follows that $L_n(G/H) \equiv L_n(G) - L_n(H)$ commutes with H currents and therefore also with $L_m(H)$. Thus the G -Virasoro algebra breaks into two commuting pieces

$$\hat{T}_G = \hat{T}_H + \hat{T}_{G/H} \quad (C.63)$$

Since central charges are additive, it follows that for the coset space G/H

$$c_{G/H} = c_G - c_H = \dim G \frac{k}{k+g_G} - \dim H \frac{k'}{k'+g_H} \quad (C.64)$$

Assume unitary representations. It follows that, if $c_{G/H} = 0$, the corresponding representation is trivial, and the representations of the combined Virasoro-Kac-Moody algebras for G and H are equivalent in spite of their very different appearance. This leads to nontrivial operator constructions. For instance if G is a simply laced Lie group, if $k = 1$, and if H is the maximal abelian torus (corresponding to the Cartan subalgebra), we have $c_G = c_H = \text{rank } G = \ell$. Thus $c_{G/H} = 0$, and this leads to a presentation of the nonabelian theory in terms of ℓ periodic scalar fields and an associated construction for the remaining ("off-diagonal") currents in terms of exponentials of these scalar fields.

Another application is a construction of the series of minimal unitary representations of the Virasoro algebra. One possibility is to take the sequence

$$G = Sp(2(m-1)) \supset Sp(2(m-2)) \otimes Sp(2) = H$$

with the original level being $k = 1$ and the subgroup inheriting in this case a level $k' = k = 1$. We have

$$\begin{aligned}c_G &= \frac{(m-1)(2m-1)}{m+1} = 2m - 5 + \frac{6}{m+1} \\ c_H &= \frac{(m-2)(2m-3)}{m} + 1 = 2m - 6 + \frac{6}{m}\end{aligned}$$

and therefore,

$$c_{Sp(2(m-1))_1 / (Sp(2(m-2)) \otimes Sp(2))_1} = 1 - \frac{6}{m(m+1)} \quad (C.65a)$$

Another possibility is to use for G the group $SU(2) \times SU(2)$ and for H the diagonal subgroup $SU(2)$. In G the levels of the two factors can be taken at will. The choice $(m-2, 1)$ yields for H the level $m-1$ and for the central charge

$$\begin{aligned} c_{(SU(2)_{m-2} \times SU(2)_1) / SU(2)_{m-1}} &= \frac{3(m-2)}{m} + 1 - \frac{3(m-1)}{m+1} \\ &= 1 - \frac{6}{m(m+1)} \end{aligned} \quad (C.65b)$$

i.e. once more the unitary-minimal series.

These observations can be used to complete the proof of the Friedan-Qiu-Shenker theorem that Virasoro representations with the above values for m integral larger or equal to 2 are the only unitary ones of central charge smaller than 1.

C.3. Representations and characters of Kac-Moody algebras

We can only sketch a few results from the representation theory of Kac-Moody algebras. We restrict ourselves to unitary representations and show first how one recovers in an algebraic setting the fact that k must be an integer (as opposed to the above topological derivation) and also derive the further restrictions on the representation $\mathcal{R}$ of G which lead to the so-called integrable representations of the Kac-Moody algebra.

Recall that for the Lie algebra of G in the adjoint representation the dominant weight is the positive (long) root

$$\psi = -\alpha_0 \quad (C.66)$$

With this notation, let us indeed show that k is a nonnegative integer satisfying

$$k \geq \frac{2(\psi, q_{\mathcal{R}})}{(\psi, \psi)} \geq 0 \quad (C.67)$$

Although our expression is written with the convention that $(\psi, \psi) = 1$, we have reinstated an arbitrary normalization to remind us that $2(\psi, q_{\mathcal{R}})/(\psi, \psi)$ is a nonnegative integer. To

prove this, consider the $SU(2)_{\alpha}$ subalgebra in the Kac-Moody algebra, in the obvious generalization of the Chevalley-Serre basis, resulting from the commutation rules

$$\begin{aligned} [e_1^{\alpha}, e_{-1}^{-\alpha}] &= k_0^{\alpha} - \frac{1}{2}k (e^{\alpha}, e^{-\alpha}) = k_0^{\alpha} - \frac{k}{(\alpha, \alpha)} \\ [k_0^{\alpha}, e_1^{\alpha}] &= 2e_1^{\alpha} \quad [k_0^{\alpha}, e_{-1}^{-\alpha}] = -2e_{-1}^{-\alpha} \end{aligned}$$

This shows that the generators

$$\frac{k}{(\alpha, \alpha)} - k_0^{\alpha}, e_{-1}^{-\alpha}, e_1^{\alpha}$$

can play the role of

$$k_{\alpha}, e_{\alpha}, e_{-\alpha}$$

respectively, occurring as $SU(2)$ generators in equations (C.5)–(C.6). The Kac-Moody representation can be decomposed into states with definite weights for the subalgebra $\{J_0^{\alpha}\}$ i.e. satisfying

$$k_0^{\alpha} |q\rangle = \frac{2(\alpha, q)}{(\alpha, \alpha)} |q\rangle$$

We see that for any q in this decomposition $k - 2(\alpha, q)$ must be an integer. Applying this to the dominant root $\alpha = -\alpha_0 = \psi$ taking into account $(\psi, \psi) = 1$, we find that $k - 2(\psi, q)/(\psi, \psi)$ has to be integral and thus k is also an integer. Furthermore if the highest weight state corresponds to a representation $\mathcal{R}$ with dominant weight $q_{\mathcal{R}}$, we must have

$$e_1^{-\alpha} |q_{\mathcal{R}}\rangle = 0$$

For a unitary representation $(e_1^{-\alpha})^{\dagger} = e_{-1}^{\alpha}$. As a consequence

$$\begin{aligned} 0 \leq \langle q_{\mathcal{R}} | e_1^{-\alpha} e_{-1}^{\alpha} | q_{\mathcal{R}} \rangle &= \langle q_{\mathcal{R}} | [e_1^{-\alpha}, e_{-1}^{\alpha}] | q_{\mathcal{R}} \rangle \\ &= \langle q_{\mathcal{R}} | \left(\frac{k}{(\alpha, \alpha)} - k_0^{\alpha} \right) | q_{\mathcal{R}} \rangle = \frac{k - 2(\alpha, q_{\mathcal{R}})}{(\alpha, \alpha)} \langle q_{\mathcal{R}}, q_{\mathcal{R}} \rangle \end{aligned}$$

Thus $k \geq 2(\alpha, q_{\mathcal{R}})$, and the most stringent condition is for the dominant root ψ normalized to unity, leading to the restriction (C.67) on the possible representations $\mathcal{R}$.

For instance for $SU(2)$, we have $(\psi, q_{\mathcal{R}})/(\psi, \psi) = j$ the integral or half integral angular momentum, and it follows for the corresponding Kac-Moody algebra, denoted $A_1^{(1)}$, that k must be a nonnegative integer such that $k \geq 2j \geq 0$.

To conclude this appendix, we describe the Kac-Moody characters of the corresponding unitary representations on the set of states obtained by acting with a string of operators J_{-n}^a ($n > 0$) on the highest weight state $|q_R\rangle$ and its partners giving a representation $\mathcal{R}$ of the underlying finite Lie group G .

Recall that we have an associated representation of the Virasoro algebra characterized by the values

$$c = \dim(G) \frac{k}{k+g} \quad h = \frac{C_R}{k+g} \quad (C.68)$$

for the central charge and the conformal weight h , where $C_R = (q_R, q_R + 2q_m)$. Of course the level k is a specified integer.

As in the finite-dimensional case, the character is a generating function for the multiplicities of the G weights, occurring at a degree n in a decomposition over states diagonalizing the Cartan subalgebra of $\{J_0^a\}$. Let these multiplicities be written $\text{mult}_n(q)$, and note that the corresponding eigenvalue of the grading operator L_0 is $h_n = h + n$.

With x an arbitrary ℓ -dimensional vector in root space and τ a complex number in the upper half plane, define the character as

$$\chi_{\mathcal{R},k}(x, \tau) = \sum_{n \geq 0} \sum_q \text{mult}_n(q) \exp 2i\pi\tau \left[(h+n) - \frac{1}{24}c \right] \exp 2i\pi(q, x) \quad (C.69)$$

where the sum over q is on the weight lattice of the finite Lie algebra of G , and we have included a factor $\exp(-\frac{1}{24}2i\pi\tau c)$, as in the text, in such a way that specialized to $x = 0$, the above formula reads

$$\chi_{\mathcal{R},k}(0, \tau) = \text{Tr}_{\mathcal{R},k} \exp 2i\pi\tau (L_0 - \frac{1}{24}c) \quad (C.70)$$

The Weyl-Kac expression for characters (generalizing equation (C.20)) yields χ as a ratio of two sums over the (finite) Weyl group of G , through

$$\chi_{\mathcal{R},k}(x, \tau) = \frac{\sum_{w \in W} \varepsilon(w) \theta_{w(q_R + q_m), k+g}(x, \tau)}{\sum_{w \in W} \varepsilon(w) \theta_{w(q_m), g}(x, \tau)} \quad (C.71)$$

where as above q_m is the sum of fundamental weights (equal to half the sum over positive roots) and the generalized θ -functions stand for

$$\theta_{q,k}(x, \tau) = \sum_{\beta \in q+k\mathcal{L}} \exp \left\{ \frac{2i\pi\tau}{k} \frac{(\beta, \beta)}{(\alpha_\ell, \alpha_\ell)} + 2i\pi(\beta, x) \right\} \quad (C.72)$$

Here $\mathcal{L}$ is the lattice generated on integers by the long roots and for simply laced groups corresponds to the full root lattice. We have again reinstated an arbitrary normalization for the invariant quadratic form. Of course when $k = 0$ and $\mathcal{R}$ is the identity, χ reduces to unity.

There exists a factorized form for the denominator in equation (C.71), analogous to equation (C.21), which leads to remarkable identities on θ -functions generalizing Jacobi's triple product identity (Appendix A). Inserting the definition (C.72) into (C.71) and dividing by a common factor precisely equal to the quantity $\psi_{q_m}(x)$ occurring in Weyl's formula, we obtain more explicitly

$$\chi_{\mathcal{R},k}(x, \tau) = \frac{\sum_{\gamma \in q_R + (k+g)\mathcal{L}} \exp \left[\frac{2i\pi\tau}{k+g} \cdot \frac{(q_m + \gamma)^2}{\alpha_\ell^2} \right] \chi_\gamma(x)}{\sum_{\gamma \in g\mathcal{L}} \exp \left[\frac{2i\pi\tau}{g} \cdot \frac{(q_m + \gamma)^2}{\alpha_\ell^2} \right] \chi_\gamma(x)} \quad (C.73)$$

Specializing to $x = 0$, with $\dim_q$ the dimension of the $\mathcal{R}$ -representation of dominant weight q , we find

$$\chi_{\mathcal{R},k}(0, \tau) = \frac{\text{Tr}_{\mathcal{R},k} \exp 2i\pi\tau (L_0 - \frac{1}{24}c)}{\sum_{\gamma \in g\mathcal{L}} \exp \left[\frac{2i\pi\tau}{g} \cdot \frac{(q_m + \gamma)^2}{\alpha_\ell^2} \right] \dim_\gamma} = \frac{\sum_{\gamma \in q_R + (k+g)\mathcal{L}} \exp \left[\frac{2i\pi\tau}{k+g} \cdot \frac{(q_m + \gamma)^2}{\alpha_\ell^2} \right] \dim_\gamma}{\sum_{\gamma \in g\mathcal{L}} \exp \left[\frac{2i\pi\tau}{g} \cdot \frac{(q_m + \gamma)^2}{\alpha_\ell^2} \right] \dim_\gamma} \quad (C.74)$$

As a particular example, let us take the case where $G \equiv SU(2)$, with a unique pair of roots $\pm\alpha$, $\mathcal{L}$ is the root lattice $\alpha\mathbb{Z}$, the dual Coxeter number is $g = 2$, and $q_m = \frac{1}{2}\alpha$. Then $q_R = 2j(\frac{1}{2}\alpha)$, $q_R + q_m = (2j+1)(\frac{1}{2}\alpha)$ and $\gamma = p\alpha$ with p ranging over integers. Finally $\dim q_R + (k+g)\gamma = 2j+1+2(k+2)p$. The character reads

$$\chi_{j,k}(0, \tau) = \frac{\sum_{p \in \mathbb{Z}} \exp \left\{ \frac{2i\pi\tau[2j+1+2(k+2)p]^2}{4(k+2)} \right\} [(2j+1) + 2(k+2)p]}{\sum_{p \in \mathbb{Z}} \exp \left\{ \frac{2i\pi\tau[1+4p]^2}{8} \right\} [1 + 4p]} \quad (C.75)$$

Setting as in section (3.4)

$$N = 2(k+2) \quad \lambda = 2j+1 \quad (C.76)$$

this is seen to coincide with the expression quoted in equation (268) as

$$\begin{aligned}\chi_{j,k}(0, \tau) &\equiv \chi_{\lambda, N}^{\text{aff}}(\tau) = \frac{\sum_{p \in \mathbb{Z}} \exp[2i\pi\tau(\lambda + pN)^2/2N] (Np + \lambda)}{\sum_{p \in \mathbb{Z}} \exp[2i\pi\tau(1 + 4p)^2/8] (4p + 1)} \\ &= \frac{\sum_{p \in \mathbb{Z}} \exp[2i\pi\tau(\lambda + pN)^2/2N] (Np + \lambda)}{\eta^3(\tau)}\end{aligned}\quad (C.77)$$

upon using Jacobi's identity.

The Goddard-Kent-Olive construction of unitary representations of the minimal unitary series, translates into an identity among characters pertaining both to the affine Lie algebra of $SU(2)$ and to the Virasoro algebra. With $\epsilon = 0$ or $\frac{1}{2}$, this reads

$$\begin{aligned}\chi_{j, k=m-2}^{\text{aff}}(\mathbf{x}, \tau) \chi_{\epsilon, k=1}^{\text{aff}}(\mathbf{x}, \tau) \\ = \sum'_{1 \leq 2j'+1 \leq m} \chi_{c=1-6/m(m+1), h_{2j+1, 2j'+1}}^{\text{Vir}}(\tau) \chi_{j', k=m-1}^{\text{aff}}(\mathbf{x}, \tau)\end{aligned}\quad (C.78)$$

where the prime means that the sum runs over those j' such that $j - j'$ is an integer if $\epsilon = 0$ and a half integer if $\epsilon = \frac{1}{2}$. Specializing to $\mathbf{x} = 0$, and using the (λ, N) notation as in (C.76), this reads for $1 \leq \lambda \leq m-1$ and $2\epsilon = 0$ or 1

$$\begin{aligned}\chi_{\lambda, 2m}^{\text{aff}}(\tau) \chi_{2\epsilon+1, 6}^{\text{aff}}(\tau) \\ = \sum_{\substack{1 \leq \lambda' \leq m \\ \lambda - \lambda' \equiv 2\epsilon \pmod{2}}} \chi_{\lambda(m+1) - \lambda', 2m(m+1)}^{\text{Vir}}(\tau) \chi_{\lambda', 2(m+1)}^{\text{aff}}(\tau)\end{aligned}\quad (C.79)$$

Notes

An original reference on conformal invariance is A.M. Polyakov, *JETP Lett.* **12**, 381 (1970). The associated infinite Lie algebra in the two-dimensional case was introduced in the framework of dual models by M.A. Virasoro *Phys. Rev. D* **1**, 2933 (1970). Additional material on previous attempts at consistent conformally invariant theories can be found in the book by I.T. Todorov, M.C. Mintchev and V.B. Petkova, *Conformal Invariance in Quantum Field Theory*, Pubblicazione della classe di scienze della scuola normale superiore, Pisa (1978). The fundamental papers, at the origin of present applications, are those of A.A. Belavin, A.M. Polyakov

and A.B. Zamolodchikov, *J. Stat. Phys.* **34**, 763 (1984) and *Nucl. Phys. B* **241**, 333 (1984).

The collective work *Vertex Operators in Mathematics and Physics* edited by J. Lepowsky, S. Mandelstam and I.M. Singer, Springer Verlag, New York, (1985) contains many contributions, among them by D. Friedan, Z. Qiu and S. Shenker on unitary representations of the Virasoro and superconformal algebras, and their relevance to physics, and by A. Rocha Carli on the corresponding characters.

The formula for the Kac determinant is announced in V.G. Kac, Springer Lecture Notes in Physics **94**, 441 (1979). We follow the derivation of B.L. Feigin and D.B. Fuchs in *Funct. Anal. and Appl.* **16**, 114 (1984), **17**, 241 (1983).

The articles by V.S. Dotsenko, *Nucl. Phys. B* **235** [FS 11] 54 (1984) and by the same author with V.A. Fateev, *Nucl. Phys. B* **240** [FS 12] 312 (1984) and **251** [FS 13] 691 (1985) discuss degenerate models and correlations.

The Ising partition function on a torus was derived in the work of Ferdinand and Fisher quoted in chapter 2. The relation between the central charge and the Casimir effect in a strip is discussed in H.W.J. Blöte, J.L. Cardy, M.P. Nightingale, *Phys. Rev. Lett.* **56**, 742 (1986) and I. Affleck, *Phys. Rev. Lett.* **56**, 746 (1986).

J.L. Cardy has written a review article in the Domb and Lebowitz series vol. 11 Academic Press, New York (1986). Some of his important contributions are *J. Phys. A* **17**, L385 (1984), *Nucl. Phys. B* **270** [FS 16] 186 (1986) on modular invariance, *Nucl. Phys. B* **240** [FS 12] 514 (1984) on the semi-infinite geometry, *Phys. Rev. Lett.* **54**, 1354 (1985) on the Lee-Yang edge singularity. The consequences of scaling in this case had previously been discussed by M.E. Fisher, *Phys. Rev. Lett.* **40**, 1610 (1978).

The interpretation of the principal series of minimal models in the unitary case by D.A. Huse, *Phys. Rev. B* **30**, 3908 (1984) is based on work by R.J. Baxter, *J. Phys. A* **13**, L61 (1980) and G.E. Andrews, R.J. Baxter and P.J. Forrester, *J. Stat. Phys.* **35**, 193 (1984). On this subject, see also A.B. Zamolodchikov, *Sov. J. Nucl. Phys.* **44**, 529 (1986).

Field theory on a torus is presented in a joint work with J.B. Zuber in *Nucl. Phys. B* **275** [FS 17] 580 (1986) while the A-D-E classification of minimal models stated as a conjecture in a work with A. Cappelli and J.B. Zuber, *Nucl. Phys. B* **280** [FS

18] 445 (1987) was proved by the same authors in *Comm. Math. Phys.* **113**, 1 (1987), and independently by A. Kato, *Mod. Phys. Lett. A2*, 585 (1987). An important relation with affine characters is discussed in D. Gepner and E. Witten, *Nucl. Phys. B278*, 493 (1986), D. Gepner, *Nucl. Phys. B287*, 111 (1987), and a piece of the puzzle proved by D. Gepner and Z. Qiu, *Nucl. Phys. B285* [FS 19] 423 (1987) as part of a larger discussion. Further work along the lines of integrable models is due to V. Pasquier, *Nucl. Phys. B285* [FS19], 162 (1987).

T. Eguchi and H. Ooguri have generalized the Belavin-Polyakov-Zamolodchikov equations to Riemann surfaces of arbitrary genus in *Nucl. Phys. B282*, 308 (1987). Explicit formulae for Ising correlation functions are found in P. Di Francesco, H. Saleur and J.-B. Zuber, *Nucl. Phys. B290* [FS20], 527 (1987).

Frustrations and symmetries are studied in G. von Gehlen, V. Rittenberg, *J. Phys. A19* L625 (1986), J.-B. Zuber, *Phys. Lett. E176*, 127 (1986) and J.L. Cardy, *Nucl. Phys. B275* [FS 17] 200 (1986).

The Ashkin-Teller (J. Ashkin and E. Teller, *Phys. Rev.* **64**, 178 (1943)), Potts, six and eight vertex models are presented in R.J. Baxter's book, *Exactly Solved Models in Statistical Mechanics* Academic Press, New York (1982). The interpretation of statistical models in terms of Coulomb gases is reviewed in B. Nienhuis, *J. Stat. Phys.* **34**, 731 (1984), with credit given to numerous contributors. In particular the combined charge-monopole operators and their correlations appear in the work of L.P. Kadanoff and A.C. Brown, *Ann. Phys. (New York)* **121**, 318 (1979). Their study in the conformal context is found in P. Di Francesco, H. Saleur and J.B. Zuber, *J. Stat. Phys.* **49**, 57 (1987), S.K. Yang, *Nucl. Phys. B285* [FS19] 183 (1987), S.K. Yang and H.B. Zheng, *Nucl. Phys. B285* [FS19] 410 (1987).

Deviations from criticality are discussed by A.B. Zamolodchikov *JETP Letters* **43**, 730 (1986) and in the work of one of the authors with H. Saleur, *J. Stat. Phys.* **48**, 449 (1987).

A reference to elliptic and modular functions, including Kronecker's limit formulas, is A. Weil *Elliptic Functions according to Eisenstein and Kronecker*, Springer Verlag, Berlin (1976).

The treatment of fermions in the context of conformal invariance originates in the work of P. Ramond, *Phys. Rev. D3*, 2415 (1971) and A. Neveu and J.H. Schwarz, *Nucl. Phys. B31*, 86 (1971).

For superconformal field theory and its applications to statistical physics, see the contribution of D. Friedan, Z. Qiu and S. Shenker in the book on *Vertex Operators in Mathematics and Physics* quoted above and their article in *Phys. Lett.* **151B**, 37 (1985) as well as M.A. Bershadsky, V.G. Knizhnik, M.G. Teitelman, *ibid.* **31** (1985), and H. Eichenherr, *ibid.* **26** (1985). P. Goddard, A. Kent and D. Olive, *Comm. Math. Phys.* **103**, 105 (1986) and A. Meurman and A. Rocha-Caridi, *Comm. Math. Phys.* **107**, 263 (1986) discuss the construction of characters. The classification of minimal superconformal models is investigated in D. Kastor, *Nucl. Phys. B280* [FS 18] 304 (1987), Y. Matsuo and S. Yahikozawa, *Phys. Lett.* **178B**, 211 (1986), and A. Cappelli, *Phys. Lett.* **185B**, 82 (1987).

The model of appendix C is described in A. Polyakov and P.B. Wiegmann, *Phys. Lett. B131*, 121 (1983), and E. Witten, *Comm. Math. Phys.* **92**, 455 (1984). The conformal structure is elaborated by V. Knizhnik and A.B. Zamolodchikov, *Nucl. Phys. B247*, 83 (1984), D. Gepner and E. Witten, *Nucl. Phys. B278*, 493 (1986), P. Goddard and D. Olive, *Int. J. of Mod. Phys. A1*, 303 (1983), where the last paper includes a review on Lie and Kac-Moody algebras. This subject is covered in many books among which J.P. Serre *Algèbres de Lie semi-simples complexes*, Benjamin, New York (1966) and J.E. Humphreys *Introduction to Lie Algebras and Representation Theory*, Springer Verlag, Heidelberg (1972). Kac-Moody algebras and their representation theory are described by V.G. Kac *Infinite Dimensional Lie Algebras*, 2nd edition, Cambridge University Press (1985), which contains numerous references to the mathematical literature. Finally it is worth recalling that a field theory in terms of currents was proposed by H. Sugawara, *Phys. Rev.* **170**, 1659 (1968) and developed by many authors, as part of current algebra, with results which anticipated many of the present ones, as testified in R. Dashen and Y. Frishman, *Phys. Rev. D11*, 2781 (1975).

Among important developments not covered here is the study of current correlations and their remarkable monodromy properties, as well as the relation between fermionic and bosonic constructions. The whole subject of two-dimensional conformal field theory has important applications to string field theory, an aspect briefly described in chapter 11.

DISORDERED SYSTEMS AND FERMIONIC METHODS

Real systems may involve various types of defects. It is therefore important to assess their relevance and to estimate their effect on the results established in pure cases. The emphasis here is on stability and robustness versus random perturbations. We present in the last section of this chapter the Harris criterion for estimating the effect of weak disorder on a critical system. At a more fundamental level, one can also look for new phenomena originating in the very existence of defects. For instance, the dynamics of dislocations in a crystal can explain the solid-liquid transition. Similarly, we were led to analyze the role of vortices when studying the XY -model. Random impurity potentials produce the localization of quantum wavefunctions (Anderson, 1958), which enables one to understand the transition between insulators and conductors. This same localization phenomenon appears in other contexts involving classical waves, such as optics or sound propagation. This subject has given rise to an intense activity. One cannot claim that it is fully understood at present, even though the weak disordered case seems under control, thanks to renormalization group arguments. The role of magnetic fields has opened a new area of research centered on the quantum Hall effect. Another domain which has stimulated vigorous and original developments is the one of magnetic systems with random and/or frustrated interactions, leading to the search for a spin glass phase, where magnetic moments can become frozen in random directions, with a plethora of metastable states, or valleys, in free energy. Applications of spin glass models range from neural systems, and optimization problems to the newly discovered high T_c superconductors. This list illustrates the diversity of the subject.

The archetype of systems with a Hamiltonian involving random parameters is the theory of random matrices elaborated by Wigner and Dyson in the context of nuclear spectra.

In this chapter, we offer simple illustrations of field theoretic methods, particularly the use of anticommuting variables. Our subjective choice of topics is mostly meant to invite the interested reader to pursue the study of this important branch of statistical physics.

10.1 One-dimensional models

10.1.1 Gaussian random potential

Consider a one-dimensional quantum Hamiltonian for a particle moving in a random potential with a Gaussian distribution. An ingenious method is known to obtain the spectrum (Halperin, 1965). The availability of an exact result enables one to test the heuristic replica trick, a method generally used in this context. We will also encounter in this formulation an elementary example of supersymmetry appearing when dealing with constrained systems, the Becchi-Rouet-Stora (BRS) supersymmetry, originally formulated in the quantization of gauge fields.

One is of course interested in various other properties, such as the structure of wavefunctions or the average of products of Green functions (to obtain transport coefficients). Unfortunately these are much more delicate questions, for which analytic answers are not known. We shall however be able to show that, quite generally, wavefunctions are localized in one dimension, when studying a discrete disordered system. This property holds true provided the random part of the Hamiltonian does not involve long-range correlations. This seems at first to confront us with a paradox. On the one hand, we shall obtain a continuous average spectrum, but localized wavefunctions indicate that they are normalizable, a property which under ordinary circumstances corresponds to a discrete spectrum. Here we have to face the unhappy fact that the classification of spectra involves not only a continuous and a discrete part, but also a singular one, the latter being infrequently encountered in unsophisticated applications. It is nonetheless most likely true that we encounter in disordered systems a typical situation where a singular part appears generically in the

spectrum, and that its character is not accessible while computing an average value for the latter. It is also worthwhile to mention that, while we concentrate here on random Hamiltonians, similar properties would show up in the case where part of the operator, the potential say, would imply several incommensurate periods. This should be the case for instance for spectra pertaining to quasicrystals.

Let the Hamiltonian be

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \quad (1)$$

where the potential V is random, with zero mean, and no correlation between distinct points. This is an extreme case, but, as was said above, amenable to an exact treatment. We have to distinguish between quantum averages, which we denote here by round brackets, and averages over disorder, which we denote with usual brackets. The potential is described by a Gaussian process with a correlation written as

$$\langle V(x)V(y) \rangle = D\delta(x-y) \quad (2)$$

More precisely, all averages can be derived from Wick's theorem in the form

$$\left\langle \exp i \int dx g(x)V(x) \right\rangle = \exp -\frac{1}{2}D \int dx g^2(x) \quad (3)$$

assuming $g(x)$ square integrable.

We want to compute the average of the Green function at coinciding points. For complex z , introduce the quantity

$$G(z) = \left\langle \left(x \left| \frac{1}{z-H} \right| x \right) \right\rangle \quad (4)$$

It is x -independent due to invariance under translations (and reflection) after averaging over V . The imaginary part of $G(z)$, as z tends to the real axis, yields the average density of energy levels per unit length, as the total number of levels is proportional to the total length. To be rigorous, the system ought to be enclosed in a finite box of size L , and boundary conditions imposed on wavefunctions, for instance periodicity. Then $\rho(E)$, the average spectral density (per unit length), in the limit of an infinite box, is given by

$$\rho(E) = \text{Im} \frac{1}{\pi} G(E - i0) = \langle (x | \delta(E-H) | x) \rangle \quad (5)$$

while the number of levels per unit length, with energies smaller than E , is

$$N(E) = \int_{-\infty}^E dE' \rho(E') \quad (6)$$

Since the potential is unbounded, we expect a spectrum extending from minus to plus infinity. It is convenient to introduce dimensionless units for lengths and energies, according to

$$\begin{aligned} \lambda &= \left(\frac{\hbar}{Dm^{\frac{1}{2}}} \right)^{\frac{1}{2}} & \rho(E) &= \left(\frac{m}{\hbar^2 D} \right)^{\frac{1}{2}} \bar{\rho}(\bar{E}) \\ V &= \lambda^{-2} \bar{V} & H &= \lambda^{-2} \bar{H} = \lambda^{-2} \left(-\frac{1}{2} \bar{\Delta} + \bar{V} \right) \\ x &= \frac{\hbar}{m^{\frac{1}{2}}} \lambda \bar{x} & \langle \bar{V}(\bar{x}) \bar{V}(\bar{y}) \rangle &= \delta(\bar{x} - \bar{y}) \\ E &= \lambda^{-2} \bar{E} \end{aligned} \quad (7)$$

Henceforth we omit the bars, and note that in reduced units, D and $m/\hbar^2$ are effectively equal to unity.

At very high energy, the potential becomes negligible, so that the total number of states tends asymptotically to its classical limit (one state per cell $\Delta x \Delta p / 2\pi$)

$$\begin{aligned} N(E) &\underset{E \rightarrow \infty}{\sim} \int_{p^2 < 2E} \frac{dp}{2\pi} = \frac{(2E)^{\frac{1}{2}}}{\pi} \\ \rho(E) &\underset{E \rightarrow \infty}{\sim} \frac{1}{\pi(2E)^{\frac{1}{2}}} \end{aligned} \quad (8)$$

The behaviour for $E \rightarrow -\infty$ is less intuitive, and will result from the exact expression.

10.1.2 Fokker-Planck equation

With $V(x)$ chosen according to the Gaussian probability law, let $\psi(x)$ be the wavefunction corresponding to energy E , a solution of Schrödinger's equation. Of course $\psi(x)$ becomes a random functional of V . The logarithmic derivative f of ψ satisfies a first order Ricatti equation

$$\begin{aligned} f(x) &= \psi'(x)/\psi(x) \\ f' + f^2 &= 2(V - E) \end{aligned} \quad (9)$$

An analogy with Brownian motion arises if we interpret the spatial coordinate x as playing the role of time.

We want to exploit the fact that the values of the potential at distinct points x are independent variables. Let $P(\xi, x)d\xi$ be the probability that $f(x)$ lies in the interval $\xi, \xi + d\xi$. Thus

$$P(\xi, x) = \langle \delta(f(x) - \xi) \rangle = \frac{1}{2\pi} \int d\alpha e^{-i\alpha\xi} \langle e^{i\alpha f(x)} \rangle \quad (10)$$

Taking the x derivative of the last expectation value and using equation (9), we find

$$\begin{aligned} \frac{\partial}{\partial x} \langle e^{i\alpha f(x)} \rangle &= i\alpha \langle f'(x) e^{i\alpha f(x)} \rangle \\ &= -i\alpha \langle (f^2(x) + 2E - 2V(x)) e^{i\alpha f(x)} \rangle \end{aligned}$$

The characteristic property of the Gaussian law (3), with D now set equal to one, is

$$\langle V(x)\phi \rangle = \int dy \delta(x-y) \left\langle \frac{\delta\phi}{\delta V(y)} \right\rangle = \left\langle \frac{\delta\phi}{\delta V(x)} \right\rangle$$

Applying this in the above formula yields

$$\frac{\partial}{\partial x} \langle e^{i\alpha f(x)} \rangle = -i\alpha \left\langle \left(f^2(x) + 2E - 2i\alpha \frac{\delta f(x)}{\delta V(x)} \right) e^{i\alpha f(x)} \right\rangle \quad (11)$$

Let us now pretend that we integrate the Ricatti equation forward in x , starting from the value of f at the point x_0 , which we let recede to $-\infty$. Under such circumstances, $f(x)$ only depends on the values of V at points $y < x$. According to equation (9), the retarded function

$$\Gamma(x, y) = \frac{\delta f(x)}{\delta V(y)} \quad y \leq x \quad (12)$$

vanishing for $y > x$, satisfies the equation

$$\left\{ \frac{\partial}{\partial x} + 2f(x) \right\} \Gamma(x, y) = 2\delta(x-y) \quad (13)$$

with a solution given by

$$\Gamma(x, y) = 2\theta(x-y) \exp - \int_y^x dx' 2f(x') = 2\theta(x-y) \frac{\psi^2(y)}{\psi^2(x)} \quad (14)$$

When $y \rightarrow x$, $\Gamma(x, y)$ seems to have the undefined limit

$$\Gamma(x, x) = 2\theta(0) \quad (15)$$

which occurs frequently when dealing not too carefully with such problems. To handle this apparent difficulty, while maintaining

the symmetry $x \rightarrow -x$, one can either discretize properly the model, or else replace in the $V-V$ correlation the singular δ -function by an approximation $u_\epsilon(x-y)$ with a function u_ϵ symmetric, positive, strongly peaked at the origin, of unit integral and tending towards δ as ϵ tends to zero. This amounts to replacing the factor $\theta(x-y)$ in equation (14) by the integral of $u_\epsilon(x')$ from $-\infty$ to $x-y$, and $\theta(0)$ by the same integral up to zero. Consequently, $\theta(0)$ is replaced by $\frac{1}{2}$, the average between $\theta(-0)$ and $\theta(+0)$, and therefore $\delta f(x)/\delta V(x) = \Gamma(x, x) = 1$. With this understood, we return to $P(\xi, x)$ and obtain, by integrating (11) with respect to α , the Fokker-Planck equation

$$\frac{\partial}{\partial x} P(\xi, x) = \frac{\partial}{\partial \xi} \left(\xi^2 + 2E + 2 \frac{\partial}{\partial \xi} \right) P(\xi, x) \quad (16)$$

with ξ playing the role of configuration variable, and x the role of time. The structure of this equation entails the conservation of the total probability, $\int d\xi P(\xi, x)$, under mild assumptions on the behaviour of P for large ξ .

For large x , P tends to a stationary process, with

$$\lim_{x \rightarrow \infty} P(\xi, x) = p(\xi)$$

satisfying

$$\frac{\partial}{\partial \xi} \left(\xi^2 + 2E + 2 \frac{\partial}{\partial \xi} \right) p(\xi) = 0 \quad (17)$$

Thus $p(\xi)$ describes an x -independent distribution of $\xi = \psi'/\psi$. Integrating equation (17) with respect to ξ , we have

$$\left(\xi^2 + 2E + 2 \frac{\partial}{\partial \xi} \right) p(\xi) = p_0 \quad (18)$$

The constant p_0 is determined by the normalization condition

$$\int d\xi p(\xi) = 1 \quad (19)$$

which requires

$$\lim_{\xi \rightarrow \infty} \xi^2 p(\xi) = p_0 \quad (20)$$

We wish to obtain the relation between $p(\xi)$ and the spectral density. To do so, reinstate the box with x ranging from 0 to L , and, given a value ξ_0 for $f(0)$, require that $f(L) = \xi_L$. This

leads to a discrete spectrum $E_1, E_2, \dots$. We let L tend to infinity and find

$$\rho(E) = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_n \delta(E - E_n) = \lim_{L \rightarrow \infty} \frac{1}{L} \delta(f(L) - \xi_L) \left| \frac{\partial f}{\partial E}(L) \right| \quad (21)$$

We expect the limit on the r.h.s. to be finite. From equation (9)

$$\frac{\partial}{\partial x} \frac{\partial f}{\partial E} + 2f \frac{\partial f}{\partial E} = -2 \quad \frac{\partial f}{\partial E}(x=0) = 0$$

If we substitute for $\partial f / \partial E$ the function $g(x)$ defined through

$$\frac{\partial f}{\partial E} = g(x) \exp \left\{ -2 \int_0^x dy f(y) \right\} = g(x) \frac{\psi^2(0)}{\psi^2(x)}$$

the equation becomes

$$\frac{\psi^2(0)}{\psi^2(x)} g'(x) = -2$$

Thus

$$g(x) = -2 \int_0^x dy \frac{\psi^2(y)}{\psi^2(0)}$$

and

$$\frac{\partial f}{\partial E} = -2 \int_0^x dy \frac{\psi^2(y)}{\psi^2(x)} \quad (22)$$

We note that $\partial f / \partial E$ is of constant negative sign. Consequently, the quantity

$$\mathcal{P}(\xi, x) = \left\langle \left| \frac{\partial f}{\partial E} \right| \delta(f(x) - \xi) \right\rangle = - \left\langle \frac{\partial f}{\partial E} \delta(f(x) - \xi) \right\rangle \quad (23)$$

is related to the spectral density through

$$\rho(E) = \lim_{L \rightarrow \infty} \frac{1}{L} \mathcal{P}(\xi_L, L) \quad (24)$$

Proceeding as we did before for $P(\xi, x)$, one establishes that

$$\begin{aligned} \frac{\partial}{\partial x} \mathcal{P}(\xi, x) &= \left\langle \left(2 + 2\xi \frac{\partial f}{\partial E} - \frac{\partial f}{\partial E} \frac{\partial}{\partial x} \right) \delta(f(x) - \xi) \right\rangle \\ &= 2\mathcal{P}(\xi, x) + \left[(\xi^2 + 2E) \frac{\partial}{\partial \xi} + 2 \frac{\partial^2}{\partial \xi^2} \right] \mathcal{P}(\xi, x) \end{aligned}$$

According to equation (24), $\mathcal{P}(\xi, x)$ behaves as $x\rho(E)$ for large positive x , hence the left hand side of this equation tends to $\rho(E)$

as $x \rightarrow +\infty$, independently of ξ . Let us multiply both sides by $p(-\xi)$ and integrate over ξ . Taking into account the normalization of $p(\xi)$, we get

$$\begin{aligned} \rho(E) &= 2 \int d\xi p(\xi) p(-\xi) \\ &\quad + \lim_{x \rightarrow +\infty} \int d\xi p(-\xi) \left[(\xi^2 + 2E) \frac{\partial}{\partial \xi} + 2 \frac{\partial^2}{\partial \xi^2} \right] \mathcal{P}(\xi, x) \end{aligned}$$

Integrating the second term by parts, we find from (17) that it vanishes, and we obtain the required relation between the spectral density and the probability $p(\xi)$ in the form

$$\rho(E) = 2 \int d\xi p(\xi) p(-\xi) \quad (25)$$

A more explicit formula follows from equation (18). Taking in the latter a derivative with respect to E , one has

$$\frac{\partial p_0}{\partial E} = 2 \frac{\partial}{\partial \xi} \frac{\partial p}{\partial E} + (\xi^2 + 2E) \frac{\partial p}{\partial E} + 2p$$

Premultiply by $p(\xi)$, integrate with respect to ξ , use equation (18) again and the normalization condition to find

$$\begin{aligned} \frac{\partial p_0}{\partial E} &= \rho(E) + \int d\xi p(-\xi) \left\{ 2 \frac{\partial}{\partial \xi} + (\xi^2 + 2E) \right\} \frac{\partial p}{\partial E} \\ &= \rho(E) + p_0 \int d\xi \frac{\partial p}{\partial E} = \rho(E) \end{aligned}$$

Hence

$$p_0 = \mathcal{N}(E) \quad (26)$$

Inserting this relation in equation (18), we see finally that $\mathcal{N}(E)$ is determined by the conditions

$$\begin{aligned} \mathcal{N}(E) &= \left[2 \frac{\partial}{\partial \xi} + (\xi^2 + 2E) \right] p(\xi) \\ \int d\xi p(\xi) &= 1 \quad \lim_{\xi \rightarrow \infty} \xi^2 p(\xi) = \mathcal{N}(E) \end{aligned} \quad (27)$$

Equation (27) is nothing but the Fourier transform of Airy's equation. Taking into account the asymptotic condition, its solution reads

$$p(\xi) = \frac{\mathcal{N}(E)}{2} e^{-\frac{1}{3}\xi^3 - E\xi} \int_{-\infty}^{\xi} d\eta e^{\frac{1}{3}\eta^3 + E\eta} \quad (28)$$

Normalization requires

$$\frac{1}{N(E)} = \frac{1}{2} \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{\xi} d\eta e^{-\frac{1}{2}(\xi^3 - \eta^3) - E(\xi - \eta)}$$

If we choose $\xi - \eta$ as an integration variable, we obtain the final expression as a single integral

$$\frac{1}{N(E)} = \sqrt{\frac{1}{2}\pi} \int_0^{\infty} \frac{du}{\sqrt{u}} \exp\left(-\left(\frac{1}{24}u^3 + Eu\right)\right) \quad (29)$$

When $E \rightarrow \infty$, u^3 can be neglected as compared to Eu , and we recover the semiclassical asymptotic form for $N(E)$, as expected in (8). For $E \rightarrow -\infty$, the integral can be estimated using the saddle point method, with the result

$$N(E) \underset{E \rightarrow -\infty}{\sim} \frac{\sqrt{-8E}}{2\pi} \exp\left(-\frac{1}{12}(-8E)^{\frac{3}{2}}\right) \quad (30)$$

It is not too surprising to find that $N(E)$ vanishes rapidly for $E \rightarrow -\infty$. The precise behaviour is somehow less expected. It is sometimes called a Lifshitz tail. The shape of $N(E)$ is depicted on figure 1. Observe that $N(E)^{-1}$ is an entire function with a series expansion

$$N(E)^{-1} = \frac{\pi^{\frac{1}{2}}}{3^{\frac{1}{2}}} \sum_0^{\infty} \frac{(-2 \cdot 3^{\frac{1}{2}} E)^n}{n!} \Gamma\left(\frac{2n+1}{6}\right) \quad (31)$$

It is instructive to obtain the same Fokker-Planck equation (16), using a slightly different reasoning, valid for more general dynamical systems. We pursue the analogy between the coordinate x and a temporal variable. Let us start from equation (9), which we use in the form

$$V(x) = E + \frac{1}{2} \left(\frac{\partial f}{\partial x} + f^2 \right)$$

To obtain the probability $P(\xi, x; \xi_0, x_0)$, we insert the above constraint in the distribution law for V , hence deriving the functional representation

$$P(\xi, x; \xi_0, x_0) = \int DV \int D(\xi) \prod_{x'} \delta\left(V - E - \frac{1}{2} \left[\frac{\partial \xi}{\partial x'} + \xi^2 \right]\right) \left| \frac{DV}{D\xi} \right| \exp\left(-\frac{1}{2} \int_{x_0}^x dy V^2(y)\right) \quad (32)$$

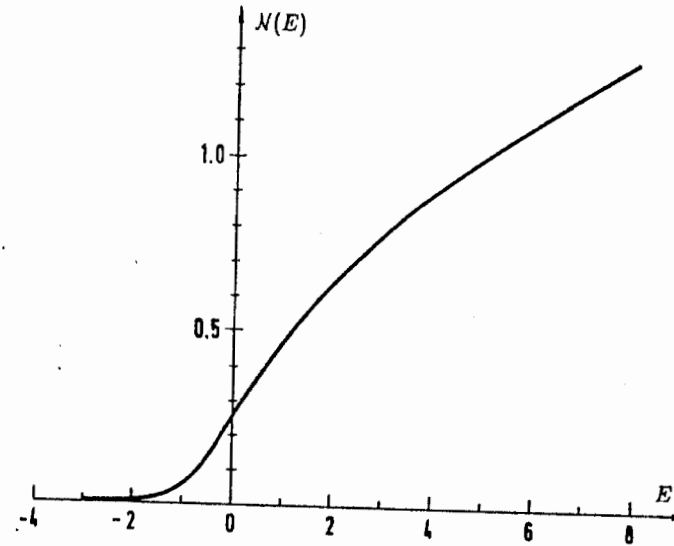


Fig. 1 The integrated density of states per unit length for a Gaussian random potential.

The integration over ξ satisfies the boundary conditions $\xi(x_0) = \xi_0$, $\xi(x) = \xi$. The Jacobian $J = |DV/D\xi|$ can be evaluated up to a multiplicative constant, assuming the mapping $\xi \leftrightarrow V$ to be one-to-one

$$J = J_0 \det \left(\frac{d/dx + 2\xi}{d/dx} \right) = J_0 \det \left(I + 2 \frac{1}{d/dx} \xi \right)$$

where $(d/dx)^{-1}$ is understood as the free retarded Green function $\Gamma_0(x, y)$ satisfying $\partial/\partial x \Gamma_0(x, y) = \delta(x - y)$, i.e. $\Gamma_0(x, y) = \theta(x - y)$. Thus

$$J = J_0 \exp \text{Tr} \ln \left(1 + \frac{1}{d/dx} 2\xi \right) = J_0 \exp 2 \int_{x_0}^x dy \Gamma_0(y, y) \xi(y)$$

Only the first term in the series expansion of the logarithm contributes in virtue of the retardation property of Γ_0 . We have already settled the question of the ill-defined $\Gamma_0(y, y) = \theta(0)$ by

arguing that it ought to be understood as $\frac{1}{2}$. Consequently

$$J = J_0 \exp \int_{x_0}^x dy \xi \quad (33)$$

which we now insert in equation (32). Let us absorb J_0 into the normalization of the functional integral, and integrate over V using the δ -functions. If $\dot{\xi}(y)$ stands for $d\xi/dy$, we obtain

$$\begin{aligned} P(\xi, x; \xi_0, x_0) &= \int \mathcal{D}\xi \exp - \int_{x_0}^x dy \left\{ \frac{1}{8} \dot{\xi}^2 + \frac{1}{2} \xi \left(\frac{1}{2} \xi^2 + E \right) + \frac{1}{2} \left(\frac{1}{2} \xi^2 + E \right)^2 - \xi \right\} \\ &= \exp - \left\{ \frac{1}{12} \xi^3 + \frac{1}{2} E \xi \right\} + \left\{ \frac{1}{12} \xi_0^3 + \frac{1}{2} E \xi_0 \right\} \\ &\quad \int \mathcal{D}\xi \exp - \int_{x_0}^x dy \left\{ \frac{1}{8} \dot{\xi}^2 + \frac{1}{2} \left(\frac{1}{2} \xi^2 + E \right)^2 - \xi \right\} \end{aligned} \quad (34)$$

The remaining path integral is a familiar expression for a quantum transition amplitude corresponding to the Euclidean dynamics of a fictitious particle of coordinate ξ with mass $\frac{1}{4}$ evolving in a potential $\frac{1}{2} \left(\frac{1}{2} \xi^2 + E \right)^2 - \xi$. As a result of this interpretation, it follows that P satisfies an analog of Schrödinger's equation

$$\begin{aligned} -\frac{\partial}{\partial x} P(\xi, x; \xi_0, x_0) &= e^{-\left\{ \frac{1}{12} \xi^3 + \frac{1}{2} E \xi \right\}} \\ &\quad \left[-2 \frac{\partial^2}{\partial \xi^2} + \frac{1}{2} \left(\frac{1}{2} \xi^2 + E \right)^2 - \xi \right] e^{\left\{ \frac{1}{12} \xi^3 + \frac{1}{2} E \xi \right\}} P(\xi, x; \xi_0, x_0) \end{aligned} \quad (35)$$

Since

$$e^{-\left\{ \frac{1}{12} \xi^3 + \frac{1}{2} E \xi \right\}} \frac{\partial}{\partial \xi} e^{\left\{ \frac{1}{12} \xi^3 + \frac{1}{2} E \xi \right\}} = \frac{\partial}{\partial \xi} + \frac{1}{4} \xi^2 + \frac{1}{2} E$$

we can rewrite equation (35) as

$$\frac{\partial}{\partial \xi} \left[\xi^2 + 2E + 2 \frac{\partial}{\partial \xi} \right] P(\xi, x; \xi_0, x_0) = \frac{\partial}{\partial x} P(\xi, x; \xi_0, x_0) \quad (36)$$

which is identical to the Fokker-Planck equation (16). We have also obtained in this way a path integral representation for the solution, as well as a natural connection between Euclidean quantum dynamics and the probabilistic theory of Brownian motion.

We can take advantage of this example to make a point concerning the treatment of the Jacobian. In the present case, we were fortunate in being able to compute it in closed form, due to its simple expression. However, this is the exception rather than the rule, and the variant treated below indicates a possibility of taking the calculations further, even without an analytic form for J . We

want to replace J by a path integral over the exponential of an action. Let us assume again that the change of variable $V \leftrightarrow \xi$ is well-defined. This is the crucial assumption and (unfortunately) is not always true in more complex situations. If the hypothesis is fulfilled, the Jacobian J never vanishes, and the absolute value sign can be ignored, meaning that it can be taken as a determinant. It admits therefore a fermionic path integral representation over anticommuting variables ψ and $\bar{\psi}$

$$\det \frac{\delta V}{\delta \xi} = \int \mathcal{D}(\psi \bar{\psi}) \exp - \int_{x_0}^x dy_1 \int_{x_0}^x dy_2 \bar{\psi}(y_1) \frac{\delta V(y_1)}{\delta \xi(y_2)} \psi(y_2) \quad (37)$$

The added fermionic part in the action reads

$$\int_{x_0}^x dy_1 \int_{x_0}^x dy_2 \bar{\psi}(y_1) \frac{\delta V(y_1)}{\delta \xi(y_2)} \psi(y_2) = \int_{x_0}^x dy \bar{\psi}(y) \left(\frac{1}{2} \frac{\partial}{\partial y} + \xi \right) \psi(y) \quad (38)$$

Hence we obtain an effective Lagrangian

$$\mathcal{L}_{\text{eff}} = \frac{1}{2} \left(\dot{\xi} + \xi^2 + E \right)^2 + \bar{\psi} \left(\frac{1}{2} \frac{\partial}{\partial y} + \xi \right) \psi \quad (39a)$$

For a general variation of ξ , ψ and $\bar{\psi}$

$$\begin{aligned} \delta \mathcal{L}_{\text{eff}} &= \left(\frac{1}{2} \left(\dot{\xi} + \xi^2 + E \right) \right) \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \delta \xi + \\ &\quad + \delta \bar{\psi} \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \psi + \bar{\psi} \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \delta \psi + \bar{\psi} \psi \delta \xi \end{aligned} \quad (40a)$$

This vanishes identically when

$$\delta \xi = \psi \delta \bar{a} \quad \delta \psi = 0 \quad \delta \bar{\psi} = \left(\frac{1}{2} \left(\dot{\xi} + \xi^2 + E \right) \right) \delta \bar{a} \quad (41a)$$

where $\delta \bar{a}$ is an anticommuting variable, as if $\mathcal{L}_{\text{eff}}$ were "supersymmetric". This is an example of a BRS supersymmetry. It occurs whenever a constraint is introduced in a path integral, with an accompanying Jacobian, and the latter is represented using a Gaussian integral over anticommuting variables. Here the constraint is the dynamical equation (9). While the BRS symmetry originated in the context of the quantization of gauge fields to control the gauge fixing procedure, it is really of a general nature, independent of the geometric origin of the problem.

The structure of the variations (41) enable one to define a derivative operator s , such that for any functional of the fields

$$s\mathcal{A} = \int dx \frac{\delta \mathcal{A}}{\delta \bar{a}} \delta \bar{a} \equiv \int dx s\mathcal{A} \delta \bar{a} \quad (42a)$$

$$s\xi = \psi \quad s\psi = 0 \quad s\bar{\psi} = \left(\frac{1}{2} \left(\dot{\xi} + \xi^2 + E \right) \right)$$

If one iterates s , which interchanges bosonic and fermionic fields, one has

$$s^2 \xi = 0 \quad s^2 \psi = 0 \quad s^2 \bar{\psi} = \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \psi \quad (43)$$

Thus s is nilpotent (of square equal to zero) provided ψ satisfies the equation of motion $(\frac{1}{2}d/dx + \xi)\psi = 0$. This is improperly called an on-shell condition, with reference to S -matrix concepts. In certain applications, it can be advantageous to define a genuine nilpotent operator, for instance when dealing with renormalization, where one studies counterterms satisfying $s\mathcal{A} = 0$. One knows *a priori* that an $\mathcal{A}$ of the form $s\mathcal{B}$ will solve this equation, and, under favorable circumstances, this is the most general solution. To ensure that $s^2 = 0$, the trick is to use a Fourier representation of the δ -functions which enforces the constraint. This amounts to adding a term in the Lagrangian with a Lagrange multiplier (field) α , initially pure imaginary, and to avoiding performing the integral over V . Thus the new effective Lagrangian reads

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}V^2 + \alpha \left(\frac{1}{2}(\dot{\xi} + \xi^2) + E - V \right) + \bar{\psi} \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \psi \quad (39b)$$

and its variation

$$\begin{aligned} \delta \mathcal{L}_{\text{eff}} = & (V - \alpha) \delta V + \delta \alpha \left(\frac{1}{2}(\dot{\xi} + \xi^2) + E - V \right) \\ & + \alpha \left(\frac{1}{2} \frac{d}{dx} \xi \right) \delta \xi + \delta \bar{\psi} \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \psi \\ & + \bar{\psi} \left(\frac{1}{2} \frac{d}{dx} + \xi \right) \delta \psi + \delta \xi \bar{\psi} \psi \end{aligned} \quad (40b)$$

vanishes identically when

$$\delta \bar{\psi} = \alpha \delta \bar{a} \quad \delta \xi = \psi \delta \bar{a} \quad \delta \psi = 0 \quad \delta \alpha = 0 \quad \delta V = 0 \quad (41b)$$

The operator s becomes

$$s \bar{\psi} = \alpha \quad s \xi = \psi \quad s \psi = 0 \quad s \alpha = 0 \quad s V = 0 \quad (42b)$$

and $s^2 = 0$. In this second form the definition of s is independent of the constraints. These are in fact coded in the effective Lagrangian. This hidden supersymmetry of the Fokker-Planck equation is only meaningful in the framework of a path integral representation.

10.1.3 The replica trick

The replica trick is a useful device when dealing with averages such as the one of the resolvent (4). Let us replace the complex

argument z by $E - i0$ with E real and use the shorthand notation E , with an infinitesimal imaginary part implicitly understood. Also write $G(E)$ for $\langle (x | [H - E]^{-1} | x) \rangle$. If, instead of operators in an infinite dimensional space, we were dealing with finite matrices, the quantity to be averaged would be the ratio of a minor by the determinant of the matrix. It therefore involves the random potential both in the numerator and denominator and leads *a priori* to untractable calculations. One looks for a mean to avoid this difficulty.

First, we interpret $(H - E)^{-1}$ as the propagator of an auxiliary scalar field. Using translational invariance of the average, we write

$$\begin{aligned} G(E) &= \lim_{L \rightarrow \infty} \left\langle \frac{1}{L} \int_{-\frac{1}{2}L}^{\frac{1}{2}L} dx \left(x \left| \frac{1}{H - E} \right| x \right) \right\rangle \\ &= \lim_{L \rightarrow \infty} \frac{2}{iL} \int_{-\frac{1}{2}L}^{\frac{1}{2}L} dx \\ &\quad \times \left\langle \frac{\int \mathcal{D}\varphi \varphi^2(x) \exp i \int_{-\frac{1}{2}L}^{\frac{1}{2}L} dx dy \varphi(x) (x | (H - E) | y) \varphi(y)}{\int \mathcal{D}\varphi \exp i \int_{-\frac{1}{2}L}^{\frac{1}{2}L} dx dy \varphi(x) (x | (H - E) | y) \varphi(y)} \right\rangle \\ &= \lim_{L \rightarrow \infty} \frac{2}{L} \frac{\partial}{\partial E} \left\langle \ln \left\{ \int \mathcal{D}\varphi \right. \right. \\ &\quad \left. \left. \exp i \int_{-\frac{1}{2}L}^{\frac{1}{2}L} dx dy \varphi(x) (x | (H - E) | y) \varphi(y) \right\} \right\rangle \quad (44) \end{aligned}$$

The i in the exponential is chosen in agreement with the $E - i0$ prescription to ensure convergence of the path integral. It would seem that we have made little progress, since we now have to take the average of a logarithm. However, one notices that, with n a continuous variable,

$$\frac{\partial}{\partial n} Z^n \xrightarrow{n \rightarrow 0} \ln Z \quad (45)$$

For any integer n , Z^n is the partition function of n identical noninteracting replicas. They involve the same random potential linearly in the action, and this allows an explicit average over a Gaussian distribution. Provided an analytic continuation in n is meaningful, it will be possible to obtain the desired quantity (44). The method relies on the existence of this analytic continuation. It is therefore interesting to pursue the exercise in the present case

under this assumption, to see whether it leads to an agreement with the previous result. The use of the replica trick in disordered problems has much in common with the technique used for polymer physics.

Let Φ denote an n -component field (of course n is thought at first as an integer). We have

$$G(E) = \lim_{L \rightarrow \infty} \frac{2}{L} \frac{\partial}{\partial E} \lim_{n \rightarrow 0} \frac{\partial}{\partial n} \left\langle \int \mathcal{D}\Phi \exp i \int \int dx dy \Phi(x) (x || [H - E] y) \Phi(y) \right\rangle \quad (46)$$

In effect the action reads

$$\int \int dx dy \Phi(x) (x || [H - E] y) \Phi(y) = \int dx \mathcal{L}(\Phi) \quad (47)$$

with a Lagrangian

$$\mathcal{L}(\Phi) = \frac{1}{2}(\partial\Phi)^2 + (V - E)\Phi^2 \quad (48)$$

The average over V is readily obtained for any integer n since it simply requires a Gaussian integral to be performed. We assume that the result continues to make sense for arbitrary n . It reads

$$G(E) = \lim_{L \rightarrow \infty} \frac{2}{L} \frac{\partial}{\partial E} \lim_{n \rightarrow 0} \frac{\partial}{\partial n} \int \mathcal{D}\Phi \exp i \int_{-\frac{1}{2}L}^{\frac{1}{2}L} dx \left[\frac{1}{2}(\partial\Phi)^2 - E\Phi^2 + \frac{i}{2}(\Phi^2)^2 \right] \quad (49)$$

In dimension higher than one, a similar expression would lead to a variant of φ^4 field theory. Luckily, in one dimension, we can turn this into a Schrödinger problem as follows. As $L \rightarrow \infty$ we expect the path integral to be dominated by a term of the form

$$\exp -\frac{1}{2}iLe(n) \quad (50)$$

Here $e(n)$ is the "ground state energy" of an associated Schrödinger equation, where the configuration variables Φ are the coordinates in a n -dimensional space, and once more x is interpreted as a temporal variable. Thus

$$e(n)\psi(\Phi) = \{-\Delta_\Phi + 2E\Phi^2 - i(\Phi^2)^2\} \psi(\Phi) \quad (51)$$

Since the potential term is complex, the notion of a ground state energy requires some qualification. If $e(n)$ admits an analytic

continuation, we expect that

$$e(n) = ne_0 + n^2e_1 + \dots \quad (52)$$

$$G(E) = -i \frac{\partial}{\partial E} e_0(E) \quad \pi\rho(E) = \frac{\partial}{\partial E} \text{Re}e_0(E) \quad (53)$$

Assuming that $e_0(E)$ vanishes for $E \rightarrow -\infty$, we obtain for the integrated density of states per unit length

$$\mathcal{N}(E) = \frac{1}{\pi} \text{Re}e_0(E) \quad (54)$$

We require that the ground state be $O(n)$ invariant. The appearance of a symmetry among identical replicas is quite natural, and no symmetry breaking is expected in this one-dimensional context. Let us use instead of the vector argument Φ a scalar variable

$$q = \frac{1}{2}\Phi^2 \quad (55)$$

and rewrite equation (51), expanding every quantity in powers of n , thus exhibiting the continuation in this variable. In particular we write $\psi = \psi_0 + n\psi_1 + \dots$, and

$$\begin{aligned} & \left\{ n \frac{d}{dq} + ne_0 + n^2e_1 + \dots \right\} \{\psi_0 + n\psi_1 + \dots\} \\ & = 2q \left\{ -\frac{d^2}{dq^2} + 2E - 2iq \right\} \{\psi_0 + n\psi_1 + \dots\} \end{aligned} \quad (56)$$

Hence

$$\begin{aligned} & \left\{ -\frac{d^2}{dq^2} + 2E - 2iq \right\} \psi_0 = 0 \\ & \left\{ e_0 + \frac{d}{dq} \right\} \psi_0 = 2q \left\{ -\frac{d^2}{dq^2} + 2E - 2iq \right\} \psi_1 \\ & \dots \end{aligned} \quad (57)$$

We look for a solution $\psi_0(q)$ bounded at the origin, together with its derivative, and vanishing at infinity. It follows from the second equation that e_0 is minus the logarithmic derivative of ψ_0 at the origin

$$e_0 = - \frac{d\psi_0/dq}{\psi_0} \Big|_{q=0} \quad (58)$$

The first equation in (57) admits a continuation for $q < 0$, provided we define it as

$$\psi_0(-q) = \psi_0(q)^* \quad (59)$$

a condition compatible with the equation. We normalize the solution by requiring

$$\psi_0(0) = 1 \quad (60)$$

which insures continuity at the origin. Taking into account equations (58) and (54), the discontinuity of the derivative of ψ_0 at the origin produces an inhomogeneous additional term when extending the equation for ψ_0 in the range $-\infty < q < +\infty$

$$\left\{ -\frac{d^2}{dq^2} + 2E - 2iq \right\} \psi_0(q) = 2\pi \mathcal{N}(E) \delta(q) \quad (61)$$

We recognize that this is a Fourier transformed version of equation (27), if we set

$$p(\xi) = \int \frac{dq}{2\pi} e^{-iq\xi} \psi_0(q) \quad (62)$$

The condition $\psi_0(0) = 1$ is equivalent to the statement that the integral over $p(\xi)$ is normalized to unity. Equation (59) implies that $p(\xi)$ is real (and positive) and equation (61) translates into a real first order equation for $p(\xi)$

$$\left(\xi^2 + 2E + 2\frac{d}{d\xi} \right) p(\xi) = \mathcal{N}(E) \quad (63)$$

Thus, in the limit $n \rightarrow 0$, the variables $q = \frac{1}{2}\Phi^2$ and ξ are conjugate, and $\psi_0(q)$ and $p(\xi)$ are Fourier transforms of each other, a slightly surprising result for which we do not have a simple interpretation. We conclude that the use of the replica trick has been fully justified in the present example.

i) Compute $\text{Im}e_0(E)$ and show that

$$\begin{aligned} e_0(E) &= \pi \mathcal{N}(E) - \frac{1}{2} i \frac{\mathcal{N}'(E)}{\mathcal{N}(E)} \\ G(E) &= -i \frac{d}{dE} e_0(E) = \frac{1}{2} \left(\frac{\rho(E)^2}{\mathcal{N}(E)^2} - \frac{\rho'(E)}{\mathcal{N}(E)} \right) - i\pi \rho(E) \end{aligned} \quad (64)$$

The solution $\psi_0(q)$ of equation (57) in the range $0 < q < \infty$, normalized to $\psi_0(0) = 1$, is a stationary point of the effective

action

$$S(\eta) = \int_0^\infty dq \left\{ \left(\frac{d\eta}{dq} \right)^2 + 2(E - iq)\eta^2 \right\} \quad (65)$$

If we substitute ψ_0 and integrate the kinetic term by parts, we find

$$\begin{aligned} S(\psi_0) &= -\psi_0(0) \frac{d\psi_0(0)}{dq} \\ &+ \int_0^\infty dq \psi_0(q) \left[-\frac{d^2}{dq^2} + 2(E - iq) \right] \psi_0(q) = e_0(E) \end{aligned} \quad (66)$$

The derivative $de_0(E)/dE$ is given by the partial derivative of $S(\eta)$ with respect to E evaluated for $\eta = \psi_0$

$$\begin{aligned} \frac{de_0(E)}{dE} &= 2 \int_0^\infty dq \psi^2(q) \\ &= 2 \int_0^\infty dq \int_{-\infty}^{+\infty} d\xi_1 d\xi_2 e^{iq(\xi_1 + \xi_2)} p(\xi_1) p(\xi_2) \\ &= 2i \int_{-\infty}^{+\infty} d\xi_1 d\xi_2 \frac{p(\xi_1) p(\xi_2)}{\xi_1 + \xi_2 + i\epsilon} \end{aligned} \quad (67)$$

Taking the real part leads to

$$\rho(E) = \text{Re} \frac{1}{\pi} \frac{de_0(E)}{dE} = 2 \int_{-\infty}^{+\infty} d\xi p(\xi) p(-\xi) \quad (68)$$

which reproduces formula (25). Let us now use the virial theorem in the form

$$\begin{aligned} 0 &= \int_0^\infty dq \frac{d\psi_0(q)}{dq} \left[-\frac{d^2}{dq^2} + 2(E - iq) \right] \psi_0(q) \\ &= -E\psi_0^2(0) + \frac{1}{2} \left(\frac{d\psi_0(0)}{dq} \right)^2 - 2i \int_0^\infty dq q \psi_0 \frac{d\psi_0}{dq} \end{aligned}$$

where the values at $q = 0$ are understood in the limit $q \rightarrow 0$. But $\psi_0(0)^2 = 1$, $(d\psi_0(0)/dq)^2 = e_0^2$, and an integration by part shows that the last integral is $-\frac{1}{2} de_0(E)/dE$ according to (67). Thus

$$i \frac{de_0}{dE} + e_0^2 = 2E \quad (69)$$

which is the required relation between the real part of $e_0(E)$, i.e. $\pi \mathcal{N}(E)$, and its imaginary part. Taking the imaginary part of both sides yields

$$\frac{d\mathcal{N}}{dE} + 2\mathcal{N} \text{Im}e_0 = 0$$

which agrees with equation (64).

- ii) Express $\mathcal{N}(E)$ with the help of the two independent solutions $Ai(x)$ and $Bi(x)$ of Airy's equation

$$\frac{d^2 u}{dx^2} = xu \quad (70)$$

and show that

$$\mathcal{N}(E) = \frac{2^{\frac{1}{2}}}{\pi^2 \left\{ Ai(-2^{\frac{1}{2}} E)^2 + Bi(-2^{\frac{1}{2}} E)^2 \right\}} \quad (71)$$

- iii) Show that the second coefficient $e_1(E)$ in the expansion (52) for $e(n)$ is given by

$$e_1(E) = - \int_0^\infty dq \psi_0(q) \frac{1}{2q} \left[e_0 + \frac{d}{dq} \right] \psi_0(q) \quad (72)$$

and that its value allows an estimate of the fluctuations of the spectrum around its mean value $\rho(E)$. More precisely in a box of size L , with L tending to infinity, let $\nu_L(E) dE$ be the total number of energy levels in the range $E, E + dE$. We have

$$\lim_{L \rightarrow \infty} \frac{1}{L} \langle \nu_L(E) \rangle = \rho(E)$$

$$\lim_{L \rightarrow \infty} \frac{1}{L} \{ \langle \nu_L(E_1) \nu_L(E_2) \rangle - \langle \nu_L(E_1) \rangle \langle \nu_L(E_2) \rangle \} = \sigma(E_1, E_2) \quad (73)$$

with

$$e_0(E) = \lim_{\epsilon \rightarrow +0} \frac{1}{i} \int dE' \rho(E') \ln(E' - E + i\epsilon)$$

$$e_1(E) = \lim_{\epsilon_1 \rightarrow +0} \frac{1}{2i} \int \int dE'_1 dE'_2 \sigma(E'_1, E'_2) \ln(E_1 - E + i\epsilon_1) \ln(E_2 - E + i\epsilon_2) \quad (74)$$

where we take the principal branch of the logarithm in the complex plane with a cut along the negative real axis. In particular for $E \rightarrow +\infty$

$$e_1 \sim \frac{-i}{4(2E)} + \frac{9}{16(2E)^{\frac{3}{2}}} + \dots \quad (75)$$

10.1.4 Random one-dimensional lattice

As a second example, we investigate the Laplacian on a random one-dimensional lattice. Such lattices have been introduced in

any dimension to restore continuous symmetries on average, while keeping an ultraviolet cutoff. We present a detailed discussion in chapter 11. For our purpose, the random lattice is only used as an elementary illustration of one-dimensional disordered discrete systems. The model is a variant of random harmonic chains studied by Dyson (1953).

Let us consider a set of N points drawn at random independently and with uniform probability in an interval L . In the limit where both N and L tend to infinity in a fixed ratio $\delta = 1/a = N/L$, δ being the density, the successive intervals ℓ between the ordered points become independent variables with a common Poisson distribution

$$p(\ell) d\ell = \frac{d\ell}{a} e^{-\ell/a} \quad (76)$$

Henceforth we use the microscopic scale a as a unit of length, such that effectively $\delta^{-1} = a = 1$. Our points generate a lattice and we take as a definition of the Laplacian of a scalar field φ defined on the ordered vertices

$$(\Delta\varphi)_n = \frac{\varphi_{n+1} - \varphi_n}{\ell_n} - \frac{\varphi_n - \varphi_{n-1}}{\ell_{n-1}} \quad (77)$$

We study the spectrum of eigenvalues $\Omega \geq 0$ such that

$$(\Delta + \Omega)\varphi = 0 \quad (78)$$

Let $\rho(\Omega) d\Omega$ denote the spectral density per unit length, i.e. per site in the present units. It satisfies the normalization condition

$$\int_0^\infty d\Omega \rho(\Omega) = 1 \quad (79)$$

which states that, for N sites and given boundary conditions, we have N eigenvalues. This reflects the existence of a short-distance cutoff, and implies that the spectral density decreases fast enough for large Ω . In contradistinction to the case of a regular lattice, the spectrum is here unbounded. We shall see that the high Ω tail results from strongly localized states around (unprobable) regions with small separations between lattice points.

We can transform the problem into one with a random potential on a regular lattice. Define

$$Q_n = \frac{\varphi_{n+1} - \varphi_n}{\ell_n} \quad (80)$$

with the obvious meaning that Q_n corresponds to the gradient of the field φ . These quantities satisfy the equation

$$Q_{n+1} - 2Q_n + Q_{n-1} = -\Omega \ell_n Q_n \quad (81)$$

which is a Schrödinger equation on a regular lattice with potential term $-\Omega \ell_n$, at zero effective energy, where Ω plays the role of coupling constant. Set

$$\omega = -\Omega \quad (82)$$

in such a way that for ω real and positive we are certainly off the spectrum and the solutions grow exponentially with n

$$Q_n \underset{n \rightarrow \infty}{\sim} \exp \gamma(\omega) n \quad \omega > 0 \quad (83)$$

A possible, but irrelevant, prefactor is omitted in (83).

Quantities such as $\gamma(\omega)$ are called Liapunov exponents, reference being made to the study of instabilities in classical dynamical systems (as previously, n may be thought of as an evolution variable). The knowledge of $\gamma(\omega)$ enables one to recover the spectral density. Assume Q_n to be given at two initial successive points, for instance corresponding to the values $n = -1$ and 0 . Let us solve equation (81) iteratively. It follows that for $n > 0$, Q_n is an n -th degree polynomial in ω . In the large n limit, the roots of this polynomial tend to (minus) the eigenvalues Ω

$$Q_n \underset{n \rightarrow \infty}{\sim} \prod_{s=1}^n (\omega + \Omega_s) = \exp \sum_{s=1}^n \ln(\omega + \Omega_s) \quad (84)$$

Thus

$$\begin{aligned} \gamma(\omega) &= \int_0^\infty d\Omega \rho(\Omega) \ln(\omega + \Omega) \\ \rho(\Omega) &= \frac{1}{\pi} \lim_{\epsilon \rightarrow +0} \text{Im} \frac{d\gamma}{d\omega}(-\Omega - i\epsilon) \end{aligned} \quad (85)$$

The second formula (85) assumes an analytic continuation of γ (with a cut along the negative real axis). As above, we transform the linear second order difference equation (81) into a first order nonlinear one, a discrete analog of Ricatti's equation. With $- \omega$

replacing Ω , we set

$$\begin{aligned} R_n &= \frac{Q_{n+1}}{Q_n} \\ R_n &= 2 + \omega \ell_n - \frac{1}{R_{n-1}} \end{aligned} \quad (86)$$

For $n \rightarrow \infty$, it follows from equation (83) that R_n is expected to behave as $\exp \gamma(\omega)$. This holds for a given realization of the set $\{\ell_n\}$. Thus on the average

$$\begin{aligned} \gamma(\omega) &= \lim_{n \rightarrow \infty} \langle \ln R_n \rangle \\ \rho(\Omega) &= \frac{1}{\pi} \text{Im} \frac{d}{d\omega} \lim_{n \rightarrow \infty} \langle \ln R_n \rangle \Big|_{\omega = -\Omega - i\epsilon} \end{aligned} \quad (87)$$

From equation (86), we can also derive an integral representation for the probability $P_n(R_n) dR_n$ that R_n takes values in the interval dR_n around R_n , given R_0 (equal to one say). This results from the observation that R_{n-1} only depends on the variables $\ell_{n-1}, \ell_{n-2}, \dots$ which are independent from ℓ_n . Thus

$$\begin{aligned} P_n(R_n) &= \int_0^\infty dR_{n-1} P_{n-1}(R_{n-1}) \\ &\quad \times \int_0^\infty d\ell e^{-\ell \delta} \left(R_n + \frac{1}{R_{n-1}} - 2 - \omega \ell \right) \end{aligned} \quad (88)$$

a discrete analog of the Fokker-Planck equation. The condition $R_0 = 1$ and the restriction $\omega > 0$ entail that the support of P_n is in the interval $[1, \infty]$.

For $n \rightarrow \infty$ assume that P_n tends to a stationary distribution P . The latter is given by the fixed point equation

$$\begin{aligned} P(R) &= \int_1^\infty dR' P(R') \int_0^\infty d\ell e^{-\ell \delta} \left(R + \frac{1}{R'} - 2 - \omega \ell \right) \\ &= \int_1^\infty dR' P(R') \frac{1}{\omega} \exp -\frac{1}{\omega} \left[R + \frac{1}{R'} - 2 \right] \theta(R + R'^{-1} - 2) \end{aligned} \quad (89)$$

The reasoning would be similar had the Poisson law for successive intervals be replaced by an arbitrary probability law. The difficulty in solving equation (89) is due to the θ -function under the integral sign. Once $P(R)$ is obtained from (89), the Liapunov

exponent is given by

$$\gamma(\omega) = \int_1^\infty dR P(R) \ln R \quad (90)$$

Of course P depends parametrically on ω . Rather than attempting to find an explicit solution of equation (89), it is interesting to understand approximation methods which have a broader range of applicability.

Obtain $\gamma_0(\omega)$ and $\rho_0(\Omega)$ in the absence of disorder, i.e. when ℓ_n instead of being a fluctuating quantity is fixed to its mean value equal to unity. Thus $P(R) = \delta(R - A)$ where A is the largest root of the equation corresponding to (86), namely

$$A^2 - (2 + \omega)A + 1 = 0 \quad A = 1 + \frac{1}{2}\omega + (\omega + \frac{1}{4}\omega^2)^{\frac{1}{2}} \quad (91)$$

Therefore

$$\begin{aligned} \gamma_0 &= \ln A \\ \rho_0(\Omega) &= \frac{1}{2\pi} \frac{1}{[\Omega(1 - \frac{1}{4}\Omega)]^{\frac{1}{2}}} \\ &\underset{\Omega \rightarrow 0}{\approx} \frac{1}{2\pi\Omega^{\frac{1}{2}}} [1 + \frac{1}{8}\Omega + \dots] \quad 0 < \Omega < 4 \end{aligned} \quad (92)$$

We shall study two opposite limits, the first one assumes weak disorder, the second a large frequency. If we pretend that the variables ℓ_n are weakly fluctuating, we expect to find an approximation valid for large wavelengths, hence small frequencies, which are the least sensitive to local disorder. We introduce a parameter λ by setting

$$\begin{aligned} \ell_n &= \langle \ell_n \rangle + \lambda(\ell_n - \langle \ell_n \rangle) = \langle \ell_n \rangle + \lambda z_n \\ R_n &= A_n \exp(\lambda B_n + \lambda^2 C_n + \lambda^3 D_n + \lambda^4 E_n + \dots) \end{aligned} \quad (93)$$

with $\langle z_n \rangle = 0$ and, as $n \rightarrow \infty$, A_n tends to the amplitude A given by (91). With these definitions and $\langle B \rangle, \langle C \rangle \dots$ standing for averages of $B_n, C_n, \dots$,

$$\gamma(\omega) = \gamma_0(\omega) + \lambda \langle B \rangle + \lambda^2 \langle C \rangle + \dots \quad (94)$$

Of course we are interested in the case $\lambda = 1$. Substitute the expression (93) (with A_n replaced by A) into equation (86). Using

(91) we find

$$\begin{aligned} A^2 [\exp(\lambda B_n + \lambda^2 C_n + \dots) - 1] \\ = \lambda(A-1)^2 z_n - [\exp - (\lambda B_{n-1} + \lambda^2 C_{n-1} + \dots) - 1] \end{aligned} \quad (95)$$

Identifying the coefficients of successive powers of λ , one derives the recursion relation

$$\begin{aligned} A^2 B_n &= (A-1)^2 z_n + B_{n-1} \\ A^2 (C_n + \frac{1}{2} B_n^2) &= C_{n-1} - \frac{1}{2} B_{n-1}^2 \\ &\dots \end{aligned} \quad (96)$$

which show the usefulness of the expansion for small $(A-1)$, i.e. small ω . Set

$$y = \frac{A-1}{A+1} z \quad (97)$$

then

$$\begin{aligned} \langle B \rangle &= \langle y \rangle = 0 & \langle C \rangle &= -\frac{1}{2} \langle y^2 \rangle & \langle D \rangle &= \frac{1}{3} \langle y^3 \rangle \\ \langle E \rangle &= -\frac{1}{4} \langle y^4 \rangle - \frac{1}{2} \frac{3 + 2A^2}{A^4 - 1} \langle y^2 \rangle^2 & \dots \end{aligned} \quad (98)$$

- i) Obtain equations (98)
- ii) Show that if

$$X_n = B_n + C_n + D_n + E_n + \dots \quad (99)$$

one has

$$\gamma(\omega) = \ln A + \langle X \rangle \quad (100)$$

$$\langle X \rangle = \langle \ln(1+y) \rangle - \frac{1}{2} \frac{3 + 2A^2}{A^4 - 1} \langle y^2 \rangle^2 + \dots \quad (101)$$

where the dots include terms involving products of two or more mean values of powers of the random variable y .

- iii) Show that the most singular behaviour in $(A-1)$ of terms of fifth and sixth order is

$$\begin{aligned} \langle F \rangle &= \frac{3}{A^2 - 1} \langle y^2 \rangle \langle y^3 \rangle + \dots \\ \langle G \rangle &= -\frac{15}{2} \frac{1}{(A^2 - 1)^2} \langle y^3 \rangle^2 + \dots \end{aligned} \quad (102)$$

From these expressions we derive the behaviour of $\gamma(\omega)$ for ω tending to zero. Since

$$A - 1 = \omega^{\frac{1}{2}} + \frac{1}{2}\omega + \frac{1}{8}\omega^{\frac{3}{2}} + \dots \quad (103)$$

we obtain up to and including terms of order ω^2

$$\begin{aligned} \gamma(\omega) = & \omega^{\frac{1}{2}} - \frac{1}{8}\omega \langle z^2 \rangle - \frac{1}{384}\omega^{\frac{3}{2}} \left[16 - 16 \langle z^3 \rangle + 15 \langle z^2 \rangle^2 \right] \\ & + \frac{1}{512}\omega^2 \left[16 \langle z^2 \rangle - 8 \langle z^4 \rangle + 24 \langle z^2 \rangle^2 + 24 \langle z^2 \rangle \langle z^3 \rangle - 15 \langle z^2 \rangle^3 \right] \\ & + O(\omega^{\frac{5}{2}}) \end{aligned} \quad (104)$$

This can now be specialized to the Poissonian case, setting $\lambda = 1$, with

$$\begin{aligned} \langle z^2 \rangle = 1 \quad \langle z^3 \rangle = 2 \quad \langle z^4 \rangle = 9 \quad \dots \\ \gamma(\omega) = \omega^{\frac{1}{2}} - \frac{1}{8}\omega + \frac{1}{384}\omega^{\frac{3}{2}} + \frac{1}{512}\omega^2 + \dots \end{aligned} \quad (105)$$

By analytic continuation

$$\begin{aligned} \operatorname{Re} \gamma(-\Omega - i\varepsilon) &= \frac{1}{8}\Omega + \frac{1}{512}\Omega^2 + O(\Omega^3) \\ \operatorname{Im} \gamma(-\Omega - i\varepsilon) &= \Omega^{\frac{1}{2}} - \frac{1}{384}\Omega^{\frac{3}{2}} + O(\Omega^{\frac{5}{2}}) \end{aligned} \quad (106)$$

The leading correction to the average spectral density is therefore

$$\rho(\Omega) = \frac{1}{2\pi\Omega^{\frac{1}{2}}} \left[1 - \frac{1}{128}\Omega + O(\Omega^2) \right] \quad (107)$$

The method is a systematic one and would allow the computation of higher order terms in the expansion of $\rho(\Omega)$. The comparison between (107) and (92) is instructive. In both cases, the dominant term is the one corresponding to the spectrum of the Laplacian in the continuum. The first correction is of opposite sign, being sixteen times smaller for the random lattice. Thus even in one dimension, the spectrum on a random lattice approximates the continuous spectrum at low frequency much better than on a regular one. One expects this effect to be much stronger the higher the dimension. This provides a definite motivation for the use of such lattices.

Equations (104) and (106) also reveal a new phenomenon. For a regular lattice the analytic continuation $\gamma(-\Omega - i\varepsilon)$ is pure imaginary as ε tends to zero and Ω is positive. For a random lattice, there appears a positive real part. The natural interpretation of this quantity is that it is the inverse of finite

localization length. With unit probability, for a frequency Ω belonging to the spectrum, not only do wavefunctions oscillate, but they also decrease exponentially. The behaviour of this localization length for small Ω is

$$L(\Omega) = \frac{8}{\Omega} - \frac{1}{8} + O(\Omega) \quad (108)$$

Naturally $L(\Omega)$ goes to infinity as $\Omega \rightarrow 0$, expressing that as one approaches the continuum limit the localization phenomenon disappears.

The magnitude of the localization length provides an empirical rule to obtain the range of validity of the weak disorder expansion. The latter makes sense as long as $L(\Omega)$ is large compared to the average spacing $\langle \ell_n \rangle$. We remark that, since ℓ_n can be arbitrarily small, the spectrum is unbounded.

To obtain some information for large Ω we need an alternative method. The above discussion suggests that, when Ω tends to infinity, states are strongly localized, so that γ is determined by small clusters of sites. We translate this for ω large by truncating the continued fraction expansion derived from equation (86)

$$R_n = 2 + \omega \ell_n - \frac{1}{2 + \omega \ell_{n-1} - \frac{1}{2 + \omega \ell_{n-2} - \dots}} \quad (109)$$

We factor out $2 + \omega \ell_n$ and take the average of $\ln R_n$ over the Poisson distribution. If $\gamma = 0.57722\dots$ stands for Euler's constant, we have

$$\begin{aligned} \langle \ln(2 + \omega \ell) \rangle &= \ln 2 + \left\langle \frac{\omega}{2 + \omega \ell} \right\rangle = \ln 2 \\ &+ e^{2/\omega} \left[\ln \frac{1}{2}\omega - \gamma + \frac{2}{\omega} - \frac{1}{2 \cdot 2!} \left(\frac{2}{\omega} \right)^2 + \frac{1}{3 \cdot 3!} \left(\frac{2}{\omega} \right)^3 - \dots \right] \end{aligned} \quad (110)$$

We set

$$C = e^\gamma = 1.78107\dots$$

and observe that for $p \geq 1$

$$\left\langle \frac{1}{(2 + \omega \ell)^{p+1}} \right\rangle = \frac{1}{p 2^p \omega} + O\left(\frac{1}{\omega^2}\right) \quad p \geq 1 \quad (111)$$

Thus

$$\gamma(\omega) = \ln 2 + \left\langle \frac{\omega}{2 + \omega \ell} \right\rangle - \sum_{p=1}^{\infty} \frac{1}{p} \left\langle \frac{1}{(2 + \omega \ell)^p} \right\rangle^2 - \left\langle \frac{1}{2 + \omega \ell} \right\rangle^2 \left\langle \frac{1}{(2 + \omega \ell)^2} \right\rangle + \dots \quad (112)$$

The last term, the first one to involve the distribution of three successive intervals, is of order $(\ln \omega)^3 / \omega^3$ and can be neglected in a computation up to and including terms of order ω^{-2} . We find therefore in the large ω limit

$$\gamma(\omega) = \ln \frac{\omega}{C} + \frac{2}{\omega} \left(\ln \frac{\omega}{2C} + 1 \right) - \frac{1}{\omega^2} \left[\left(\ln \frac{\omega}{2C} - 1 \right)^2 + \eta - 4 \right] + O \left(\left[\frac{\ln \omega}{\omega} \right]^3 \right) \quad (113)$$

with a constant η given by

$$\eta = \sum_{n=1}^{\infty} \frac{1}{4^n n^2 (n+1)^2} = 0.13070 \dots \quad (114)$$

We conclude that as $\Omega \rightarrow \infty$

$$\begin{aligned} \rho(\Omega) &= \frac{1}{\pi} \frac{d}{d\Omega} \text{Im} \gamma(-\Omega - i\epsilon) \underset{\Omega \rightarrow \infty}{\sim} \frac{2}{\Omega^2} + \frac{4}{\Omega^3} \left[\ln \frac{\Omega}{2C} - \frac{3}{2} \right] + \dots \\ \frac{1}{L(\Omega)} &= \text{Re} \gamma(-\Omega - i\epsilon) \underset{\Omega \rightarrow \infty}{\sim} \ln \frac{\Omega}{C} - \frac{2}{\Omega} \left(\ln \frac{\Omega}{C} + 1 \right) - \frac{1}{\Omega^2} \left[\left(\ln \frac{\Omega}{2C} - 1 \right)^2 + \eta - 4 - \pi^2 \right] + \dots \end{aligned} \quad (115)$$

The logarithms occurring in these expressions are due to the finite limit of the probability distribution $p(\ell)$ as $\ell \rightarrow 0$. For Ω large, the localization length tends to zero as $1/\ln \Omega$, thus justifying the method. One can obtain more terms in the expansion at the price of rather heavy computation.

The results can be compared to a numerical simulation as shown on figure 2. The behaviour of $\rho(\Omega)$ in $2/\Omega^2$ as $\Omega \rightarrow \infty$ is in agreement with the fact that $\rho(\Omega)$ is integrable.

- i) We have not made use of equation (89) which contains more information. It is possible to extract from this equation an

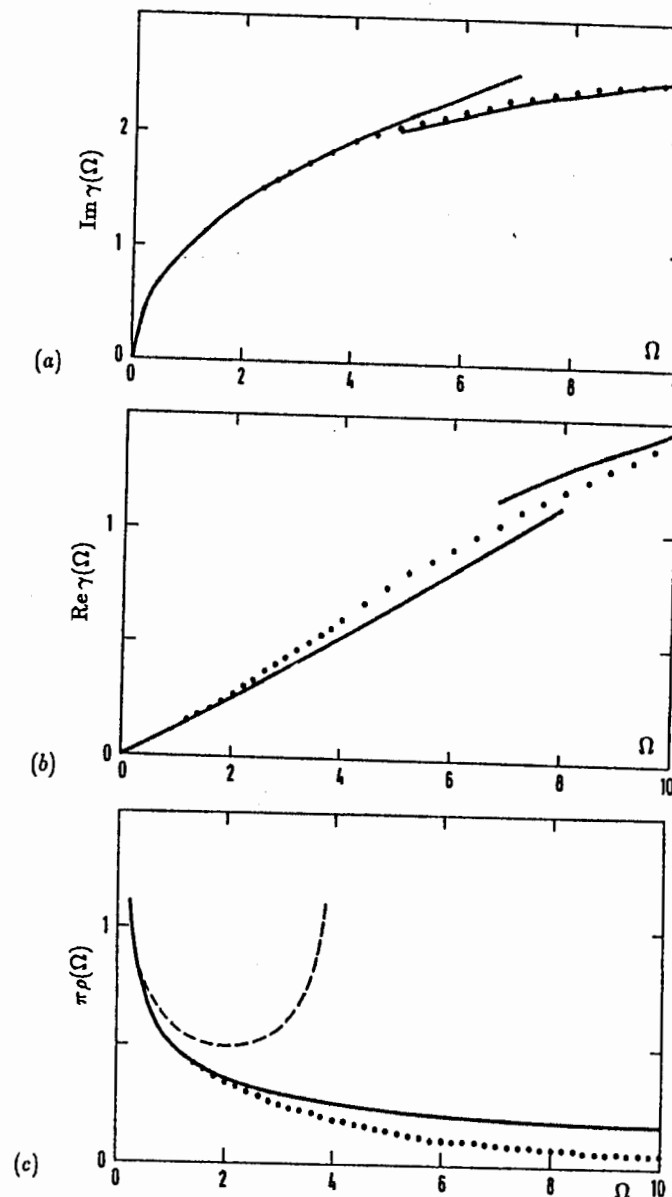


Fig. 2 (a) Comparison of small and large Ω expansions (full line) with the data (dots) for the integrated density. (b) Same comparison for the inverse correlation length. (c) The density of states in the continuum (full line), on a regular lattice (dashed line) and on a random lattice (dotted line).

analytic solution. For that purpose, set

$$R = 1 + x^{-1} \quad P(R) dR = p(x) dx \quad x \geq 0 \quad (116)$$

With these notations we have

$$\omega x^2 p(x) = \int_{\text{Sup}(0, x-1)}^{\infty} dx' p(x') \exp -\frac{1}{\omega} \left[\frac{1}{x} - \frac{1}{1+x'} \right] \quad (117)$$

and

$$\int_0^{\infty} dx p(x) = 1 \quad (118)$$

For x in the interval $(0,1)$

$$0 \leq x < 1 \quad p(x) = \frac{1}{\alpha \omega x^2} \exp -\frac{1}{\omega x} \quad (119)$$

where α is a positive constant to be determined from the normalization condition (118)

$$\frac{1}{\alpha} = \int_0^{\infty} dx p(x) \exp \left[\frac{1}{\omega(1+x)} \right] \quad (120)$$

To obtain $p(x)$ for $x \geq 1$, we trade $p(x)$ for the function $q(x)$ defined as

$$p(x) = \frac{1}{\alpha \omega x^2} \exp \left[-\frac{1}{\omega x} \right] q(x) \quad (121)$$

Thus $q(x)$ is equal to 1 for $0 \leq x < 1$ and is given in general by the integrodifferential equation

$$\exp \left(-\frac{1}{\omega(x+1)} \right) dq(x+1) + q(x) d \left(\exp \left[-\frac{1}{\omega x} \right] \right) = 0 \quad (122)$$

This shows that $q(x)$ is determined in the interval $n \leq x < n+1$ through its values for $0 \leq x < n$, from

$$q(x+1) = 1 - \int_0^x dy q(y) \frac{1}{\omega y^2} \exp -\frac{1}{\omega y(y+1)} \quad (123)$$

The function $q(x)$ is thus positive and decreasing. It behaves as $\omega^{1-x}/\Gamma(x)$ for $x \rightarrow \infty$. Also

$$\alpha = \sum_{n=1}^{\infty} q(n) e^{-1/n\omega} \quad (124)$$

Derive the expression for $\gamma(\omega)$ and re-obtain its high frequency expansion. Express the solution of equation (122) using a Laplace transform.

ii) Recover all the above results using the replica method.

10.2 Two-dimensional electron gas in a strong field

10.2.1 Landau levels - Quantum Hall effect

Let us illustrate the use of anticommuting variables on an example inspired by the study of the quantum Hall effect. The latter corresponds to the existence of plateaux in the Hall conductivity as a function of a (very strong) magnetic field at sufficiently low temperature. In such experiments, the electrons are trapped in a two-dimensional layer. The quantization of the Hall conductivity is accompanied by a sharp drop in the longitudinal resistance (or conductance!) showing that dissipative effects vanish at the center of the plateaux.

Let us recall the description of noninteracting electrons (charge e , mass m) constrained to move in the (x, y) plane, in the presence of a transversal field B . Energy levels are quantized and strongly degenerate (Landau). Choose a vector potential

$$A_x = -\frac{1}{2}By \quad A_y = \frac{1}{2}Bx \quad (125)$$

in the so-called symmetric gauge. The Hamiltonian reads

$$H_0 = \frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 = -\frac{\hbar^2}{2m} \left[\left(\partial_x + \frac{ie}{2\hbar}By \right)^2 + \left(\partial_y - \frac{ie}{2\hbar}Bx \right)^2 \right] \quad (126)$$

Let us pretend that the electronic charge is positive, and introduce dimensionless units such that

$$X = \sqrt{\frac{eB}{2\hbar}} x \quad Y = \sqrt{\frac{eB}{2\hbar}} y \quad (127a)$$

and trade the magnetic field for the Larmor frequency

$$\omega = \frac{eB}{2m} \quad (127b)$$

The Hamiltonian takes the form

$$H_0 = -\frac{1}{2}\hbar\omega \left[\left(\frac{\partial}{\partial X} + iY \right)^2 + \left(\frac{\partial}{\partial Y} - iX \right)^2 \right] \quad (128)$$

$$= \hbar\omega \left[\left(-\partial + \frac{1}{2}z \right) \left(\bar{\partial} + \frac{1}{2}z \right) + \frac{1}{2} \right]$$

where we use the complex notation

$$z = X + iY \quad \partial = \frac{\partial}{\partial z} \quad \bar{\partial} = \frac{\partial}{\partial \bar{z}} \quad (129)$$

The operators

$$a = \bar{\partial} + \frac{1}{2}z \quad a^\dagger = -\partial + \frac{1}{2}\bar{z} \quad (130)$$

satisfy the bosonic commutation relations

$$[a, a] = [a^\dagger, a^\dagger] = 0 \quad [a, a^\dagger] = 1 \quad (131)$$

Hence the Landau spectrum reads

$$E_\ell = \hbar\omega \left(\ell + \frac{1}{2}\right) \quad \ell = 0, 1, 2, \dots \quad (132)$$

corresponding to infinitely degenerate states

$$u_{\ell,m}(z, \bar{z}) = \frac{a^\ell}{\sqrt{\ell!}} u_{0,m}(z, \bar{z}) \quad \ell, m = 0, 1, 2, \dots \quad (133)$$

The ground state wavefunctions satisfy

$$a u_{0,m} = (\bar{\partial} + \frac{1}{2}z) u_{0,m} = 0$$

with orthonormal solutions

$$u_{0,m} = \frac{z^m}{\sqrt{m!}} \frac{e^{-z\bar{z}/2}}{\sqrt{\pi}} \quad m = 0, 1, 2, \dots \quad (134)$$

These functions generate the Fock-Bargmann space of entire functions, after factorization of $e^{-\frac{1}{2}z\bar{z}}/\sqrt{\pi} \equiv u_{0,0}$ (chapter 2 in volume 1). We absorb this factor into the measure and write

$$d\mu(z) = \frac{d^2z}{\pi} e^{-z\bar{z}} \quad u_{0,m} = \frac{e^{-\frac{1}{2}z\bar{z}}}{\sqrt{\pi}} v_m \quad (135)$$

$$\int d\mu(z) \overline{v_m(z)} v_{m'}(z) = \delta_{m,m'}$$

where it is understood that $d^2z = d\text{Re } z d\text{Im } z$.

- i) Give the explicit form of $u_{\ell,m}$ defined by equation (133).
- ii) Investigate the meaning of the quantum number m . What are the consequences of translational and rotational invariance? Pay attention to changes of gauge.
- iii) Show that for large m , the state $u_{0,m}$ corresponds to a probability distribution concentrated in an annulus of width $\Delta(r^2) \simeq \sqrt{m}$ around a mean radius $r^2 = m$

$$|u_{0,m}|^2 = \frac{e^{-r^2} r^{2m}}{\pi m!} \quad m \rightarrow \infty \quad \frac{1}{\pi} \frac{e^{-(r^2-m)^2/2m}}{\sqrt{2\pi m}} \quad (136)$$

Deduce that the number of independent states in the fundamental Landau level concentrated inside a circular disk of area

$\mathcal{A} = \pi R^2$ is given asymptotically by the ratio of the flux BA to the quantum of flux $\varphi_0 = h/e$

$$N = \frac{eBA}{2\pi\hbar} \quad (137)$$

if we return to original units. This result is independent of the choice of basis and of the shape of the surface provided N is large. Its interpretation is clear. The magnetic flux through the area $\mathcal{A}$, multiplied by the electric charge, is N times Planck's constant h . Alternatively, N is equal to the increase $(1/2\pi) \oint (e/\hbar) \mathbf{A} \cdot d\mathbf{x}$ in the phase of a wavefunction, measured in units of 2π , as one winds around the boundary of the sample. Show that these results extend to excited states.

What is significant, in the absence of any other interaction, is the macroscopic degeneracy of Landau levels whenever the magnetic length $\sqrt{\hbar}/eB$ (i.e. the square root of the area embracing a unit magnetic flux) is small compared to the overall scale (i.e. when N given by equation (137) is large). This quantum degeneracy leads to nontrivial quantum effects.

Consider an ideal situation where we neglect interactions among electrons as well as with the substrate. We also ignore magnetic contributions of electronic spins, the latter being assumed frozen along the strong magnetic field. Finally, pretend that the motion is constrained in the confining plane. Let us introduce an infinitesimal electric field $\mathbf{E}$ in this plane (hence perpendicular to the magnetic field $\mathbf{B}$). In classical terms, this field can be compensated for by a uniform displacement of the electron gas at a speed $\mathbf{v}$ adjusted in such a way that the Lorentz force $e(\mathbf{E} + \mathbf{v} \wedge \mathbf{B})$ vanishes. This is a purely kinematical effect, independent of the charge, such that

$$\mathbf{v} = \frac{\mathbf{E} \wedge \mathbf{B}}{B^2} \quad (138)$$

and generates a Hall current given by

$$\mathbf{j}_H = en\mathbf{v} = \frac{en}{B^2} \mathbf{E} \wedge \mathbf{B} \quad (139)$$

where n is the (surface) density of charges. This current does not lead to any dissipative effect ($\mathbf{j}_H \cdot \mathbf{E} = 0$). In matrix notation,

$$\mathbf{j}_H = \bar{\sigma}_H \mathbf{E} \quad (140)$$

with an antidiagonal conductivity matrix

$$\bar{\sigma}_H = \begin{pmatrix} 0 & en/B \\ -en/B & 0 \end{pmatrix} \quad (141)$$

One can also add to this expression a phenomenological diagonal ohmic term σ_0 to restore the effects of impurities and interactions. Thus the total conductivity matrix looks like

$$\bar{\sigma} = \begin{pmatrix} \sigma_0 & en/B \\ -en/B & \sigma_0 \end{pmatrix} \quad (142)$$

and its inverse, the resistivity, reads

$$\bar{\rho} = \bar{\sigma}^{-1} = \frac{1}{\sigma_0^2 + e^2 n^2 / B^2} \begin{pmatrix} \sigma_0 & -en/B \\ en/B & \sigma_0 \end{pmatrix} \quad (143)$$

In the limit where σ_0 is negligible, we have the apparent paradox that both σ_{xx} and ρ_{xx} tend simultaneously to zero. Let us look at the case where the chemical potential lies between two Landau levels, meaning that those levels with quantum number $\ell = 0, 1, \dots, L$ are totally occupied. According to equation (137), the electronic density is

$$n = \frac{eB}{2\pi\hbar}(L+1) \quad (144)$$

Hence the Hall constant, i.e. the antidiagonal term $\sigma_H = \sigma_{xy}$, is quantized according to

$$\sigma_H = \frac{en}{B} = \frac{e^2}{2\pi\hbar}(L+1) \quad (145)$$

Let ξ be the transverse section of a sample, and I the longitudinal current. The current density (ignoring edge effects) is $j = I/\xi$, while, if V_H is the transverse difference of potential, the electric field is $E = V_H/\xi$. As a result, and independently of the geometry or any other parameter,

$$\sigma_H = \frac{I}{V_H} = \frac{e^2}{h}(L+1) \quad (146)$$

These are the plateaux of the integral quantum Hall effect. Additional plateaux are observed at fractional occupation values of the Landau levels, and are believed to arise from interactions among electrons.

The quantized Hall effect, apart from being a spectacular manifestation of quantum mechanics at a macroscopic level,

important, if only because it involves fundamental constants and can lead to an improved accuracy on their measurement. What is truly remarkable is that the values (146) appear unmodified even in the presence of various impurities, (small) temperature effects, etc, and are observed with an amazing accuracy. The above description is of course somehow caricatural. One may naively think that a number of states become localized as a result of various kinds of defects – impurities, etc – and therefore cannot contribute to the conductivity. Arguments involving gauge invariance and topology show that the domain of validity of relation (146) extends beyond the ideal model and does apply to the real situation under much weaker assumptions (appendix A). However the width of the plateaux varies with experimental conditions and a complete theory should explain the observed behaviour.

We shall only study here how the density of states is affected by random impurities (Wegner 1983). It turns out that, with the help of simplifying assumptions, this problem admits an interesting analytic solution.

10.2.2 One particle spectrum in the presence of impurities

We want to understand how impurities affect the Landau levels. In practice, we shall limit ourselves to the lowest one with $\ell = 0$, and for a strong enough field, ignore excitations to higher levels, the gap between levels $\hbar\omega$ being much bigger than the perturbing potentials. Impurities are modelled by adding a random potential $V(x)$ to the Hamiltonian (126). Using reduced units defined in (127), we write

$$H = H_0 + V \quad (147)$$

One assumes that correlations in the potential are of short-range, and in the limiting case that the potential is uncorrelated at distinct points. The origin of energies is chosen in general by requiring that

$$\langle V \rangle = 0 \quad (148)$$

although in some instances it is more convenient to relax this condition. Our hypothesis on the locality of correlations translates

into the characteristic functional

$$\left\langle \exp -i \int d^2x \alpha(x) V(x) \right\rangle = \exp \int d^2x g(\alpha(x)) \quad (149)$$

The function $g(\alpha)$ represents the Fourier transform of a distribution of potential at a site

$$\exp g(\alpha) = \int dV P(V) e^{-i\alpha V} \quad (150)$$

Thus a Gaussian noise will correspond to

$$g(\alpha) = -\frac{1}{2} w \alpha^2 \quad (151)$$

i) A Poisson model of random impurities corresponds to a uniform density ρ of zero-range scattering centers (Friedberg and Luttinger, 1975). The probability density to find N impurities at points $\mathbf{x}_1, \dots, \mathbf{x}_N$, in an area $\mathcal{A}$ is given by

$$P(\mathbf{x}_1, \dots, \mathbf{x}_N) = e^{-\rho \mathcal{A}} \rho^N / N! \quad (152)$$

The corresponding potential of strength λ is

$$V(\mathbf{x}) = \lambda \sum_{i=1}^N \delta(\mathbf{x} - \mathbf{x}_i) \quad (153)$$

Show that

$$g(\alpha) = \rho [e^{-i\lambda\alpha} - 1] \quad (154)$$

How should one modify this expression, when the coupling constant are also random variables?

ii) A Lorentzian distribution of the potential

$$P(V) = \frac{\lambda}{\pi} \frac{1}{(V^2 + \lambda^2)} \quad (155)$$

leads to a nonanalytic function

$$g(\alpha) = -\lambda |\alpha| \quad (156)$$

As before, we obtain the average spectral density per unit area using the resolvent

$$\rho(E) = -\frac{1}{\pi} \text{Im} \left\langle \left(\mathbf{x} \left| \frac{1}{E + i0 - H} \right| \mathbf{x} \right) \right\rangle \quad (157)$$

The latter is in turn represented by a functional integral over a complex field φ through

$$\begin{aligned} & \left(\mathbf{x} \left| \frac{1}{E + i0 - H} \right| \mathbf{x}' \right) \\ &= \frac{1}{iZ} \int \mathcal{D}(\varphi \bar{\varphi}) \varphi(\mathbf{x}') \bar{\varphi}(\mathbf{x}) \exp i \int d^2x \bar{\varphi} (E + i0 - H) \varphi \end{aligned} \quad (158)$$

with a normalization factor

$$Z = \int \mathcal{D}(\varphi \bar{\varphi}) \exp i \int d^2x \bar{\varphi} (E + i0 - H) \varphi \quad (159)$$

This shorthand notation anticipates the fact that the effective action is the integral over a local density. Henceforth the infinitesimal imaginary part in E will be understood, but not written explicitly. It accounts for the choice of sign in the exponential. Instead of using the replica trick, in order to average over disorder, let us present a variant, which looks more satisfactory from a mathematical point of view. We represent Z^{-1} , which is nothing but $\det(E - H)$, by a path integral over an auxiliary complex fermionic field $(\psi, \bar{\psi})$

$$Z^{-1} = \det(E - H) = \int \mathcal{D}(\psi \bar{\psi}) \exp i \int d^2x \bar{\psi} (E - H) \psi \quad (160)$$

We may understand this by noting that a complex Fermi field counts for minus two real bosonic components.

Therefore

$$\begin{aligned} & \left(\mathbf{x} \left| \frac{1}{E - H} \right| \mathbf{x}' \right) = -i \int \mathcal{D}(\varphi \bar{\varphi} \psi \bar{\psi}) \varphi(\mathbf{x}) \bar{\varphi}(\mathbf{x}') \\ & \times \exp i \int d^2x [\bar{\varphi} (E - H_0) \varphi + \bar{\psi} (E - H_0) \psi - V(\bar{\varphi} \varphi + \bar{\psi} \psi)] \end{aligned} \quad (161)$$

In this form we can compute the average over V using equation (149) with $\alpha(\mathbf{x})$ replaced by $\bar{\varphi} \varphi + \bar{\psi} \psi$. Before doing so, we introduce the hypothesis that the potential should not excite transitions between distinct Landau levels, and that in (161) the only states to contribute correspond to the lowest level. This amounts to integrating only over fields φ and ψ satisfying the conditions

$$\begin{aligned} (E - H_0) \varphi &= \mathcal{E} \varphi \\ (E - H_0) \psi &= \mathcal{E} \psi \end{aligned} \quad \mathcal{E} = E - \frac{1}{2} \hbar \omega \quad (162)$$

Thus

$$\varphi = e^{-\frac{1}{2}z\bar{z}}u(z) \quad \psi = e^{-\frac{1}{2}z\bar{z}}v(z) \quad (163)$$

where u (bosonic) and v (fermionic) are holomorphic in z ($\bar{u}$ and $\bar{v}$ are antiholomorphic). We do not include a factor $\pi^{-1/2}$ in (163) in order to maintain a symmetry between bosonic and fermionic integrations. The Jacobian corresponding to a transformation on the basis of all Landau levels is unity, with fermionic contributions compensating bosonic ones. Of course, we truncate this expansion by keeping only the lowest level. After averaging over V , we obtain

$$\begin{aligned} & \left\langle \left(\mathbf{x} \left| \frac{1}{E-H} \right| \mathbf{x}' \right) \right\rangle \\ &= -i \exp -\frac{1}{2}(z\bar{z} + z'\bar{z}') \int \mathcal{D}(u\bar{u}v\bar{v}) u(z)\bar{u}(z') \exp S \end{aligned} \quad (164)$$

where the action S reads

$$S = \int d^2z e^{-z\bar{z}} \{ i\mathcal{E}(\bar{u}u + \bar{v}v) + g(e^{-z\bar{z}}(\bar{u}u + \bar{v}v)) \} \quad (165)$$

Use translational invariance (accompanied by gauge transformations) to show that, within the approximation to the lowest Landau level, all the spatial dependence of the average of the resolvent can be factored out as

$$\left\langle \left(\mathbf{x} \left| \frac{1}{E-H} \right| \mathbf{x}' \right) \right\rangle = \exp \{ z\bar{z}' - \frac{1}{2}z\bar{z} - \frac{1}{2}z'\bar{z}' \} \left\langle \left(\mathbf{x} \left| \frac{1}{E-H} \right| \mathbf{x} \right) \right\rangle \quad (166)$$

while

$$\left\langle \left(\mathbf{x} \left| \frac{1}{E-H} \right| \mathbf{x} \right) \right\rangle = -ie^{-z\bar{z}} \int \mathcal{D}(u\bar{u}v\bar{v}) u(z)\bar{u}(z) \exp S \quad (167)$$

is $\mathbf{x}$ -independent.

The action S possesses a supersymmetry which will become explicit if we introduce, besides the commuting coordinates $z, \bar{z}$, anticommuting ones θ and $\bar{\theta}$. We normalize the integral as

$$\int d\theta d\bar{\theta} e^{-\theta\bar{\theta}} = 1/\pi \quad (168a)$$

to maintain the parallel with

$$\int d^2z e^{-z\bar{z}} = \pi \quad (168b)$$

We define a holomorphic superfield ϕ (and its antiholomorphic conjugate $\bar{\phi}$), a function of z and θ which unites u and v , as

$$\phi(z, \theta) = u(z) + \theta v(z) \quad \bar{\phi}(z, \bar{\theta}) = \bar{u}(z) + \bar{\theta} \bar{v}(z) \quad (169)$$

Thus

$$S = i\pi\mathcal{E} \int d\theta d\bar{\theta} d^2z e^{-(z\bar{z} + \theta\bar{\theta})} \bar{\phi}\phi + \int d^2z g \left(e^{-z\bar{z}} \pi \int d\theta d\bar{\theta} e^{-\theta\bar{\theta}} \bar{\phi}\phi \right) \quad (170)$$

The first term is invariant under rotations in superspace, which contain transformations which do not mix z and θ as well as transformations which read in infinitesimal form

$$\begin{aligned} dz &= \bar{\omega}\theta & d\theta &= z\omega \\ d\bar{z} &= \bar{\theta}\bar{\omega} & d\bar{\theta} &= z\bar{\omega} \end{aligned} \quad (171)$$

and leave invariant the quadratic form $z\bar{z} + \theta\bar{\theta}$. Here $(\omega, \bar{\omega})$ is an "infinitesimal" anticommuting parameter. What is remarkable is that the second term of the action (170) is also invariant. To see this, let us assume that $g(\alpha)$ admits an expansion

$$g(\alpha) = \sum_{n=1}^{\infty} g_n \alpha^n \quad (172)$$

Indeed $g(0)$ vanishes as a consequence of the normalization of the probability distribution of the potential. If an expansion of the form (172) is not available, we approximate $g(\alpha)$ by functions with this property. We have a series of interaction terms of the type

$$\sum_1^{\infty} g_n e^{-n z\bar{z}} \left[\pi \int d\theta d\bar{\theta} e^{-\theta\bar{\theta}} \bar{\phi}\phi \right]^n \quad (173)$$

However the following identity holds

$$\left[\pi \int d\theta d\bar{\theta} e^{-\theta\bar{\theta}} \bar{\phi}\phi \right]^n = \frac{1}{n} \pi \int d\theta d\bar{\theta} e^{-n\theta\bar{\theta}} (\bar{\phi}\phi)^n \quad (174)$$

which can be checked by an explicit computation

Prove relation (174).

This introduces a function $h(\alpha)$, associated to $g(\alpha)$, through

$$h(\alpha) = \sum_{n=1}^{\infty} \frac{g_n}{n} \alpha^n = \int_0^{\alpha} \frac{d\beta}{\beta} g(\beta) \quad (175)$$

where the property $g(0) = 0$ is crucial to yield a finite $h(\alpha)$. With this notation, the action reads

$$S = \pi \int d\theta d\bar{\theta} d^2 z \left[i\bar{\epsilon} \bar{\phi} \phi e^{-(z\bar{z} + \theta\bar{\theta})} + h(\bar{\phi} \phi e^{(z\bar{z} + \theta\bar{\theta})}) \right] \quad (176)$$

exhibiting explicitly rotational invariance in superspace. This is at the origin of spectacular simplifications in the evaluation of the path integral which follows.

Derive from translational invariance in superspace of the path integral, with

$$\phi(z, \theta) \rightarrow \phi(z - a, \theta - \omega) \exp \left\{ z\bar{a} + \theta\bar{\omega} - \frac{1}{2}a\bar{a} - \frac{1}{2}\omega\bar{\omega} \right\}$$

that the two-point function takes the form

$$\langle \phi(z, \theta) \bar{\phi}(z', \bar{\theta}') \rangle = C \exp \{ z\bar{z}' + \theta\bar{\theta}' \} \quad (177)$$

where C is the quantity, independent from configuration variables, that we wish to evaluate

$$C = \left\langle \left(x \left| \frac{1}{E - H} \right| x \right) \right\rangle \quad (178)$$

Note that equation (177) means

$$\langle u(z) \bar{u}(z') \rangle = \langle v(z) \bar{v}(z') \rangle = C \exp z\bar{z}' \quad \langle u(z) \bar{v}(z') \rangle = 0 \quad (179)$$

To compute the two-point function $\langle \phi \bar{\phi}' \rangle$, we consider a perturbative expansion, making the best use of the supersymmetric holomorphic formalism. The propagator is the inverse of the operator occurring in the quadratic part of the action. For any superholomorphic function, we have the identity

$$f(z', \theta') = \int d^2 z d\theta d\bar{\theta} f(z, \theta) \exp \{ z'\bar{z} + \theta'\bar{\theta} - z\bar{z} - \theta\bar{\theta} \} \quad (180)$$

which generalizes the reproducing kernel in Bargmann's space. We deduce that the propagator is proportional to $\exp(z'\bar{z} + \theta'\bar{\theta})$. The interactions are given by the second term in the action (176). An interaction vertex of degree $2n$ contains a Gaussian factor $\exp -n \{ z\bar{z} + \theta\bar{\theta} \}$. We illustrate the Feynman rules by computing the contribution of the one-loop diagram to the mean value $\langle \phi \bar{\phi}' \rangle$ (figure 3).

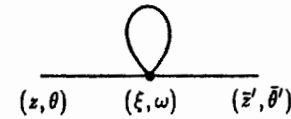


Fig. 3 One loop diagram contribution to the mean value $\langle \phi \bar{\phi}' \rangle$.

In the white noise case, i.e. when

$$g(\alpha) = -\frac{1}{2}w\alpha^2 \quad h(\alpha) = -\frac{1}{4}w\alpha^2 \quad (181)$$

the computation involves

i) the product of three propagators

$$(i/\pi\epsilon)^3 \exp \{ z\bar{\xi} + \theta\bar{\omega} + \xi\bar{\xi} + \omega\bar{\omega} + \xi\bar{z}' + \omega\bar{\theta}' \}$$

ii) a vertex factor

$$-\pi w \exp -2 \{ \xi\bar{\xi} + \omega\bar{\omega} \}$$

iii) a symmetry factor equal to unity (the field is complex)

iv) finally an integral of the variables ξ, ω , leading to

$$(i/\pi\epsilon)^3 (-\pi w) \exp \{ z\bar{z}' + \theta\bar{\theta}' \} \times \int d^2 \xi d\omega d\bar{\omega} \exp \{ -(\xi - z)(\bar{\xi} - \bar{z}') - (\omega - \theta)(\bar{\omega} - \bar{\theta}') \} \quad (182)$$

We obtain the expected exponential prefactor containing all the dependence on external variables. What remains is to perform the Gaussian integral in superspace. Its value is unity. This is due to an exact cancellation between a bosonic and a fermionic determinant which are identical by virtue of the symmetry between the variables z and θ . Here we find the previous integrals

$$\int d^2 \xi \exp -\xi\bar{\xi} = \pi \quad \int d\omega d\bar{\omega} \exp -\omega\bar{\omega} = 1/\pi \quad (183)$$

The phenomenon observed in this example extends to all orders. The combination of propagators and vertex factors leads to Gaussian integrals in superspace. In the exponential there appear quadratic and linear terms. When one factors out the expected prefactor, one recovers a supersymmetric Gaussian integral with value unity, as was the case above, and for the same reasons. In

short, any diagram yields $\exp(z\bar{z}' + \theta\bar{\theta}')$ times a factor which would result in a similar zero-dimensional theory. We can therefore substitute a simple integral for a path integral. The dimensional reduction $d \rightarrow d - 2$ is typical of the existence of supersymmetry in such questions. It was first noticed by Parisi and Sourlas in a different context (section 5).

The conclusion of this discussion is the following expression

$$\begin{aligned} \left\langle \left(\mathbf{x} \left| \frac{1}{E - H} \right| \mathbf{x} \right) \right\rangle &= \left\langle \phi(z, \theta) \overline{\phi(z', \theta')} \right\rangle \exp - \{z\bar{z}' + \theta\bar{\theta}'\} \\ &= \frac{\int d\varphi d\bar{\varphi} \varphi \bar{\varphi} \exp \pi \{i\mathcal{E}\bar{\varphi}\varphi + h(\bar{\varphi}\varphi)\}}{\int d\varphi d\bar{\varphi} \exp \pi \{i\mathcal{E}\bar{\varphi}\varphi + h(\bar{\varphi}\varphi)\}} \end{aligned} \quad (184)$$

The required quantity is obtained as the ratio of two integrals each one extending over the complex φ -plane. We recall that $\mathcal{E} = E - \frac{1}{2}\hbar\omega$. Since

$$\left\langle \left(\mathbf{x} \left| \frac{1}{E - H} \right| \mathbf{x} \right) \right\rangle = \frac{1}{i\pi} \frac{\partial}{\partial \mathcal{E}} \ln \int_0^\infty dt \exp \pi \{i\mathcal{E}t + h(t)\} \quad (185)$$

and the mean density of states is obtained in the explicit form

$$\rho(E) = \frac{1}{\pi^2} \text{Im} \frac{\partial}{\partial E} \ln \left[\int_0^\infty dt \exp \pi \left\{ i \left(E - \frac{1}{2}\hbar\omega \right) t + \int_0^t \frac{d\alpha}{\alpha} g(\alpha) \right\} \right] \quad (186)$$

where $g(\alpha)$ is given by equation (150) which defines the local disorder.

As an illustration, we evaluate the expression (186) for the examples introduced above. The reader will check the corresponding integrals. We also restore the dimensional factor

$$K = \frac{eB}{\hbar} \quad (187)$$

which up to now was set equal to unity.

i) White noise (figure 4)

$$\begin{aligned} g(\alpha) &= -w \frac{1}{2} \alpha^2 & \nu &= \sqrt{\frac{2\pi}{wK^2}} \left(E - \frac{1}{2}\hbar\omega \right) \\ \rho(E) &= \frac{K}{\pi^2} \sqrt{\frac{2}{w}} \frac{e^{\nu^2}}{1 + \left(2\pi^{-\frac{1}{2}} \int_0^\nu dx e^{x^2} \right)^2} \end{aligned} \quad (188)$$

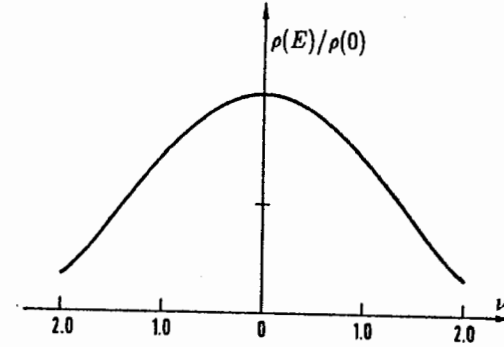


Fig. 4 Plot of the normalized density of states in the Gaussian case $\rho(E)/\rho(0)$ versus ν (equation (188)).

The total number of states is $\int dE \rho(E) = K^2/2\pi$. For large E , $\rho(E)$ decreases as

$$\rho(E) \underset{E \rightarrow \infty}{\sim} \frac{K}{\pi} \sqrt{\frac{2}{w}} \nu^2 e^{-\nu^2} \left(1 + O\left(\frac{1}{\nu^2}\right) \right) \quad (189)$$

in agreement with a semiclassical approximation. At the center of the band ($\nu \rightarrow 0$), we have

$$\rho(E) \underset{E \rightarrow \hbar\omega/2}{\sim} \frac{K}{\pi^2} \sqrt{\frac{2}{w}} \left[1 + \left(1 - \frac{4}{\pi} \right) \nu^2 + O(\nu^4) \right] \quad (190)$$

giving a distribution which resembles a Gaussian but is flatter.

ii) Poisson distribution of short-range scatterers

$$\begin{aligned} g(\alpha) &= \rho(e^{-i\lambda\alpha} - 1) & \nu &= \frac{2\pi}{\lambda K^2} \left(E - \frac{1}{2}\hbar\omega \right) & f &= \frac{2\pi}{K^2} \rho \\ \rho(E) &= \frac{1}{\pi\lambda} \text{Im} F(\nu) \end{aligned} \quad (191)$$

$$F(\nu) = \frac{\partial}{\partial \nu} \ln \left[\int_0^\infty dt \exp \left\{ i\nu t - f \int_0^t \frac{d\alpha}{\alpha} (1 - e^{-i\alpha}) \right\} \right]$$

We have assumed that the parameter λ , which characterizes the strength of the potential, is positive. The dimensionless ratio f is a measure of the impurity density in the natural unit of area embracing a unit magnetic flux.

One observes singularities when f goes through integral values, as shown on figure 5, where we have plotted the integrated density $N(E)$. The spectrum extends in the range

$\nu > 0$. As $\nu \rightarrow 0$, we find

$$\lambda \rho(E) \underset{\nu \rightarrow 0}{\sim} \begin{cases} (1-f)\delta(\nu) + A(f)\nu^{-f} + \dots & 0 < f < 1 \\ \frac{1}{\nu[(\ln \nu/\nu_0)^2 + \pi^2]} + \dots & f = 1 \\ B(f)\nu^{2-f} + \dots & 1 < f < 2 \\ B(2) + \dots & f = 2 \\ B(f)\nu^{f-2} + \dots & f > 2 \end{cases} \quad (192)$$

Evaluate the constants appearing in the above singular behaviour at $\nu \rightarrow 0$,

$$\begin{aligned} A(f) &= \frac{f \sin \pi f}{\Gamma(1-f)} \int_0^\infty \frac{dt}{t^f} \exp - \left\{ t + f \int_t^\infty \frac{d\alpha}{\alpha} e^{-\alpha} \right\} \quad 0 < f < 1 \\ \ln \nu_0 &= \int_0^\infty dt \exp - \left\{ \int_t^\infty \frac{d\alpha}{\alpha} e^{-\alpha} + 1 e^{-t} \right\} \\ 1/B(f) &= \Gamma(f-1) \int_0^\infty \frac{dt}{t^f} \exp - f \int_t^\infty \frac{d\alpha}{\alpha} e^{-\alpha} \quad f > 1 \end{aligned} \quad (193)$$

The most striking phenomenon occurs for $0 < f < 1$, where a fraction $(1-f)$ of the states in the lowest Landau level is unaffected by the presence of the scatterers. There are weaker singularities in the density of states at integrals values of ν as shown on figure 5.

iii) Lorentzian distribution (equations (155)–(156)).

The density of states is again a Lorentzian as was $P(V)$

$$g(\alpha) = -\lambda |\alpha| \quad \rho(E) = \frac{K^2}{2\pi^2} \frac{\lambda}{\lambda^2 + (E - \frac{1}{2}\hbar\omega)^2} \quad (194)$$

The above discussion, even though it leads to interesting results for the density of states, does not illuminate the question of the robustness of the quantum Hall effect. It does not either clarify what is exactly meant by localization in the presence of a magnetic field. In appendix A we present a partial answer to the first question by exhibiting the Hall conductance as a topological invariant. This is an elaboration of an argument originally presented by Laughlin.

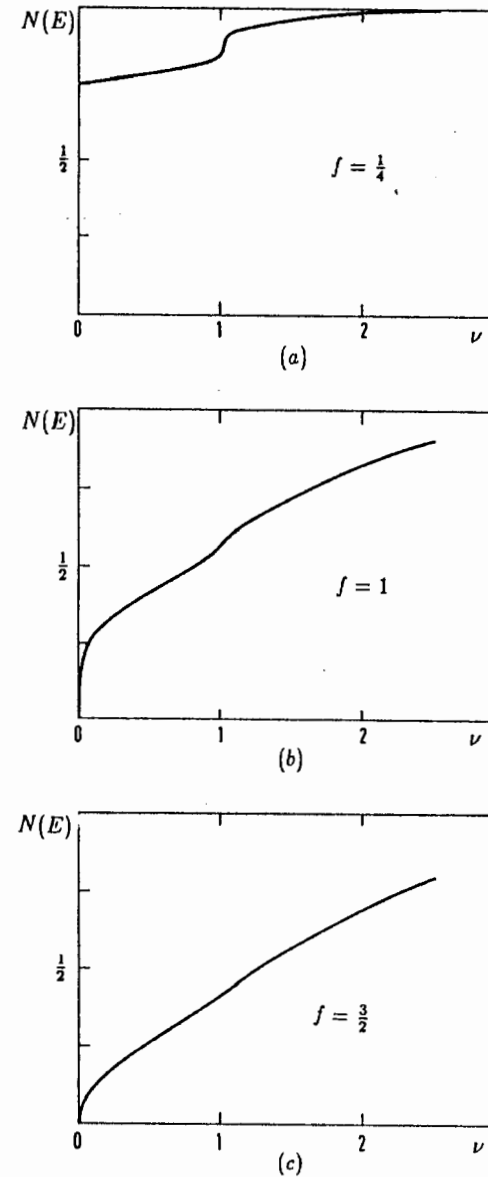


Fig. 5 The integrated density of states for the Poissonian model for various values of the density of impurities f . (a) $f = \frac{1}{4}$ (b) $f = 1$ (c) $f = \frac{3}{2}$.

10.3 Random matrices

Wigner, Dyson and others suggested comparing the spectrum of complex systems, such as heavy nuclei, with those of random Hamiltonians, with distribution laws as unbiased as possible, except for possible symmetries. The basic principle requires these distribution laws to be invariant under a transformation group derived from the existence of a Hilbert space. This implies in general the full unitary group, the real orthogonal group in the case where some reality condition (such as time reversal) is assumed, or else the unitary symplectic group. If one introduces an extra - ad hoc - hypothesis that matrix elements remain independent variables in any reference frame, the probability law becomes unique except for a scaling factor. Surprisingly enough, some characteristic properties of these ensembles, in particular the spacing between successive levels, give a fair representation of many observed spectra which range from atoms and nuclei to the zeroes of Riemann's ζ -function. Some universality is at work, of a nature specific to discrete systems, the precise mechanism of which has not yet been fully elucidated.

10.3.1 Semicircle law

We start with the simplest example of $N \times N$ Hermitian matrices. The matrix elements satisfy $H_{ij} = H_{ji}^*$, $1 \leq i, j \leq N$. We are interested in the limit $N \rightarrow \infty$. From the above hypothesis, the distribution law is a Gaussian

$$P(H) dH = \text{cst} \exp - \left\{ \frac{1}{2} \beta N \text{Tr} H^2 \right\} \prod_{i=1}^N dH_{ii} \times \prod_{1 \leq i < j \leq N} d\text{Re} H_{ij} d\text{Im} H_{ij} \quad (195)$$

The constant prefactor is determined by the normalization condition. The introduction of a factor N in the exponential implies that the large N limit is to be taken for fixed β . This will be justified later. The ensemble described by (195) is the unitary Gaussian ensemble. We are interested by the distribution of eigenvalues $p(\Lambda) d\Lambda$ where Λ is the real diagonal matrix of the eigenvalues λ_i , H being diagonalized with the help of a unitary matrix

U in the form

$$H = U^\dagger \Lambda U \quad (196)$$

To be specific, we can assume U to be of unit determinant although this still leaves $N - 1$ phases undetermined. To obtain $p(\Lambda) d\Lambda$ from $p(H) dH$, we express H in the vicinity of a diagonal Λ by writing $U = I + i\delta K$, where δK is an infinitesimal Hermitian traceless matrix. To first order, $H = \Lambda + i[\Lambda, \delta K]$. We introduce this parametrization in (195) and compute the corresponding Jacobian. Separating the integral over the (special) unitary group, we obtain

$$p(\Lambda) d\Lambda = \text{cst} \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^2 \exp \left[-\frac{1}{2} \beta N \sum_{i=1}^N \lambda_i^2 \right] \prod_{1 \leq i \leq N} d\lambda_i \quad (197)$$

Show that in the real orthogonal case the eigenvalue distribution is

$$p(\Lambda) d\Lambda = \text{cst} \prod_{1 \leq i < j \leq N} |\lambda_i - \lambda_j| \exp \left[-\frac{1}{2} \beta N \sum_{1 \leq i \leq N} \lambda_i^2 \right] \prod_{1 \leq i \leq N} d\lambda_i \quad (198)$$

while for the unitary symplectic ensemble

$$p(\Lambda) d\Lambda = \text{cst} \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^4 \exp \left[-\frac{1}{2} \beta N \sum_{1 \leq i \leq N} \lambda_i^2 \right] \prod_{1 \leq i \leq N} d\lambda_i \quad (199)$$

In all cases, we study the asymptotic properties of these distributions as N tends to infinity. These expressions suggest a statistical interpretation in terms of a one-dimensional Coulomb gas with logarithmic pair interactions. We define in general a partition function

$$Z(\alpha, \beta) = \frac{1}{N!} \int \prod_{1 \leq i \leq N} d\lambda_i \exp -S = \int_{\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N} \prod_{1 \leq i \leq N} d\lambda_i \exp -S$$

$$S = \frac{1}{2} \beta N \sum_{1 \leq i \leq N} \lambda_i^2 - \alpha \sum_{1 \leq i < j \leq N} \ln |\lambda_i - \lambda_j| \quad (200)$$

The values $\alpha = 1, 2, 4$ correspond respectively to the orthogonal, unitary and symplectic ensembles. The computation of Z gives the

normalization of the distribution. Observe that $\alpha > 0$ corresponds to a short-distance repulsive potential.

Rather than trying to compute Z explicitly, we obtain its asymptotic behaviour using the familiar saddle point method. We order the λ_i in increasing order as indicated in (200) and introduce a variable $x = i/N$ ranging through equally spaced values in the interval $[0, 1]$. For N large, we approximate x by a continuous variable, and write

$$\frac{1}{N} \sum_{1 \leq i \leq N} F(\lambda_i) = \frac{1}{N} \sum_{1 \leq i \leq N} F\left(\lambda\left(\frac{i}{N}\right)\right) \simeq \int_0^1 dx F(\lambda(x)) = \int_{-\infty}^{+\infty} d\lambda \rho(\lambda) F(\lambda) \quad (201)$$

defining a density $\rho(\lambda)$ of eigenvalues such that

$$d\lambda \rho(\lambda) = dx \quad \int_{-\infty}^{+\infty} d\lambda \rho(\lambda) = 1 \quad (202)$$

In this approximation the action takes the form

$$\frac{1}{N^2} S = \beta \int d\lambda \frac{1}{2} \lambda^2 \rho(\lambda) - \frac{1}{2} \alpha \iint d\lambda d\lambda' \rho(\lambda) \rho(\lambda') \ln |\lambda - \lambda'| \quad (203)$$

The stationary point in ρ yields a thermodynamic approximation for Z as well as for the distribution of eigenvalues. Using a Lagrange multiplier for the normalization constraint (202), we obtain

$$\frac{\lambda^2}{\lambda_0^2} = \int d\lambda' \rho(\lambda') \ln |\lambda - \lambda'| - \xi \quad (204)$$

$$\lambda_0^2 = 2\alpha/\beta > 0$$

provided that λ belongs to the support of $\rho(\lambda)$. The condition of normalization suggests that this support be bounded. We now seek for such a solution. Taking the derivative of both sides of equation (204) with respect to λ gives

$$2 \frac{\lambda}{\lambda_0^2} = \text{PP} \int d\lambda' \frac{\rho(\lambda')}{\lambda - \lambda'} \quad (205)$$

where the integration symbol means principal part and λ belongs to the support of ρ . It is convenient to introduce a function $F(z)$ analytic outside the support of ρ and behaving as z^{-1} at infinity,

through

$$F(z) = \int \frac{d\lambda' \rho(\lambda')}{z - \lambda'} \quad (206)$$

According to (205), when z tends to a point λ of the support of ρ with positive infinitesimal imaginary part, $F(\lambda + i0)$ tends to $2\lambda/\lambda_0^2 - i\pi\rho(\lambda)$, while outside this support $F(\lambda)$ is real. Clearly $\text{Im} F(z) < 0$ for $\text{Im} z > 0$. The ensembles are invariant under reflection $H \leftrightarrow -H$, so that $\rho(\lambda) = \rho(-\lambda)$. A unique function satisfying all the above properties and analytic in a plane cut along $[-\lambda_0, \lambda_0]$ is

$$F(z) = \frac{2}{z + \sqrt{z^2 - \lambda_0^2}} = \frac{2}{\lambda_0^2} \left[z - \sqrt{z^2 - \lambda_0^2} \right] \quad (207)$$

and the required asymptotic distribution with support in $[-\lambda_0, \lambda_0]$ is

$$\rho(\lambda) = \frac{2}{\pi \lambda_0} \sqrt{1 - \frac{\lambda^2}{\lambda_0^2}} \quad \lambda^2 \leq \lambda_0^2 \quad (208)$$

The solution (208) depends only on the ratio α/β , and in a reduced scale it is independent of the random ensemble, with $\alpha = 1, 2$ or 4. The initial choice of the factor N in the Boltzmann weight was designed to lead to this limiting distribution, called Wigner's semicircle law for obvious reasons, since

$$\left(\frac{1}{2} \pi \lambda_0 \rho(\lambda)\right)^2 + \left(\frac{\lambda}{\lambda_0}\right)^2 = 1$$

- i) Show that the distribution is correctly normalized.
- ii) Verify that the saddle point equation (204) is satisfied and compute $Z(\alpha, \beta)$. Starting from equation (205) and integrating in the interval $[-\lambda_0, \lambda_0]$ one obtains equation (204) with a constant ξ given by

$$\xi = \int_{-\lambda_0}^{\lambda_0} d\lambda \rho(\lambda) \ln |\lambda|$$

$$= \ln \lambda_0 + \frac{2}{\pi} \int_{-1}^{+1} d\mu \sqrt{1 - \mu^2} = \ln \lambda_0 + \frac{1}{2} [\psi(\frac{1}{2}) - \psi(2)]$$

where $\psi(x)$ is the logarithmic derivative of Euler's $\Gamma(x)$ function

$$\psi(x) = -\gamma + \sum_{n=0}^{\infty} \left(\frac{1}{n+1} - \frac{1}{n+x} \right)$$

and γ is Euler's constant. It follows that $\psi(2) = 1 - \gamma$ and $\psi(\frac{1}{2}) = -\gamma + 2 \ln \frac{1}{2}$. Thus

$$\xi = \ln \left(\frac{\lambda_0}{2e^{\frac{1}{2}}} \right) \quad (209)$$

Neglecting fluctuations around the saddle point value, the partition function is given by

$$Z(\alpha, \beta) \simeq \exp -S_0(\alpha, \beta) \quad (210a)$$

$$\begin{aligned} N^{-2} S_0(\alpha, \beta) &= \frac{1}{2} \beta \lambda_0^2 \int d\lambda \rho(\lambda) \left(\frac{\lambda^2}{\lambda_0^2} - \frac{1}{2} \int d\lambda' \rho(\lambda') \ln |\lambda - \lambda'| \right) \\ &= \frac{1}{4} \beta \int d\lambda \rho(\lambda) (\lambda^2 - \xi \lambda_0^2) = \frac{1}{8} \alpha \left[3 - \ln \frac{\alpha}{2\beta} \right] \end{aligned} \quad (210b)$$

10.3.2 The fermionic method

It is interesting to return to the previous discussion and show how one can make efficient use of Grassmann variables, one of the themes of this chapter. We consider as we did earlier the mean of the resolvent $(z-H)^{-1}$ in the unitary ensemble ($\alpha = 2$, $\lambda_0^2 = 4/\beta$). Set

$$\left\langle \left(\frac{1}{z-H} \right)_{jk} \right\rangle = \delta_{jk} F(z) \quad (211a)$$

The structure of the r.h.s. follows from the invariance under unitary transformations. Thus $NF(z)$ is the average of the trace. This entails

$$\rho(\lambda) = -\frac{1}{\pi} \text{Im} F(\lambda + i0) \quad (211b)$$

For $\text{Im} z > 0$, we express $F(z)$ as a mean value of the propagator for an N -component complex scalar field $(\varphi, \bar{\varphi})$. The corresponding denominator is replaced by a Grassmannian integral over an N -component fermionic field $(\psi, \bar{\psi})$. Consequently

$$\begin{aligned} F(z) &= -\frac{i}{N} \int \prod_{1 \leq j \leq N} \frac{d\varphi_j d\bar{\varphi}_j d\psi_j d\bar{\psi}_j}{\pi} \\ &\quad \bar{\varphi} \cdot \varphi \langle \exp i \{ \bar{\varphi}(z-H)\varphi + \bar{\psi}(z-H)\psi \} \rangle \end{aligned} \quad (212)$$

According to equation (197) in the Gaussian unitary ensemble

$$\langle \exp -i \text{Tr}(HK) \rangle = \exp -\frac{1}{2\beta N} \text{Tr} K^2 \quad (213)$$

We apply this identity to the matrix

$$\begin{aligned} (K)_{jk} &= \varphi_j \bar{\varphi}_k - \psi_j \bar{\psi}_k \\ \text{Tr} K^2 &= (\bar{\varphi} \cdot \varphi)^2 - (\bar{\psi} \cdot \psi)^2 + 2(\varphi \cdot \bar{\psi})(\bar{\varphi} \cdot \psi) \end{aligned}$$

Hence

$$\begin{aligned} F(z) &= -\frac{i}{N} \int \prod_{1 \leq j \leq N} \left(\frac{d\varphi_j d\bar{\varphi}_j d\psi_j d\bar{\psi}_j}{\pi} \right) \bar{\varphi} \cdot \varphi \\ &\quad \exp \{ iz(\bar{\varphi} \cdot \varphi + \bar{\psi} \cdot \psi) \\ &\quad - \frac{1}{2\beta N} [(\bar{\varphi} \cdot \varphi)^2 - (\bar{\psi} \cdot \psi)^2 + 2(\bar{\psi} \cdot \varphi)(\bar{\varphi} \cdot \psi)] \} \end{aligned} \quad (214)$$

To perform the fermionic integral, it is easier to write an integral representation of the quartic term

$$\exp \frac{1}{2\beta N} (\bar{\psi} \cdot \psi)^2 = \left(\frac{\beta N}{2\pi} \right)^{\frac{1}{2}} \int d\mu \exp -\{ \mu \bar{\psi} \cdot \psi + \frac{1}{2} \beta N \mu^2 \}$$

which enables one to perform the Grassmannian integral in the form

$$\int \prod_{1 \leq i \leq j \leq N} d\psi_j d\bar{\psi}_j \exp -\bar{\psi} M \psi = \det M$$

$$M_{jk} = (\mu - iz) \delta_{jk} + \frac{1}{\beta N} \varphi_j \bar{\varphi}_k$$

and

$$\det M = (\mu - iz)^{N-1} \left(\mu - iz + \frac{1}{\beta N} \bar{\varphi} \cdot \varphi \right)$$

As a result

$$\begin{aligned} F(z) &= -\frac{i}{N} \left(\frac{\beta N}{2\pi} \right)^{\frac{1}{2}} \int d\mu \prod_{1 \leq j \leq N} \frac{d\varphi_j d\bar{\varphi}_j}{\pi} (\bar{\varphi} \cdot \varphi) (\mu - iz)^{N-1} \\ &\quad \times \left(\mu - iz + \frac{1}{\beta N} \bar{\varphi} \cdot \varphi \right) \\ &\quad \times \exp \left\{ iz \bar{\varphi} \cdot \varphi - \frac{1}{2} \frac{1}{\beta N} (\bar{\varphi} \cdot \varphi)^2 - \frac{1}{2} \beta N \mu^2 \right\} \end{aligned} \quad (215)$$

We go over to polar coordinates in φ space using the angular integral

$$\int \prod_{1 \leq j \leq N} \left(\frac{d\varphi_j d\bar{\varphi}_j}{\pi} \right) \delta(\bar{\varphi} \cdot \varphi - 1) = \frac{2}{(N-1)!} \quad (216)$$

and obtain for $F(z)$ the following expression as a double integral, valid for any N ,

$$F(z) = -\frac{2i}{N!} \left(\frac{\beta N}{2\pi} \right)^{\frac{1}{2}} \int_{-\infty}^{+\infty} d\mu \int_0^{\infty} d\nu \nu^N (\mu - iz)^{N-1} \left(\mu - iz + \frac{\nu}{\beta N} \right) \exp \left\{ iz\nu - \frac{1}{2\beta N} \nu^2 - \frac{1}{2} \beta N \mu^2 \right\} \quad (217)$$

One of the two integrals can be performed in closed form. It is sufficient for our purpose to observe that for $N \rightarrow \infty$, we can use the saddle point method. We leave it to the reader to check that this leads readily to

$$F(z) = \frac{1}{2} \beta \left(z - \sqrt{z^2 - 4/\beta} \right) \quad (218)$$

which coincides with the value found in (207) for $\alpha = 2$ and $\lambda_0^2 = 4/\beta$. The above method has the advantage that it enables one to study the corrections to the saddle point method and in particular to investigate the exponentially small tails in the distribution of levels.

- i) Obtain the exponential tails of the level distribution.
- ii) Using the fermionic method find the asymptotic value for

$$\langle \text{Tr}(z_1 - H)^{-1} \text{Tr}(z_2 - H)^{-1} \rangle$$

10.3.3 Level spacings

While the semicircle law does not give a realistic description of most spectra, the distribution of level spacings predicted by the Gaussian ensembles seems on the contrary to have a much wider validity. We shall discuss in some detail the unitary case, leaving as an exercise the other two cases (orthogonal and symplectic).

We return to the probability law for levels ($\alpha = 2$)

$$p_N(\lambda_1, \dots, \lambda_N) = \frac{\prod_{j < k} (\lambda_j - \lambda_k)^2 \exp -\frac{1}{2} \beta N \sum_j \lambda_j^2}{\int \prod_j d\lambda_j \prod_{j < k} (\lambda_j - \lambda_k)^2 \exp -\frac{1}{2} \beta N \sum_j \lambda_j^2} \quad (219)$$

Since the scale is chosen in such a way that for fixed β as $N \rightarrow \infty$, the N levels remain with unit probability in a fixed interval, the level spacings are of order N^{-1} . If we are interested in the distribution of these spacings with a fixed average, we have to rescale the levels by a factor N .

Let $\chi(\lambda)$ be the characteristic function of an interval $[\lambda_0, \lambda_1]$

$$\chi(\lambda) = \begin{cases} 1 & \text{if } \lambda \in [\lambda_0, \lambda_1] \\ 0 & \text{if } \lambda \notin [\lambda_0, \lambda_1] \end{cases} \quad (220)$$

Therefore $1 - \chi$ is the characteristic function of the real axis with the interval $[\lambda_0, \lambda_1]$ deleted. The quantity

$$E(\lambda_0, \lambda_1) = \int \prod_{1 \leq j \leq N} [d\lambda_j (1 - \chi(\lambda_j))] p(\lambda_1, \dots, \lambda_N) \quad (221)$$

is the probability of finding no eigenvalue in the range $[\lambda_0, \lambda_1]$. The combination

$$E(\lambda_0, \lambda_1) - E(\lambda_0 - \delta\lambda_0, \lambda_1) - E(\lambda_0, \lambda_1 + \delta\lambda_1) + E(\lambda_0 - \delta\lambda_0, \lambda_1 + \delta\lambda_1) \simeq -\delta\lambda_0 \delta\lambda_1 \frac{\partial^2}{\partial \lambda_0 \partial \lambda_1} E(\lambda_0, \lambda_1)$$

gives the probability of finding no level in the interval $[\lambda_0, \lambda_1]$ but to find at least one in $[\lambda_1, \lambda_1 + \delta\lambda_1]$ as well as in $[\lambda_0 - \delta\lambda_0, \lambda_0]$. In the limit where $\delta\lambda_0, \delta\lambda_1$ tend to zero, each of the boundary interval will only contain one level with unit probability. Let us assume that in the appropriate scale, as $N \rightarrow \infty$, $E(\lambda_0, \lambda_1)$ only depends on the difference $\lambda_1 - \lambda_0$ and not on the midpoint $\frac{1}{2}(\lambda_0 + \lambda_1)$ which we shall subsequently choose at the center of the spectrum to simplify matters. We see that

$$q(\lambda_1 - \lambda_0) = \frac{\partial^2}{\partial (\lambda_1 - \lambda_0)^2} E(\lambda_1 - \lambda_0) \quad (222)$$

is the probability of finding two successive levels at a distance $\lambda_1 - \lambda_0$. Here the key word is successive. It is possible also to study the two level probability distribution and it is likely that its

short-distance behaviour coincides with the one for level spacings. The two quantities are unrelated at large separations.

To compute $E(\lambda_0, \lambda_1)$, we use an analogy between the probability law (219) and a one-dimensional gas of fermions in an harmonic potential. We stress that these fermions are unrelated to those appearing in the previous subsection. Indeed $p(\lambda_1, \dots, \lambda_N)$ is the (modulus) square of an antisymmetric wavefunction

$$\psi(\lambda_1, \dots, \lambda_N) = \mathcal{N}^{-1} \prod_{j < k} (\lambda_j - \lambda_k) \exp\left(-\frac{1}{4}\beta N \sum \lambda_i^2\right) \quad (223)$$

with $\mathcal{N}^{-1}$ a normalization factor. We can represent ψ as a Slater determinant constructed out of the first N eigenfunctions of an harmonic oscillator. They satisfy

$$\begin{aligned} (2a^\dagger a + 1)\psi_n &= (2n + 1)\psi_n & [a, a^\dagger] &= 1 \\ a &= \frac{1}{\sqrt{2}} \left(\sqrt{K}\lambda + \frac{1}{\sqrt{K}} \frac{\partial}{\partial \lambda} \right) & a^\dagger &= \frac{1}{\sqrt{2}} \left(\sqrt{K}\lambda - \frac{1}{\sqrt{K}} \frac{\partial}{\partial \lambda} \right) \\ K &= \frac{1}{2}\beta N \end{aligned} \quad (224)$$

Hence

$$\begin{aligned} \psi_0(\lambda) &= \left(\frac{K}{\pi}\right)^{\frac{1}{4}} e^{-K\frac{1}{2}\lambda^2} \\ \psi_n(\lambda) &= \frac{a^{\dagger n}}{\sqrt{n!}} \psi_0(\lambda) = \left(\frac{K}{\pi}\right)^{\frac{1}{4}} \left(\frac{2^n K^n}{n!}\right)^{\frac{1}{2}} (\lambda^n + \dots) e^{-K\frac{1}{2}\lambda^2} \\ \int_{-\infty}^{+\infty} d\lambda \psi_n(\lambda) \psi_m(\lambda) &= \delta_{n,m} \end{aligned} \quad (225)$$

Thus we find

$$\begin{aligned} \psi(\lambda_1, \dots, \lambda_N) &= \frac{1}{(N!)^{\frac{1}{2}}} \begin{vmatrix} \psi_0(\lambda_N) & \dots & \psi_{N-1}(\lambda_N) \\ \vdots & \ddots & \vdots \\ \psi_0(\lambda_1) & \dots & \psi_{N-1}(\lambda_1) \end{vmatrix} \\ \psi(\lambda_1, \dots, \lambda_N) &= \frac{1}{\left(\prod_{i=1}^N n!\right)^{\frac{1}{2}}} \left(\frac{K}{\pi}\right)^{\frac{1}{4}N} (2K)^{\frac{1}{4}N(N-1)} \\ &\quad \times \prod_{1 \leq j < k \leq N} (\lambda_j - \lambda_k) \exp\left(-\frac{1}{2}K \sum_{1 \leq j \leq N} \lambda_j^2\right) \end{aligned} \quad (227)$$

Hence $p(\lambda_1, \dots, \lambda_N)$ given by

$$p(\lambda_1, \dots, \lambda_N) = |\psi(\lambda_1, \dots, \lambda_N)|^2 \quad (228)$$

$$= \frac{1}{\prod_{i=1}^N n!} \left(\frac{K}{\pi}\right)^{\frac{1}{4}N} (2K)^{N(N-1)} \prod_{j < k} (\lambda_j - \lambda_k)^2 \exp -K \sum_j \lambda_j^2$$

satisfies the correct normalization condition

$$\int \prod_j d\lambda_j p(\lambda_1, \dots, \lambda_N) = 1 \quad (229)$$

Obtain the partition function $Z(\alpha = 2, \beta)$ using the above normalized probability.

Let us now turn to the computation of $E(\lambda_0, \lambda_1)$. From the fact that $(1 - \chi(\lambda))^2 = 1 - \chi(\lambda)$, it follows that if we set

$$f_n(\lambda) = (1 - \chi(\lambda))\psi_n(\lambda) \quad (230)$$

the expression for $E(\lambda_0, \lambda_1)$ reduces to

$$E(\lambda_0, \lambda_1) = \int \prod_{1 \leq j \leq N} d\lambda_j \frac{1}{N!} |\det f_k(\lambda_i)|^2$$

Expanding the determinant and integrating, one finds

$$E(\lambda_0, \lambda_1) = \det \left[\delta_{jk} - \int_{\lambda_0}^{\lambda_1} d\lambda \psi_j(\lambda) \psi_k(\lambda) \right]_{0 \leq j, k \leq N-1} \quad (231)$$

Gaudin has noticed that if one restricts the Darboux-Christoffel kernel

$$\begin{aligned} \mathcal{K}_N(\lambda, \lambda') &= \sum_{0 \leq j \leq N-1} \psi_j(\lambda) \psi_j(\lambda') \\ &= \sqrt{\frac{2N}{K}} \frac{\psi_N(\lambda) \psi_{N-1}(\lambda') - \psi_{N-1}(\lambda) \psi_N(\lambda')}{\lambda - \lambda'} \end{aligned} \quad (232)$$

to the interval $[\lambda_0, \lambda_1]$, and if one looks for its N eigenvalues ξ and eigenvectors φ in the subspace generated by the functions $\psi_j(\lambda)$, the eigenvalue equation

$$\xi \varphi(\lambda) = \int_{\lambda_0}^{\lambda_1} d\lambda' \mathcal{K}_N(\lambda, \lambda') \varphi(\lambda') \quad (233a)$$

can be rewritten

$$\xi \varphi^j = \sum_{0 \leq k \leq N-1} \int_{\lambda_0}^{\lambda_1} d\lambda \psi_j(\lambda) \psi_k(\lambda) \varphi^k \quad (233b)$$

when one sets $\varphi = \sum_{0 \leq j \leq N-1} \varphi^j \psi_j(\lambda)$. The characteristic equation takes the form

$$\prod_{1 \leq s \leq N} (\xi - \xi_s) = \det \left[\xi \delta_{jk} - \int_{\lambda_0}^{\lambda_1} d\lambda \psi_j(\lambda) \psi_k(\lambda) \right] \quad (234)$$

Consequently, $E(\lambda_0, \lambda_1)$ can be expressed in terms of the eigenvalues ξ_s as

$$E(\lambda_0, \lambda_1) = \prod_{1 \leq s \leq N} (1 - \xi_s) \quad (235)$$

We look for an asymptotic form for the kernel $\mathcal{K}_N(\lambda, \lambda')$ as N gets large. This is obtained using a semiclassical approximation for the functions $\psi_j(\lambda)$. We restrict ourselves to a domain of order $1/N$ around the origin and recall that $K = \frac{1}{2}\beta N$ is of order N . In the equation satisfied by $\psi_N(\lambda)$

$$\left[\frac{\partial^2}{\partial \lambda^2} - K^2 \lambda^2 + K(2N+1) \right] \psi_N(\lambda) = 0$$

we neglect $K^2 \lambda^2$ (of order one) as compared to $K(2N+1)$ (of order N^2). Thus for $\lambda(2NK)^{\frac{1}{2}}$ finite and $N \rightarrow \infty$, we find

$$\psi_N(\lambda) \sim \begin{cases} \psi_N(0) \cos \lambda(2NK)^{\frac{1}{2}} & N \text{ even} \\ \psi'_N(0)(2NK)^{-\frac{1}{2}} \sin \lambda(2NK)^{\frac{1}{2}} & N \text{ odd} \end{cases} \quad (236)$$

One can easily obtain $\psi_N(0)$ and $\psi'_N(0)$ in these expressions from the definitions (225). From $\sqrt{N+1}\psi_{N+1}(0) - \sqrt{N}\psi_{N-1}(0) = 0$ for N odd, we derive

$$\begin{aligned} \psi_{2s}(0) &= (-1)^s \left(\frac{K}{\pi} \right)^{\frac{1}{2}} \frac{(2s!)^{\frac{1}{2}}}{2^s s!} \underset{s \rightarrow \infty}{\sim} (-1)^s \left(\frac{K}{s\pi^2} \right)^{\frac{1}{2}} \\ \psi'_{2s+1}(0) &\underset{s \rightarrow \infty}{\sim} (-1)^s (4sK)^{\frac{1}{2}} \left(\frac{K}{s\pi^2} \right)^{\frac{1}{2}} \end{aligned} \quad (237)$$

Inserting these evaluations in (232), one gets asymptotically, and close to the origin,

$$\mathcal{K}_N(\lambda, \lambda') \underset{N \rightarrow \infty}{\sim} \frac{\sin \left[(2NK)^{\frac{1}{2}} (\lambda - \lambda') \right]}{\pi(\lambda - \lambda')} \quad (238)$$

According to equation (208), the density of levels at the center of the distribution is $\rho(0) = 2/\pi\lambda_0$ with $\lambda_0^2 = 4/\beta = 4N/\beta N$, hence

$$N\rho(0) = \frac{(2KN)^{\frac{1}{2}}}{\pi} = \frac{1}{D} \quad (239)$$

where D is the average level spacing. We can rewrite the asymptotic kernel $\mathcal{K}$, which only depends on the difference $\lambda - \lambda'$, in the form

$$\mathcal{K}(\lambda, \lambda') = \frac{1}{D} \frac{\sin[\pi(\lambda - \lambda')/D]}{[\pi(\lambda - \lambda')/D]} \quad (240)$$

Choose D as a scale for eigenvalues by setting

$$\lambda = Dx \quad (241)$$

and denote by $\varphi(x)$ the eigenfunction of $\mathcal{K}$ in an interval

$$-\frac{1}{2}t \leq x \leq \frac{1}{2}t \quad (t > 0)$$

We have

$$\xi \varphi(x) = \int_{-\frac{1}{2}t}^{\frac{1}{2}t} dx' \frac{\sin \pi(x - x')}{\pi(x - x')} \varphi(x') \quad (242)$$

The kernel is even, meaning that we can split the eigenfunctions into even, $\varphi^{(+)}(x)$, and odd ones, $\varphi^{(-)}(x)$. In both cases with $\varepsilon = \pm 1$

$$\varepsilon \xi^{(\varepsilon)} \varphi^{(\varepsilon)}(x) = \int_{-\frac{1}{2}t}^{\frac{1}{2}t} dx' \frac{\sin \pi(x + x')}{\pi(x + x')} \varphi^{(\varepsilon)}(x') \quad (243)$$

The identity

$$\frac{\sin \pi(x + x')}{\pi(x + x')} = \int_{-\frac{1}{2}t}^{\frac{1}{2}t} dx'' \frac{\exp 2i\pi x x''/t}{\sqrt{t}} \cdot \frac{\exp 2i\pi x'' x'/t}{\sqrt{t}} \quad (244)$$

shows that the kernel $\sin \pi(x + x')/\pi(x + x')$ is the square of the exponential kernel $t^{-\frac{1}{2}} \exp 2i\pi x x'/t$ in the interval $-\frac{1}{2}t \leq x \leq \frac{1}{2}t$. We can therefore choose $\varphi^{(\varepsilon)}(x)$ as an eigenvector of this second kernel with eigenvalues $\sqrt{t}\gamma^{(+)}$ or $i\sqrt{t}\gamma^{(-)}$

$$\begin{aligned} \int_{-\frac{1}{2}t}^{\frac{1}{2}t} \frac{dx'}{\sqrt{t}} \exp \frac{2i\pi x x'}{t} \varphi^{(\pm)}(x) &= \int_{-\frac{1}{2}t}^{\frac{1}{2}t} \frac{dx'}{\sqrt{t}} \left(\cos(2\pi x x'/t) \varphi^{(+)}(x') \right. \\ &\quad \left. + i \sin(2\pi x x'/t) \varphi^{(-)}(x') \right) \\ &= \begin{pmatrix} \sqrt{t}\gamma^{(+)} \varphi^{(+)}(x) \\ i\sqrt{t}\gamma^{(-)} \varphi^{(-)}(x) \end{pmatrix} \end{aligned} \quad (245)$$

We have the relation

$$\xi^{(\pm)} = t\gamma^{(\pm)^2} \quad (246)$$

By a change of variable

$$x = \frac{1}{2}ty \quad \varphi^{(\pm)}(x) = F^{(\pm)}(y) \quad (247)$$

this last system reads

$$\begin{aligned} \gamma^{(+)} F^{(+)}(y) &= \int_0^1 dy' \cos \frac{1}{2}\pi ty y' F^{(+)}(y') \\ \gamma^{(-)} F^{(-)}(y) &= \int_0^1 dy' \sin \frac{1}{2}\pi ty y' F^{(-)}(y') \end{aligned} \quad (248)$$

and

$$\begin{aligned} E^{(\pm)}(t) &= \prod_{s=1}^{\infty} (1 - t\gamma_s^{(\pm)^2}) \\ E(t) &= E^{(+)}(t)E^{(-)}(t) \end{aligned} \quad (249)$$

The γ 's are also functions of t . The latter is the separation of eigenvalues (of the original random matrix) in relative units. In these units, the relation (222) between $E(t)$ and $q(t)$ remains valid. The eigenvalue equations (248) appear in various other contexts, in particular in communication analysis.

To summarize, in the scale $x = \lambda/D$, and with a kernel $\mathcal{K}_t(x, x') = \sin \pi(x - x')/\pi(x - x')$ restricted to the interval $[-\frac{1}{2}t, \frac{1}{2}t]$ we have

$$\begin{aligned} E(t) &= \det(1 - \mathcal{K}_t) = \exp - \sum_1^{\infty} \frac{\text{Tr } \mathcal{K}_t^n}{n} \\ q(t) &= \frac{d^2 E(t)}{dt^2} \end{aligned} \quad (250)$$

For small t we can expand $\text{Tr } \mathcal{K}_t^n$ in powers of t , the leading term being t^n ,

$$\begin{aligned} \text{Tr } \mathcal{K}_t &= t \\ \text{Tr } \mathcal{K}_t^2 &= t^2 - \frac{1}{18}\pi^2 t^4 + \frac{1}{675}\pi^4 t^6 - \frac{1}{8820}\pi^6 t^8 + \dots \\ \text{Tr } \mathcal{K}_t^3 &= t^3 - \frac{1}{12}\pi^2 t^5 + \frac{1}{225}\pi^4 t^7 + \dots \\ \text{Tr } \mathcal{K}_t^4 &= t^4 - \frac{1}{9}\pi^2 t^6 + \frac{121}{16200}\pi^4 t^8 + \dots \\ \text{Tr } \mathcal{K}_t^5 &= t^5 - \frac{5}{36}\pi^2 t^7 + \dots \\ \text{Tr } \mathcal{K}_t^6 &= t^6 - \pi^2 t^8 \frac{1}{6} + \dots \end{aligned}$$

$$\text{Tr } \mathcal{K}_t^7 = t^7 + \dots$$

$$\text{Tr } \mathcal{K}_t^8 = t^8 + \dots \quad (251)$$

Thus

$$\begin{aligned} E(t) &= 1 - t + \frac{1}{36}\pi^2 t^4 - \frac{1}{675}\pi^4 t^6 + \frac{1}{17640}\pi^6 t^8 + O(t^{10}) \\ q(t) &= \frac{1}{3}(\pi t)^2 - \frac{2}{45}(\pi t)^4 + \frac{1}{315}(\pi t)^6 + O(t^8) \end{aligned} \quad (252)$$

The behaviour of $q(t)$ in t^2 for small t reflects the quadratic repulsive factor of eigenvalues at short distance, as shown in the probability law (219). The convergence of the expansion (252) is limited by the first zero in the determinant $E(t)$.

i) The similar result for the Gaussian orthogonal ensemble is given by $E^{(+)}(t)$ (equation (249)) with an expansion

$$\begin{aligned} E^{(+)}(t) &= 1 - t + \frac{1}{36}\pi^2 t^3 - \frac{1}{1200}\pi^4 t^5 + \frac{1}{8100}\pi^4 t^6 + \frac{1}{70560}\pi^6 t^7 + \dots \\ q^{(+)}(t) &= \frac{1}{6}\pi^2 t - \frac{1}{60}\pi^4 t^3 + \frac{1}{270}\pi^4 t^4 + \frac{1}{1680}\pi^6 t^5 + \dots \end{aligned} \quad (253)$$

Note that here $q^{(+)}(t)$ is linear in t for small t as expected.

ii) Show that asymptotically for $t \rightarrow \infty$

$$q(t) \sim e^{-\frac{1}{6}\pi^2 t^2} \quad q^{(+)}(t) \sim e^{-\frac{1}{18}\pi^2 t^2} \quad (254)$$

and try to find the prefactors.

The understanding of the universality class of these spacing laws, and knowledge whether there might exist systems exhibiting interesting deviations is still unresolved. At any rate there exists a wide range of systems exhibiting quantum fluctuations where the techniques and results of the Gaussian ensembles find interesting applications.

10.4 The planar approximation

The study of Gaussian random matrices is a particular case of matrix valued field theories when the size N of the matrices grows indefinitely. One would *a priori* expect many simplifications to arise in this limit, on the basis of the analogy with the large n -limit of the vector model. This is precisely what 't Hooft (1974) suggested in the context of gauge fields. He discussed the nature of the corresponding simplifications, showing that

it led to a new topological perturbative expansion, where the dominant term is the resummation of all diagrams which can be drawn on a plane (or by compactification on a sphere), hence the name planar approximation. Although this topic seems at first sight only loosely connected to the subject of this chapter, except by providing an extension of the Gaussian ensembles of random matrices, and extending some of the techniques used above, it turns out that it leads to an evaluation of the number of random triangulations of surfaces which ties in with another aspect of random systems (chapter 11). We shall limit ourselves to considerations in low dimension which enable one to exhibit the phenomena and to solve some combinatorial problems.

10.4.1 Combinatorics

We consider a matrix valued scalar field, denoted $M(x)$, an Hermitian $N \times N$ matrix. A general Lagrangian, invariant under the adjoint action of a global unitary group $SU(N)$ reads

$$\mathcal{L} = \frac{1}{2} \text{Tr} \partial_\mu M \partial_\mu M + \frac{1}{2} \text{Tr} M^2 + \sum_p g_p N^{1-\frac{1}{2}p} \text{Tr} M^p \quad (255)$$

We shall justify later the choice of a factor $N^{1-\frac{1}{2}p}$ associated to a term of order p . Let us show that, for fixed coupling constants and in the large N limit, only planar diagrams survive in the perturbative expansion. The crucial part in the Feynmann rules is the one giving the factors associated to the matrix nature of the field. The remaining rules are the familiar ones in a scalar theory, propagators, integrals, etc. It is convenient to represent the propagators of the field M by double lines, with an opposite orientation on each one, associated to the propagation of each of the indices. Let us study the N -dependence of a diagram with P propagators and V_p vertices of type p ($V = \sum_{3 \leq p} V_p$). To simplify, we look at connected vacuum diagrams. If one counts each propagator twice, one counts p times the vertices of type p , hence

$$2P = \sum_{3 \leq p} p V_p \quad (256)$$

Furthermore each of the B internal loops may be considered as the boundary of an oriented polyhedral surface. Joinging these faces

along the propagators, one obtains a closed oriented connected surface, satisfying Euler's formula

$$\chi = 2 - 2H = V - P + B \quad (257)$$

where χ is the Euler characteristic and we have denoted the genus by H (for handles), to avoid confusion with the coupling constants g_p . After summation over internal indices, each loop produces a factor N , so that, apart from the standard factors, the contribution of a diagram contains the product

$$\prod_{3 \leq p} (g_p N^{1-\frac{1}{2}p})^{V_p} N^B = \left(\prod_{3 \leq p} g_p^{V_p} \right) N^{2-2H} \quad (258)$$

In the limit $N \rightarrow \infty$, the leading contribution to the free energy (the generating function of connected vacuum diagrams) arises from diagrams drawn on a surface of genus zero ($\chi = 2$, $H = 0$), i.e. the planar ones. Corrections of order N^{-2} correspond to diagrams drawn on a torus, etc. The above reasoning can be generalized when source terms are added to the Lagrangian, provided that they carry the appropriate powers of N .

We set

$$E(g) = - \lim_{N \rightarrow \infty} \frac{1}{N^2} F(g) = - \lim_{N \rightarrow \infty} Z_N(g) \quad (259)$$

As an illustration, we consider a φ^4 model ($g_4 \equiv g$), and enumerate the weights corresponding to the first few planar diagrams up to order g^3

$$\begin{array}{llllll} \text{Diagram 1} & 2g & & & & \\ \text{Diagram 2} & 2g^2 & \text{Diagram 3} & 16g^2 & & \\ \text{Diagram 4} & \frac{32}{3}g^2 & \text{Diagram 5} & 64g^3 & \text{Diagram 6} & 128g^3 \\ \text{Diagram 7} & & & & \text{Diagram 8} & \frac{256}{3}g^3 \end{array} \quad (260)$$

Extend the previous considerations to the (quartic) Lagrangian

$$\mathcal{L} = \text{Tr} \partial_\mu M \partial_\mu M^\dagger + \text{Tr} M M^\dagger + \frac{\alpha g}{2N} \text{Tr} M M^\dagger M M^\dagger \quad (261a)$$

where

$$\begin{aligned} \alpha = 1 & \quad M \text{ real orthogonal} \\ \alpha = 2 & \quad M \text{ Hermitian} \\ \alpha = 4 & \quad M \text{ complex} \end{aligned} \quad (261b)$$

with corresponding symmetry groups $SO(N)$, $SU(N)$ and $SU(N) \times SU(N)$. The leading terms are the same in all cases, but the dependence on N of subleading corrections varies according to the different choices of matrices.

In the zero-dimensional case, $E(g)$ is a generating function for combinatorial factors such as those shown in equation (260). The method used in the preceding section can be extended to evaluate this quantity. We present the computation for a quartic interaction, but there is no difficulty to include any polynomial interaction, even when it is unbounded from below. It is understood that the large N limit will be taken at the end of the calculation. Up to a normalizing g -independent constant, we have

$$\begin{aligned} \exp -N^2 E(g) &= \text{cst} \int dM \exp - \left\{ \frac{1}{2} \text{Tr} M^2 + \frac{g}{N} \text{Tr} M^4 \right\} \\ &= \int \prod_{1 \leq i \leq N} d\lambda_i \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^2 \\ &\quad \times \exp - \left\{ \sum_{1 \leq i \leq N} \left(\frac{1}{2} \lambda_i^2 + \frac{g}{N} \lambda_i^4 \right) \right\} \end{aligned} \quad (262)$$

with $\{\lambda_i\}$ standing for the N real eigenvalues of the Hermitian matrix M . The saddle point method yields

$$E(g) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \left\{ \sum_{1 \leq i \leq N} \left(\frac{1}{2} \lambda_i^2 + \frac{g}{N} \lambda_i^4 \right) - \sum_{1 \leq i \neq j \leq N} |\lambda_i - \lambda_j| \right\} \quad (263a)$$

with the stationary condition

$$\frac{1}{2} \lambda_i + \frac{2g}{N} \lambda_i^3 = \sum_{1 \leq j \leq N} \frac{1}{\lambda_i - \lambda_j} \quad (263b)$$

We order the eigenvalues and, in the large N limit, substitute to the discrete sequence of λ_i 's the continuous function λ_l

$\sqrt{N} \lambda(i/N)$. This translates into the approximation

$$\begin{aligned} E(g) &= \int_0^1 dx \frac{1}{2} \lambda^2(x) + g \lambda^4(x) - \int_0^1 \int_0^1 dx dy \ln |\lambda(x) - \lambda(y)| \\ \frac{1}{2} \lambda(x) + 2g \lambda^3(x) &= PP \int_0^1 dy \frac{1}{\lambda(x) - \lambda(y)} \end{aligned} \quad (264)$$

We define the normalized density of eigenvalues $u(\lambda)$, according to

$$dx = u(\lambda) d\lambda$$

with $u(\lambda)$ even positive, with support in an interval $[-2a, 2a]$ satisfying

$$1 = \int_0^1 dx = \int_{-2a}^{+2a} d\lambda u(\lambda) \quad (265)$$

Thus

$$\frac{1}{2} \lambda + 2g \lambda^3 = PP \int_{-2a}^{2a} dx \frac{u(\mu)}{\lambda - \mu}, \quad -2a \leq \lambda \leq 2a \quad (266)$$

The analytic function

$$F(z) = \int_{-2a}^{2a} \frac{d\mu u(\mu)}{z - \mu} \quad (267a)$$

is odd, defined in the complex plane cut along $[-2a, 2a]$, real on the real axis outside this interval. It behaves as z^{-1} as z goes to infinity and satisfies

$$F(\lambda \pm i0) = \frac{1}{2} \lambda + 2g \lambda^3 \mp i\pi u(\lambda) \quad -2a \leq \lambda \leq 2a \quad (267b)$$

The unique solution is

$$F(z) = \frac{1}{2} z + 2gz^3 - \left(\frac{1}{2} + 4ga^2 + 2gz^2 \right) \sqrt{z^2 - 4a^2} \quad (268)$$

The prefactor of the square root is such that $F(z)$ decreases as z^{-1} for z infinite. This requires that

$$12ga^4 + a^2 - 1 = 0 \quad (269)$$

For small g , this means that a^2 has a convergent expansion

$$a^2 = \frac{1}{24g} \left[(1 + 48g)^{\frac{1}{2}} - 1 \right] = 1 - 12g + 2(12g)^2 - 5(12g)^3 + \dots \quad (270)$$

In equation (268), the square root is chosen to be positive for real z larger than $2a$. The discontinuity of F yields an expression for $u(\lambda)$ generalizing Wigner's semicircle law

$$u(\lambda) = \frac{1}{\pi} \left\{ \frac{1}{2} + 4ga^2 + 2g\lambda^2 \right\} \sqrt{4a^2 - \lambda^2} \quad (271)$$

Returning to (264), one finds $E(g)$ as

$$E(g) - E(0) = \frac{1}{24}(a^2 - 1)(9 - a^2) - \frac{1}{2} \ln a^2 \quad (272)$$

From the expression of a^2 as a function of g , this produces the series

$$\begin{aligned} E_0(g) \equiv E(g) - E(0) &= - \sum_1^{\infty} (-12g)^k \frac{(2k-1)!}{k!(k+2)!} \\ &= 2g - 18g^2 + 288g^3 - 6048g^4 + \dots \end{aligned} \quad (273)$$

in agreement with the first few terms exhibited in equation (260).

In contradistinction to the complete perturbative series, the planar approximation enjoys the properties that the corresponding expansion is convergent, hence analytic around zero coupling $g = 0$. The nearest singularity arises from the expression of $a^2(g)$ and is located on the real negative axis at

$$g_c = -\frac{1}{48} \quad (274)$$

As a result we obtain the asymptotic behaviour of the coefficients in the expansion (273)

$$E_0(g) = - \sum_1^{\infty} A_k (-g)^k \quad A_k \sim \frac{1}{2\sqrt{\pi}} (48)^k k^{-\frac{1}{2}} \quad (275)$$

i) The same method yields expressions for "Green functions" in the planar approximation. Thus

$$G_{2p}(g) = \langle \text{Tr } M^{2p} \rangle = \int_{-2a}^{+2a} d\lambda u(\lambda) \lambda^{2p} \quad (276)$$

Mean values of odd powers obviously vanish. From

$$zF(z) = \int_{-2a}^{+2a} d\lambda u(\lambda) \frac{z}{z-\lambda} = \sum_{p=0}^{\infty} z^{-2p} G_{2p}(g)$$

we derive

$$\begin{aligned} G_{2p}(g) &= \frac{(2p)!}{p!(p+2)!} a^{2p} (2p+2 - pa^2) \\ &= \frac{(2p)!}{p!(p-1)!} \sum_{k=0}^{\infty} \frac{(2k+p-1)!}{k!(k+p+1)!} (-12g)^k \end{aligned} \quad (277)$$

Similarly one can generate connected Green functions G_{2p}^c , defined recursively according to

$$G_{2p} = \sum_{\substack{r_q \geq 0 \\ \sum_q r_q = p}} \frac{(2p)!}{(2p+1-\sum r_q)!} \frac{(G_2^c)^{r_1}}{r_1!} \frac{(G_4^c)^{r_2}}{r_2!} \dots \frac{(G_{2q}^c)^{r_q}}{r_q!} \dots \quad (278)$$

and one finds

$$G_{2p}^c(g) = \frac{(3p-1)! 3^{1-p}}{(p-1)!(2p-1)!} \sum_{p=1}^{\infty} \frac{(-12g)^k (2k+p-1)!}{(k-p+1)!(k+2p)!} \quad (279)$$

Finally, if the one-particle irreducible functions (the vertex functions) are expressed as

$$\begin{aligned} \Gamma_2 &= [G_2^c]^{-1} \\ -\Gamma_4 &= G_4^c [G_2^c]^{-4} \\ -\Gamma_6 &= G_6^c [G_2^c]^{-6} - 3 [G_4^c]^2 [G_2^c]^{-7} \\ -\Gamma_8 &= G_8^c [G_2^c]^{-8} - 8 G_6^c G_4^c [G_2^c]^{-9} + 12 [G_4^c]^3 [G_2^c]^{-10} \\ &\dots \end{aligned} \quad (280a)$$

their generating function (analogous to the free energy expressed in terms of the magnetization after a Legendre transformation)

$$\Gamma(x) = \sum_{p=1}^{\infty} \Gamma_{2p} x^{2p} \quad (280b)$$

satisfies the algebraic equation

$$3x^2(1-a^2)(1+\Gamma)^2 + 9a^4\Gamma(1+\Gamma-\Gamma^2/x^2) - a^2(2+a^2)^2\Gamma = 0 \quad (280c)$$

from which one can derive an expansion in powers of g .

ii) It is also possible to study systematically the corrections to the planar approximation. This amounts to enumerating the diagrams which are drawn on a torus, a double torus, etc. Let the corresponding contributions be denoted $E_1(g)$, $E_2(g)$, ... In the Hermitian case ($\alpha = 2$) we have

$$E(g) - E(0) = E_0(g) + \frac{1}{N^2} E_1(g) + \frac{1}{N^4} E_2(g) + \dots \quad (281)$$

A useful technique to obtaining the successive terms relies on the asymptotic expansion for large order of the orthogonal polynomials with respect to the measure $d\lambda \exp -(\frac{1}{2}\lambda^2 + (g/N)\lambda^4)$. These polynomials generalize the Hermite polynomials used in the previous section. One is led to the expressions

$$\begin{aligned} E_1(g) &= \frac{1}{12} \ln(2 - a^2) \\ E_2(g) &= \frac{1}{6!} \frac{(1 - a^2)^3}{(2 - a^2)^5} (82 + 21a^2 - 3a^4) \\ &\dots \end{aligned} \quad (282)$$

with a^2 given by equation (270). For $E_H(g)$ the leading term in g is, for any H ,

$$E_H(g) = 2^{2H-2} \frac{(4H-3)!}{H!(H-1)!} g^{2H-1} + \dots \quad (283)$$

iii) We have discussed a quartic interaction, but the method obviously extends to any polynomial interaction in particular to a cubic one (to give a meaning to the original integral, we may assume the coupling constant to be pure imaginary). We sketch briefly the results in this case (useful by duality to study triangulations on surfaces). One wishes to evaluate for large N the integral

$$\begin{aligned} \exp -N^2 E(g) &= \int \prod_{1 \leq i \leq N} d\lambda_i \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^2 \\ &\quad \times \exp - \sum_{1 \leq i \leq N} \left(\frac{1}{2} \lambda_i^2 + \frac{g}{\sqrt{N}} \lambda_i^3 \right) \end{aligned} \quad (284)$$

We apply the saddle point method with

$$\lambda_i = \sqrt{N} \lambda \left(\frac{i}{N} \right) \quad d\lambda v(\lambda) = 2 d\lambda \quad (285)$$

This leads to the equations

$$\begin{aligned} \lambda + 3g\lambda^2 &= PP \int_{2a}^{2b} \frac{v(\mu) d\mu}{\lambda - \mu} \quad 2a \leq \lambda \leq 2b \\ \int_{2a}^{2b} d\lambda v(\lambda) &= 2 \end{aligned} \quad (286)$$

The support of $v(\lambda)$ is not symmetric except in the limit $g \rightarrow 0$. The function $F(z)$ having $v(\lambda)$ as its discontinuity along the cut $[2a, 2b]$, real outside this interval on the real axis, and vanishing

as $2/z$ at infinity, is

$$\begin{aligned} F(z) &= \int_{2a}^{2b} \frac{v(\lambda) d\lambda}{z - \lambda} \\ &= z + 3gz^2 - [3gz + 1 + 3g(a+b)] [(z-2a)(z-2b)]^{\frac{1}{2}} \end{aligned} \quad (287)$$

Set

$$\sigma = 3g(a+b) \quad (288)$$

The required behaviour of F at infinity gives the condition

$$\begin{aligned} 18g^2 + \sigma(1+\sigma)(1+2\sigma) &= 0 \\ \sigma &= -\frac{1}{4} \sum_{k=1}^{\infty} \frac{(72g^2)^k \Gamma(\frac{1}{2}(3k-1))}{k! \Gamma(\frac{1}{2}(k+1))} \end{aligned} \quad (289)$$

This yields for $E(g)$

$$\begin{aligned} E(g) - E(0) &= -\frac{\sigma(3\sigma^2 + 6\sigma + 2)}{3(1+\sigma)(1+2\sigma)^2} + \frac{1}{2} \ln(1+2\sigma) \\ &= -\frac{1}{2} \sum_{k=1}^{\infty} \frac{(72g^2)^k \Gamma(3k/2)}{(k+2)! \Gamma(k/2+1)} \end{aligned} \quad (290)$$

All these expressions are analytic around $g = 0$, the nearest singularity being

$$g_c^2 = \frac{1}{108\sqrt{3}}$$

The generating function of connected Green functions

$$\psi(j) = 1 + \sum_{p=1}^{\infty} j^p G_p^c \quad (291a)$$

obeys the equation

$$3g\psi^2 + (j-3g)\psi - \frac{1}{2}j \frac{(1+\sigma)(2+3\sigma)}{1+2\sigma} - j^3 = 0 \quad (291b)$$

One can also modify the Green functions by eliminating the tadpole contributions corresponding to $G_1 = G_1^c$. To this effect, one introduces an extra linear term $\rho \sum_{1 \leq i \leq N} \lambda_i$ in the action and one fixes the value of the parameter ρ by requiring the new one-point function G_1 to vanish. The modified generating function $\tilde{\psi}$ obeys

$$3g\tilde{\psi}^2 + (j-3g)\tilde{\psi} - j(1-\rho j + j^2) = 0 \quad (292a)$$

where ρ is related parametrically to g through

$$\begin{aligned} 3g\rho &= -\tau(1-3\tau) \\ 3g^2 &= \tau(1-2\tau)^2 \end{aligned} \quad (292b)$$

Eliminating τ , we obtain for $\tilde{\psi}$ an explicit expansion in terms of graphs with E external lines, and V vertices (the remaining topological characteristics being fixed by the conditions that we deal with connected planar graphs with vertices of valence three)

$$\begin{aligned} \tilde{\psi}(j, g) &= 1 + \sum_{E=2}^{\infty} j^E \sum_{\substack{V=E-2 \\ V-E \text{ even}}}^{\infty} (-3g)^V \tilde{\psi}_{E,V} \quad (293) \\ \tilde{\psi}_{E,V} &= \frac{2E-2}{(E-1)!(E-2)!} 2^{\frac{1}{2}(V-E)+1} \frac{(\frac{1}{2}(3V+E)-1)!}{(E+V)!(\frac{1}{2}(V-E)+1)!} \end{aligned}$$

To summarize, the method leads to algebraic equations in the pure combinatorial case (zero-dimensional field theory!), which provides the required quantities by an expansion in powers of g . Although it would be of great interest to carry out the same program in more realistic cases, this has not been fully successful, albeit in some phenomenological way, except for one-dimensional quantum field theory i.e. quantum mechanics.

10.4.2 The planar approximation in quantum mechanics

In general, the diagrams with a given topology are to be weighted by nontrivial Feynman integrals. It is therefore remarkable that, at least in the one-dimensional case, with a propagator $(p^2 + 1)^{-1}$ (p is a one-component momentum to be integrated from $-\infty$ to $+\infty$), one can obtain explicit expressions in the planar limit. As an illustration we study a set of N^2 coupled anharmonic oscillators, in such a way that the configuration variables can be arranged as the entries in an $N \times N$ hermitian matrix with the Hamiltonian invariant under the group $SU(N)$. We look at the ground state energy for

$$\begin{aligned} H &= -\frac{1}{2}\Delta + V \\ \Delta &= \sum_{1 \leq i \leq N} \frac{\partial^2}{\partial M_{ii}^2} + \frac{1}{2} \sum_{1 \leq i < j \leq N} \left[\frac{\partial^2}{(\partial \text{Re} M_{ij})^2} + \frac{\partial^2}{(\partial \text{Im} M_{ij})^2} \right] \\ V &= \frac{1}{2} \text{Tr} M^2 + \frac{g}{N} \text{Tr} M^4 \end{aligned} \quad (294)$$

This is only shown as an illustration and one could introduce more general potentials. For the ground state

$$H\psi = N^2 E(g)\psi \quad (295)$$

the corresponding wavefunction is invariant under unitary transformations

$$\psi(M) = \psi(U^\dagger M U)$$

Thus ψ is a symmetric function of the eigenvalues of M , generalizing the notion of radial variables in "polar" coordinates. Hence $\psi \equiv \psi(\{\lambda_i\})$ and

$$\begin{aligned} E(g) &= \lim_{N \rightarrow \infty} \frac{1}{N^2} \text{Inf}_{\psi} Z_0^{-1} Z \\ Z &= \int \prod_{1 \leq i \leq N} d\lambda_i \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^2 \left\{ \frac{1}{2} \sum_{1 \leq i \leq N} \left(\frac{\partial \psi}{\partial \lambda_i} \right)^2 + V \psi^2 \right\} \\ Z_0 &= \int \prod_{1 \leq i \leq N} d\lambda_i \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j)^2 \psi^2 \end{aligned} \quad (296)$$

To the symmetric wavefunction $\psi(\{\lambda_i\})$, we associate an antisymmetric one denoted ϕ , defined as

$$\phi(\lambda_1, \dots, \lambda_N) = \prod_{1 \leq i < j \leq N} (\lambda_i - \lambda_j) \psi(\lambda_1, \dots, \lambda_N) \quad (297)$$

describing a gas of noninteracting fermions submitted to a quartic potential (in the present case). From equation (296) we derive the variational Schrödinger equation

$$\sum_{1 \leq i \leq N} \left\{ -\frac{1}{2} \frac{\partial^2}{\partial \lambda_i^2} + \frac{1}{2} \lambda_i^2 + \frac{g}{N} \lambda_i^4 \right\} \phi = N^2 E(g) \phi \quad (298)$$

The factor $1/N$ scaling the potential is included to ensure a correct limiting behaviour. Going over to polar coordinates did not generate additional terms in the potential. This would not be the case if we would have studied another ensemble of matrices. The antisymmetric behaviour of ϕ is the analog of the boundary conditions at the origin in cartesian polar coordinates. Call $e_1 < e_2 < e_3 < \dots$ the successive energies corresponding to the one-body Hamiltonian

$$h = -\frac{1}{2} \frac{\partial^2}{\partial \lambda^2} + \frac{1}{2} \lambda^2 + \frac{g}{N} \lambda^4 \quad (299)$$

new subtractions. More precisely,

$$\Gamma_{2,\varphi^4}(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q})_\Lambda = a_1(\Lambda) + a_2(\Lambda)(\mathbf{p}_1^2 + \mathbf{p}_2^2) + a_3(\Lambda)\mathbf{p}_1 \cdot \mathbf{p}_2 + \text{finite} \quad (254)$$

The coefficient a_1 is quadratic, a_2 and a_3 are logarithmic in Λ . As Γ_{2,φ^2} also has a logarithmic divergence, it is possible to construct a linear combination in φ^2 and φ^4 which eliminates the singular part of a_1 . Similarly, using the operators $\varphi(\Delta\varphi)$ and $(\partial\varphi)^2$, the divergences of a_2 and a_3 can also be suppressed. The idea is thus to form linear combinations of operators with equal or lower dimensions in order to obtain finite renormalized quantities. We have also to give precise normalization conditions. For the example of operators of degree 0, the three possible candidates are

$$\mathcal{O}_1 = \varphi^4 \quad \mathcal{O}_2 = \varphi(\Delta\varphi) \quad \mathcal{O}_3 = (\partial\varphi)^2$$

Their insertion leads to renormalized vertex functions related to the bare ones through

$$\Gamma_{n,\mathcal{O}_i}^R = Z_1^{\frac{1}{2}n} \left[\sum_j z_{ij} \Gamma_{n,\mathcal{O}_j} + z_i \Gamma_{n,\frac{1}{2}\varphi^2} \right] \quad (255)$$

in which twelve new *a priori* divergent constants z_{ij} and z_i have been introduced. Hence twelve conditions are necessary, as e.g.

$$\begin{aligned} \Gamma_{4,\varphi^4}^R(0) &= 4! & \Gamma_{2,\varphi^4}^R(\mathbf{p}) &= \mathcal{O}(\mathbf{p}^4) \\ \Gamma_{4,\varphi(\Delta\varphi)}^R(0) &= 0 & \Gamma_{2,\varphi(\Delta\varphi)}^R(\mathbf{p}) &= \mathbf{p}_1^2 + \mathbf{p}_2^2 + \mathcal{O}(\mathbf{p}^4) \\ \Gamma_{4,(\partial\varphi)^2}^R(0) &= 0 & \Gamma_{2,(\partial\varphi)^2}^R(\mathbf{p}) &= 2\mathbf{p}_1 \cdot \mathbf{p}_2 + \mathcal{O}(\mathbf{p}^4) \end{aligned} \quad (256)$$

Note that each function $\Gamma_{2,\mathcal{O}_i}^R$ satisfies three conditions at $\mathbf{p} = 0$.

This example can now be generalized. The insertion of a renormalized operator is equivalent, in the bare theory, to the insertion of a linear combination of all operators of equal or lower dimension. Their number might be reduced if one takes into account symmetries or relations arising from the equations of motion. Indeed, if the regularized theory admits some symmetry, it is clear that we can only combine quantities with the same behaviour. Note also that the Green functions obey relations which result from the structure of the Lagrangian (as a consequence of the so-called *equations of motion*).

To first order, the introduction in the action of additional local terms requires that one estimates the corresponding insertions at

zero momentum in the correlation functions. The renormalized operator $\mathcal{O}_i$ is coupled to other bare $\mathcal{O}_j$'s and in particular to $\frac{1}{2}\varphi^2$ which is of degree -2. Therefore, its insertion displaces the critical temperature. This effect can be compensated by subtracting from $\mathcal{O}_i$ a term proportional to $\frac{1}{2}\varphi^2$,

$$\hat{\mathcal{O}}_i = \mathcal{O}_i - \frac{1}{2}a_i\varphi^2, \quad (257)$$

in which a_i is to be chosen in such a way that the expansion of the renormalized operator does not contain any contribution of $\frac{1}{2}\varphi^2$.

A definite advantage of this procedure is to suppress additional infrared divergences which would correspond to a perturbative shift of the critical temperature. Furthermore, it is convenient to include in $\mathcal{O}_i$ and $\hat{\mathcal{O}}_i$ the dimensionful factor $\Lambda^{-\delta_i}$ which ensures that the corresponding coupling is dimensionless. As we have seen, this modification also affects the coupling constant g_0 (coupled to φ^4) for $d < 4$.

Using these conventions, the zero momentum insertions in the critical theory read

$$\Gamma_{n,\hat{\mathcal{O}}_i}^R(\mathbf{p}; g_0, \mu) = Z_1^{n/2}(g_0, \Lambda/\mu) \sum_j (\Lambda/\mu)^{\delta_j} z_{ij}(g_0, \Lambda/\mu) \Gamma_{n,\mathcal{O}_j}(\mathbf{p}; g_0, \Lambda) \quad (258)$$

where the sum on j runs over all operators of degree less than or equal to δ_i except $\frac{1}{2}\varphi^2$. In dimension four, the quantities z_{ij} only contain logarithmic divergences.

The corresponding Callan-Symanzik equations for the bare vertex functions are

$$\begin{aligned} \left\{ \left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g_0) \frac{\partial}{\partial g_0} - \frac{1}{2}n\gamma_1(g_0) \right] \delta_{ij} - \gamma_{ij}(g_0) \right\} \Gamma_{n,\mathcal{O}_i}(\mathbf{p}; g_0, \Lambda) &= 0 \\ \gamma_{ij}(g_0) &= \{z^{-1}\}_{ik} \left[\delta_k + \Lambda \frac{\partial}{\partial \Lambda} \right] z_{kj} \end{aligned} \quad (259)$$

For small momenta, the integration of this equation generates additional powers of $\mathbf{p}/\Lambda$ with respect to functions without insertions.

The matrix γ_{ij} has a block triangular structure. Let η_a be its eigenvalues. One finds as a result of equation (259) that

$$\Gamma_{n,\mathcal{O}_i}(\lambda\mathbf{p}; g_0, \Lambda) \underset{\lambda \rightarrow 0}{\sim} \lambda^{d - \frac{1}{2}n(d-2+\eta)} \sum_a C_{ia} \lambda^{\eta_a} \quad (260)$$

According to equation (252), the eigenvalues behave as

$$\eta_a = \delta_a + \mathcal{O}(4-d) \quad (261)$$

The block triangular structure of z_{ij} shows that, when the degree δ_i of the operator increases, one obtains for its eigenvalues the previous values η_a corresponding to lower degrees together with new eigenvalues, close to the degree of the inserted operator. As we have compensated $\frac{1}{2}\varphi^2$, the lowest eigenvalues, which are the dominant ones in (260), correspond to the operators φ^4 and $(\partial\varphi)^2$ (note that, at zero momentum, the two operators $\varphi(-\Delta\varphi)$ and $(\partial\varphi)^2$ are equivalent). These operators are marginal in four dimensions. Their effect is to change the coupling constant and to renormalize the field. The corresponding eigenvalue is precisely ω (see (231)) and the correction behaves as $(|p|/\Lambda)^\omega$.

To summarize, the effect of higher local terms in the Lagrangian is to correct the scaling laws by subdominant corrections. The leading one corresponds to the exponent ω which can be directly computed within the framework of the φ^4 theory. Hence, the truncation of the Lagrangian is *a posteriori* justified.

5.5 Numerical results

It is not possible to collect here the many perturbative results about critical quantities, nor their detailed derivation. The interested reader can find a partial list of references at the end of this chapter. We only reproduce here some series. Their numerical treatment is essential from the practical point of view. Important progress was made when it was realized that the series, although divergent, are nevertheless asymptotic, and that their behaviour at large order can be analyzed using saddle point methods. The initial idea is due to Dyson in the context of electrodynamics. The large order behaviour of perturbative series in field theory has been studied extensively by Lipatov and others. This information allows one to make the best use of rather short series. The stability of the estimates (i.e. the theoretical uncertainties) are discussed in the original papers.

We first give some results on the ε -expansion. Then we also present some series in fixed dimension three.

5.5.1 ε -expansion of critical exponents

The computation scheme described in section 3 has been developed and applied by Wilson, Brézin, Le Guillou, Zinn-Justin, Nickel and others to compute the following series

$$\begin{aligned} \eta = \gamma_1(g_c) = & \frac{n+2}{2(n+8)^2}\varepsilon^2 + \frac{n+2}{8(n+8)^4}[24(3n+14) - (n+8)^2]\varepsilon^3 \\ & + \frac{n+2}{2(n+8)^6}\left[-\frac{5}{16}n^4 - \frac{115}{8}n^3 + \frac{281}{4}n^2 + 1120n \right. \\ & \left. + 2884 - 24(n+8)(5n+22)\zeta(3)\right]\varepsilon^4 + \mathcal{O}(\varepsilon^5) \end{aligned} \quad (262)$$

$$\begin{aligned} 2 - \nu^{-1} = \gamma_2(g_c) = & \frac{n+2}{n+8}\varepsilon + \frac{n+2}{2(n+8)^3}(13n+44)\varepsilon^2 \\ & + \frac{n+2}{(n+8)^4}\left[36\frac{(3n+14)(n+3)}{n+8} \right. \\ & \left. - \frac{3n^2 + 388n + 848}{8} - 12(5n+22)\zeta(3)\right]\varepsilon^3 + \mathcal{O}(\varepsilon^4) \end{aligned} \quad (263)$$

with

$$\zeta(3) = \sum_{j=1}^{\infty} \frac{1}{j^3} = 1.20206... \quad (264)$$

From the relation $\gamma = \nu(2 - \eta)$, one gets

$$\begin{aligned} \gamma = 1 + & \frac{n+2}{2(n+8)}\varepsilon + \frac{(n+2)(n^2+22n+52)}{4(n+8)^3}\varepsilon^2 \\ & + \frac{n+2}{(n+8)^5}\left[\frac{n^4}{8} + \frac{11n^3}{3} + 83n^2 + 312n \right. \\ & \left. + 388 - 6(n+8)(5n+22)\zeta(3)\right]\varepsilon^3 + \mathcal{O}(\varepsilon^4) \end{aligned} \quad (265)$$

The exponent ω governing the corrections to the dominant scaling behaviour, defined by equation (231), is given by

$$\begin{aligned} \omega = \beta'(g_c) = & \varepsilon - \frac{3(3n+14)}{(n+8)^2}\varepsilon^2 + \frac{1}{(n+8)^3}\left[\frac{33n^2}{4} + \frac{461n}{2} \right. \\ & \left. + 740 + 24(5n+22)\zeta(3) - 18\frac{(3n+14)^2}{n+8}\right]\varepsilon^3 + \mathcal{O}(\varepsilon^4) \end{aligned} \quad (266)$$

Table I. Critical exponents of the three-dimensional Ising model

	$\mathcal{O}(\varepsilon^k)$	Padé-Borel	High temp. series	$d = 3$ renorm. group
η	0.019 0.037 0.029	0.04 ^a 0.0333 ± 0.0001 ^b	0.041 ± 0.01	0.031 ± 0.001
γ	1.167 1.244 1.19	1.242 ^a 1.235 ± 0.004 ^b	1.250 ± 0.003 ^c 1.245 ± 0.003 ^d	1.241 ± 0.002
ν	0.589 0.634 0.606	0.632 ^a 0.628 ± 0.02 ^b	0.638 ± 0.002 ^c	0.630 ± 0.0015

Note: The values listed $\mathcal{O}(\varepsilon^k)$ are for $k = 1, 2, 3$ and should be understood $\mathcal{O}(\varepsilon^{k+1})$ for the exponent η .

^aJ. Zinn-Justin and J.-C. Le Guillou, *Phys. Rev. B* 21 3976 (1980).

^bG. Parisi, *J. Stat. Phys.* 23, 49 (1980).

^cC. Domb, *Phase Transitions and Critical Phenomena*, vol. 3, C. Domb and M.S. Green eds (Academic Press, London).

^dJ. Zinn-Justin, *J. Physique* 40 969 (1979).

If the corrections are expressed in terms of the deviation θ to the critical temperature, the corresponding exponent is (cf. (232))

$$\omega\nu = \frac{1}{2}\varepsilon + \frac{n^2 - 8n - 68}{4(n+8)^2}\varepsilon^2 + \frac{1}{8(n+8)^3}[(n+2)(n+3)(n+20) - 72\frac{(3n+14)^2}{n+8} + 15n^2 + 802n + 2792 + 96(5n+22)\zeta(3)]\varepsilon^3 + \mathcal{O}(\varepsilon^4) \quad (267)$$

The most naive interpretation is to insert the corresponding value of ε and to evaluate the successive contributions. Table I displays the successive approximations for the Ising model $n = 1$ in three dimensions. At second order, the values are in good agreement with the results from high temperature expansions, but the disagreement increases at the next order. The fourth column displays more accurate values based on a Padé-Borel resummation method, which uses large order estimates. Finally, the last two columns show the results of high temperature expansions and similar results of the renormalization group analysis obtained in dimension three.

5.5.2 Equation of state

The equation of state can also be obtained in an ε -expansion. Let us compute the effective potential $V(M)$ for an n -vector model as the generating functional of irreducible diagrams. Up to one-loop order,

$$V(M) = \frac{1}{2}\theta M^2 + \frac{g(M^2)^2}{4!} + \frac{1}{2} \int \frac{d^d p}{(2\pi)^d} \text{Tr} \ln[(p^2 + \theta)\delta_{ij} + \frac{1}{6}g(M^2)\delta_{ij} + 2M_i M_j] - \text{subtractions} \quad (268)$$

The trace is relative to internal indices. Subtractions contain an M -independent polynomial $P_2(\theta)$ of degree 2 in θ , which corresponds to the divergences of $\Gamma_{0,0}$, $\Gamma_{0,1}$, $\Gamma_{0,2}$. Further subtractions correspond to the renormalization of $\Gamma_{2,0}$, $\Gamma_{4,0}$, $\Gamma_{2,1}$ in the massless theory. In particular, the functions $\Gamma_{4,0}$ and $\Gamma_{2,1}$ are normalized at the symmetric point $p_i \cdot p_j = -p_i^2/3$ and we choose $\mu^2 = 4p_i^2/3$ as a mass scale. Hence

$$V(M) = \frac{1}{2}\theta M^2 + g \frac{(M^2)^2}{4!} + \frac{1}{2} \int \frac{d^d p}{(2\pi)^d} \text{Tr} \left\{ \ln \left[\delta_{ij} + \frac{1}{p^2}(\theta + \frac{1}{6}gM^2)\delta_{ij} + \frac{g}{3p^2}M_i M_j \right] - \frac{1}{p^2}(\theta + \frac{1}{6}gM^2)\delta_{ij} - \frac{g}{p^2}M_i M_j + \frac{1}{2p^2(p+p_i+p_j)^2} \left[(\theta + \frac{1}{6}gM^2)\delta_{ij} + \frac{2}{3}g(\theta + \frac{1}{3}gM^2)M_i M_j \right] \right\} + P_2(\theta) \quad (269)$$

The transverse and longitudinal parts are treated independently. Taking into account the quartic subtractions which avoid infrared divergences, one gets as $d \rightarrow 4$

$$V = \frac{1}{2}\theta M^2 + g \frac{(M^2)^2}{4!} + \frac{1}{8(4\pi)^2} \{ (\theta + \frac{1}{2}gM^2)^2 [2\ln(\theta + \frac{1}{2}gM^2) + 1] + (n-1)(\theta + \frac{1}{6}gM^2)[2\ln(\theta + \frac{1}{6}gM^2) + 1] \} + P_2(\theta) \quad (270)$$

We substitute for g its critical value (118),

$$g_c = (4\pi)^2 \frac{3}{n+8} \varepsilon + \dots \quad (271)$$

and get

$$H = \frac{\partial V}{\partial M} = M \left\{ \theta + \frac{1}{6}y + \frac{3\varepsilon}{2(n+8)} \left[(\theta + \frac{1}{2}y) \ln(\theta + \frac{1}{2}y) + 1 + \frac{1}{3}(n-1)(\theta + \frac{1}{6}y)(\ln(\theta + \frac{1}{6}y) + 1) \right] + \dots \right\} \quad (272)$$

with $y = g_c M^2$. This expression can be identified with the expansion of

$$H = M^\delta f\left(\frac{\theta}{M^{1/\beta}}\right) \quad (273)$$

To order ε ,

$$\begin{aligned} \eta &= 0 + \dots \\ \nu &= \frac{1}{2} + \frac{\varepsilon}{4} \frac{n+2}{n+8} + \dots \\ \beta &= \frac{1}{2} \nu (d-2+\eta) = \frac{1}{2} - \frac{3\varepsilon}{2(n+8)} \\ \delta &= \frac{d+2-\eta}{d-2+\eta} = 3 + \varepsilon + \dots \end{aligned} \quad (274)$$

If we insert these values in (273) and compare with (272), changing the scale of M and θ so that $f(0) = 1$ and $f(-1) = 0$, we get

$$H = M^\delta \left(1 + x + \frac{\varepsilon}{2(n+8)} [(n-1)(x+1) \ln(x+1) + 3(x+3) \ln(x+3) + 6x \ln 2 - 9(x+1) \ln 3] + \dots \right) \quad (275)$$

with

$$x = \frac{\theta}{M^{1/\beta}}$$

which gives the function $f(x)$ to order ε . Further computations at higher orders can be found in the references.

5.5.3 Amplitude ratios

It is also possible to compute other universal quantities describing the critical behaviour which are independent of the arbitrary

scale of the temperature and the external field. For instance, consider the specific heat $A_\pm (\pm\theta)^{-\alpha}$ above and below the critical temperature. The ratio of amplitudes A_+/A_- is such a universal quantity. This ratio is also the one of singular parts in the free energy in the absence of an external field. These quantities are contained in the expression of the function $f(x)$ in the equation of state (275).

Taking into account that $2-\alpha = \beta(1+\delta)$, the effective potential is of the form

$$V(M, \theta) = \theta^{2-\alpha} \varphi(x) \quad x = \frac{\theta}{M^{1/\beta}} \quad (276)$$

The free energy is a regular function of θ in presence of a field, i.e. $x^{2-\alpha} \varphi(x)$ is analytic at the origin. As

$$\varphi'(x) = -x^{\alpha-3} f(x) \quad (277)$$

with the same function f as occurred in equations (273) and (275), an integration is sufficient to determine φ . We take a convention of a phase $e^{i\pi}$ for x negative, and determine A_+/A_- as the ratio of the limits of V as $M \rightarrow 0$ for $\theta > 0$ and as $M \rightarrow M_{sp} = (-\theta)^\beta$ for $\theta < 0$. Hence

$$\frac{A_+}{A_-} = \frac{\varphi(+\infty)}{e^{i\pi(2-\alpha)} \varphi(-1)} \quad (278)$$

Integrating equation (277) assuming $0 < \alpha < 1$, one obtains

$$\varphi(x) = \frac{\beta}{\alpha-2} \left\{ \int_0^x dy y^{\alpha-2} [f'(y) - f'(0)] + f'(0) \frac{x^{\alpha-1}}{\alpha-1} - x^{\alpha-2} f(x) \right\} \quad 0 < \alpha < 1 \quad (279)$$

For an alternative derivation, we may return to equation (269) and substitute g_c as well as the corresponding scales for M and θ . In the range $0 < \alpha < 1$, we obtain

$$\frac{A_+}{A_-} = \frac{\int_0^\infty dx x^{\alpha-2} [f'(x) - f'(0)]}{\int_{-1}^0 dx (-x)^{\alpha-2} [f'(0) - f'(x)] + f'(0)/(1-\alpha)} \quad 0 < \alpha < 1 \quad (280)$$

Similar expressions exist if α lies in other intervals.

Using the result (275) as well as $\alpha = \frac{1}{2}\varepsilon(4-n)/(n+8) + \dots$, one finds that the leading term in the ratio A_+/A_- is

$$\frac{A_+}{A_-} = \frac{1}{4}n + \mathcal{O}(\varepsilon) \quad (281)$$

An elaborate computation up to order ε^2 yields

$$\frac{A_+}{A_-} = \frac{1}{4}n2^\alpha \left[1 + \varepsilon + \frac{\varepsilon^2}{2(n+8)^2} [3n^2 + 26n + 100 + \zeta(2)(4-n)(n-1) - 6\zeta(3)(5n+22) + \frac{9}{2}(4-n)I] + \mathcal{O}(\varepsilon^3) \right] \quad (282)$$

with

$$I = \int_0^1 dx \frac{\ln x(1-x)}{1-x(1-x)} = -2.349\dots$$

In the mean field approximation, the specific heat is discontinuous at the critical temperature and the above ratio vanishes. As $\varepsilon \rightarrow 0$, that is in four dimensions, the ratio is relative to the coefficient of the singularity in $(\ln|\theta|)^{(4-n)/(n+8)}$ and its value is $\frac{1}{4}n$. In three dimensions, taking $\nu = 0.63$ and therefore $\alpha = 0.11$ for $n = 1$, equation (282) predicts the value

$$d=3 \quad n=1 \quad \frac{A_+}{A_-} = 0.54 \quad (283)$$

For the Ising model, the susceptibility χ and the correlation length are finite below the critical temperature, since there are no zero modes. It is then possible to define a similar ratio χ_+/χ_- for the susceptibility, written as

$$\chi(\theta) = \chi_\pm |\theta|^{-\gamma} \quad (284)$$

From the equation of state, one derives

$$\frac{\chi_+}{\chi_-} = \lim_{x \rightarrow \infty} f'(-1) \frac{x^\gamma}{\beta \delta f(x) - x f'(x)} \quad (285)$$

which is meaningful only if $f(x)/x^\gamma$ behaves as a power series in $x^{-2\beta} = M^2/\theta^{2\beta}$ for large x . Using

$$\begin{aligned} \beta \delta &= \frac{3}{2} + \dots \\ \gamma &= 1 + \frac{1}{6}\varepsilon + \dots \end{aligned} \quad (286)$$

and the expression (275) for $f(x)$, one obtains

$$\frac{\chi_+}{\chi_-} = 2 + \varepsilon \left(1 + \frac{1}{3} \ln 2 \right) + \mathcal{O}(\varepsilon^2) \quad (287)$$

which can be rewritten as

$$\frac{\chi_+}{\chi_-} = 2^{\gamma-1} \frac{\gamma}{\beta} + \mathcal{O}(\varepsilon^2) \quad (288)$$

The mean field value 2 is obtained in the limit $\varepsilon \rightarrow 0$.

5.5.4 Three-dimensional results

The direct treatment of the critical theory in fixed dimension less than four yields more accurate results, although no small parameter analogous to ε is available. We have discussed above the computational scheme. Let us display the results obtained by Baker, Nickel, Green and Meiron for $d = 3$ and $n = 1$. They use a coupling constant normalized such that the first two terms of the β function have coefficients -1 and 1 . This means

$$\bar{g} = \frac{3g}{16\pi} \quad \bar{\beta}(\bar{g}) = \frac{3}{16\pi} \beta(g)$$

The series are

$$\begin{aligned} \bar{\beta}(\bar{g}) &= -\bar{g} + \bar{g}^2 - 0.4224965707\bar{g}^3 + 0.3510695978\bar{g}^4 \\ &\quad - 0.3765268283\bar{g}^5 + 0.49554751\bar{g}^6 - 0.749689\bar{g}^7 + \dots \end{aligned} \quad (289)$$

$$\begin{aligned} \gamma_1(\bar{g}) &= 0.0109739369\bar{g}^2 + 0.0009142223\bar{g}^3 \\ &\quad + 0.0017962229\bar{g}^4 - 0.00065370\bar{g}^5 + 0.0013878\bar{g}^6 + \dots \end{aligned} \quad (290)$$

$$\begin{aligned} \gamma_2(\bar{g}) &= \frac{1}{3}\bar{g} - 0.0631001372\bar{g}^2 + 0.0452244754\bar{g}^3 \\ &\quad - 0.0377233460\bar{g}^4 + 0.04374666\bar{g}^5 - 0.0589754\bar{g}^6 + \dots \end{aligned} \quad (291)$$

This extends considerably the meager expressions given in (115)–(118). We do not elaborate here on the methods needed to extract the information contained in these series, in particular to find $\bar{g}_c$, and which also make an essential use of the asymptotic behaviour. Table II displays the results obtained for the critical exponents.

Appendix 5.A Multicritical points

We have presented a detailed study of renormalization applied to the standard φ^4 critical point. It was observed that irrelevant

Table II. Critical exponents of n -vector models in dimension three

	$n = 0$	$n = 1$	$n = 2$	$n = 3$
$\bar{g}_c$	1.421 ± 0.004	1.416 ± 0.0015	1.406 ± 0.005	1.392 ± 0.009
ω	0.794 ± 0.006	0.788 ± 0.003	0.78 ± 0.01	0.78 ± 0.02
$\eta + \nu^{-1} - 2$	-0.274 ± 0.01	-0.382 ± 0.005	-0.474 ± 0.008	-0.550 ± 0.012
ν	0.588 ± 0.001	0.630 ± 0.002	0.669 ± 0.003	0.705 ± 0.005
γ	1.161 ± 0.003	1.241 ± 0.004	1.316 ± 0.009	1.39 ± 0.01
α	0.236 ± 0.004	0.110 ± 0.008	-0.007 ± 0.009	-0.115 ± 0.015
β	0.302 ± 0.004	0.324 ± 0.006	0.346 ± 0.009	0.362 ± 0.012
$\frac{1}{2}(3\nu + \gamma)$	1.462 ± 0.004	1.566 ± 0.006	1.662 ± 0.009	1.753 ± 0.012
η	0.026 ± 0.014	0.031 ± 0.011	0.032 ± 0.015	0.031 ± 0.022
δ	4.85 ± 0.08	4.82 ± 0.06	4.81 ± 0.08	4.82 ± 0.12

Note: From Baker, Nickel, Green and Meiron, *Phys. Rev. Lett.* **38**, 1351 (1976), *Phys. Rev.* **B17**, 1365 (1978).

operators did not affect the leading scaling behaviour, giving a foundation to the property of universality. This holds close to the upper marginal dimension for two effective macroscopic control parameters (such as the external field and the temperature). There are circumstances when this is no longer the case, as exemplified by the tricritical singularity, occurring for instance in $\text{He}^3 - \text{He}^4$ mixtures or for antiferromagnets in the presence of a strong external field (metamagnets). Here one deals with a three-dimensional phase diagram. We give as a complement a brief introduction to this subject, present some sample calculations, and quote in the notes relevant work giving both a more complete physical picture as well as more details on renormalization.

For simplicity assume a symmetric situation with a one component field φ (i.e. no explicit symmetry breaking term is present). To discuss a tricritical singularity we retain in a mean field approach terms up to order φ^6 in a Landau Lagrangian. Neglecting fluctuations, the free energy per unit volume is expressed as a potential term

$$V(\varphi) = \frac{1}{2}t\varphi^2 + \frac{1}{4}s\varphi^4 + \frac{1}{6}g\varphi^6 \quad (\text{A.1})$$

One requires g to be positive to insure that $V(\varphi)$ is bounded from below. In the $s-t$ plane, let us plot the various possible situations following the minima of $V(\varphi)$ which should satisfy

$$\varphi \{t + s\varphi^2 + g\varphi^4\} = 0 \quad (\text{A.2})$$

The origin is always an extremum but there exist other possibilities if $s^2 - 4tg > 0$ (dotted curve on figure 7, which will be the boundary of a metastable zone).

If t is negative, the absolute minimum is for

$$\varphi_+^2 = \frac{-s + \sqrt{s^2 - 4tg}}{2g}$$

in which case

$$V(\varphi_+) - V(0) = \frac{1}{3}\varphi_+^2 [t + s\frac{1}{4}\varphi_+^2] = \frac{1}{12}\varphi_+^2 [3t - g\varphi_+^4]$$

is negative, indicating a broken symmetry with a spontaneous magnetization φ_+ .

If both s and t are positive, the only real root of (A.2) is $\varphi = 0$ and, as one crosses the positive s axis, one has a typical second order mean field transition controlled by t and the term $g\varphi^6$ is

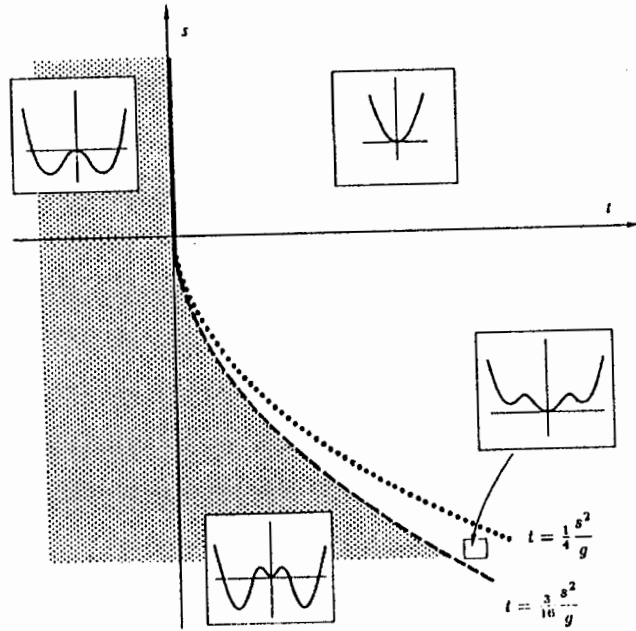


Fig. 7 Phase diagram for a tricritical point. Heavy line: second order transition; broken line: first order transition; dotted curve: limit of a metastability region. At the tricritical point, the two transition lines merge smoothly. In each region a typical behaviour of the potential is shown. The region to the left of the transition lines is the ordered one.

irrelevant. If however t is positive but s negative and $4tg < s^2$, one has two extra extrema $\varphi_{\pm}$. The sign of $V(\varphi_{\pm}) - V(0)$ is the sign of

$$8gt + s(-s \pm \sqrt{s^2 - 4gt}) = \pm s\sqrt{s^2 - 4gt} - (s^2 - 8gt)$$

Obviously $V(\varphi_-) - V(0)$ is positive but $V(\varphi_+) - V(0)$ is negative i.e. φ_+ represents a spontaneous magnetization only up to the curve $s\sqrt{s^2 - 4gt} = s^2 - 8gt$, i.e. the arc of parabola

$$s^2 = \frac{16}{3}gt \quad (A.3)$$

its negative s -branch describing a line of first order transitions.

Thus, as shown on figure 7, a line of second order transitions merges smoothly at the tricritical point ($s = t = 0$) into a line of first order transitions. The discontinuous jump in magnetization

along the first order line vanishes linearly as $\varphi_+ = \frac{3}{4}(-s)/g$ if one uses s (rather than t) as a parameter along the curve. Mean field exponents are readily defined in the vicinity of the tricritical point.

Such a mechanism can be generalized to multicritical points of order k with a multidimensional phase space ($(k-1)$ dimensions in the absence of a symmetry breaking field). Physical examples are not so readily available but this extension is interesting in view of the two-dimensional discussion to be given in chapter 9 (volume 2). At multicriticality the effective potential reduces to

$$V_{\text{crit}} = g \frac{\varphi^{2k}}{(2k)!} \quad (A.4)$$

where for convenience the coupling constant has been rescaled by a numerical factor $[(2k-1)!]^{-1}$. The upper critical dimension, such that g becomes dimensionless, is according to the familiar criterion

$$d_c = \frac{2k}{k-1} \quad (A.5)$$

Thus $d_c = 4$ for $k = 2$, $d_c = 3$ for $k = 3$, the two important physical instances, while d_c tends to 2 as k grows indefinitely.

Let us sketch the renormalization of the corresponding massless theory in an ϵ -expansion such that $d = d_c - \epsilon$. We have in mind logarithmic corrections for $d = k = 3$ (this is left as an exercise for the reader) or the approach to two dimensions in an ϵ expansion for $k \rightarrow \infty$, the rationale being that exact results are known in this case. It will be seen that this comparison is unfortunately not convincing, although it seems to offer an interesting testing ground for the concept of dimensional continuation.

The cutoff-dependent bare Lagrangian is written in terms of a renormalized field φ

$$\varphi^{\text{bare}} = Z^{1/2} \varphi \quad (A.6)$$

and renormalized coupling g (the bare one is denoted g_0), as

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}(\partial\varphi)^2 + \frac{g}{(2k)!} \mu^{\epsilon(k-1)} \varphi^{2k} + (Z-1) \frac{1}{2} \partial\varphi^2 \\ & + \frac{g}{(2k)!} \mu^{\epsilon(k-1)} (Z^k Z_g - 1) \varphi^{2k} \\ & + [\text{counterterms for } \varphi^2, \dots, \varphi^{2k-2}] \end{aligned} \quad (A.7)$$

The extra counterterms are designed to cancel the divergent parts in the irreducible functions $\Gamma^{(2)}, \dots, \Gamma^{(2k-2)}$ normalized to zero at zero momentum in Fourier transformed space. The logarithmic divergent function $\Gamma^{(2k)}$ (at $\varepsilon = 0$) is normalized at a symmetric point scaled by μ , thus defining $Z^k Z_g$. More precisely the normalization conditions read

$$\begin{aligned}\Gamma^{(2)}(0) &= \dots = \Gamma^{(2k-2)}(0) = 0 \\ \frac{\partial}{\partial p^2} \Gamma^{(2)}(p, -p) \Big|_{p^2 = \mu^2} &= 1 \\ \Gamma^{(2k)}(p_1, \dots, p_{2k})_{S_\mu} &= g\mu^{\varepsilon(k-1)}\end{aligned}\quad (A.8)$$

The symmetric point S_μ for $2k$ momenta satisfying $\sum_1^{2k} p_j = 0$ is chosen in such a way that $p_j^2 = \mu^2(2k-1)/k^2$ and for $i \neq j$ $p_i \cdot p_j = -\mu^2/k^2$, with the result that the sum of any k distinct momenta has a square equal to μ^2 . The factor $\mu^{\varepsilon(k-1)}$ insures that the renormalized coupling constant g is dimensionless. The bare functions

$$\Gamma_{\text{bare}}^{(2\ell)} = Z^{-\ell} \Gamma^{(2\ell)} \quad (A.9)$$

are of course μ -independent.

The irreducible functions $\Gamma^{(2)}$ and $\Gamma^{(2k)}$ are the only ones required to obtain the renormalization factors Z and Z_g . To lowest order, as the reader will check, the only logarithmic singularities (at $\varepsilon = 0$) are accounted for by the graphs depicted in figure 8 (dimensional regularization suppresses in particular tadpole insertions). Explicitly

$$\begin{aligned}\Gamma^{(2)}(p, -p) &= Zp^2 - \\ &\quad - \frac{g^2}{(2k-1)!} \mu^{2\varepsilon(k-1)} \int \prod_1^{2k-2} \frac{d^d q_j}{(2\pi)^d} \frac{1}{\prod_j q_j^2 (p - \sum_j q_j)^2}\end{aligned}\quad (A.10)$$

$$\begin{aligned}\Gamma^{(2k)}(\{p\}) &= g\mu^{\varepsilon(k-1)} Z_g Z^k - \\ &\quad - \frac{g^2}{k!} \mu^{2\varepsilon(k-1)} \sum \int \prod_1^{k-1} \frac{d^d q_j}{(2\pi)^d} \frac{1}{\prod_j q_j^2 (\sum \tilde{p} - \sum q_j)^2}\end{aligned}\quad (A.11)$$

On the right-hand side in equation (A.11) the sum runs over the $(2k)!/2(k!)^2$ ways of splitting the $2k$ external momenta into two groups of k momenta as $\{p\} = \{\tilde{p}\} \cup \{\bar{\tilde{p}}\}$.

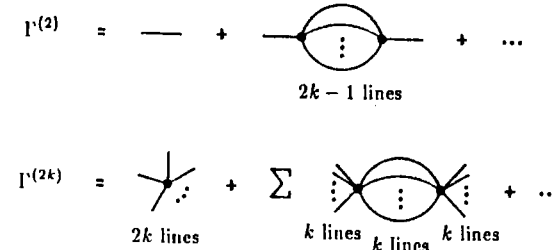


Fig. 8 Diagrams contributing to the leading singular behaviour in the renormalization factors Z and Z_g .

We use dimensional regularization in the form

$$\begin{aligned}\int \frac{d^d q}{(2\pi)^d} \frac{1}{[q^2]^\alpha} \frac{1}{[(p-q)^2]^\beta} &= \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(\frac{1}{2}d - \alpha)\Gamma(\frac{1}{2}d - \beta)}{\Gamma(d - \alpha - \beta)} \\ &\quad \times \frac{\Gamma(\alpha + \beta - \frac{1}{2}d)}{\Gamma(\alpha)\Gamma(\beta)} \frac{1}{(p^2)^{\alpha + \beta - d/2}}\end{aligned}\quad (A.12)$$

and recursively

$$\begin{aligned}\int \prod_1^{m-1} \frac{d^d q_j}{(2\pi)^d} \frac{1}{\prod_j [q_j^2]^{\alpha_j} [(p - \sum q_j)]^{\alpha_m}} &= \\ &= \frac{1}{(4\pi)^{\frac{1}{2}(m-1)d}} \frac{\Gamma(\frac{1}{2}d - \alpha_1) \dots \Gamma(\frac{1}{2}d - \alpha_m)}{\Gamma(m\frac{d}{2} - \alpha_1 - \dots - \alpha_m)} \\ &\quad \times \frac{\Gamma(\alpha_1 + \dots + \alpha_m - \frac{1}{2}(m-1)d)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_m)} \frac{1}{(p^2)^{\alpha_1 + \dots + \alpha_m - \frac{1}{2}(m-1)d}}\end{aligned}\quad (A.13)$$

In equation (A.10) we need the case $m = 2k-1$, in (A.11) the case $m = k$ with the α 's all equal to unity. Setting the derivative $\partial\Gamma^{(2)}/\partial p^2$ at $p^2 = \mu^2$ equal to unity yields in the limit $\varepsilon \rightarrow 0$

$$Z = 1 - \frac{g^2}{(4\pi)^{2k}} \frac{1}{(2k-1)!} \frac{\Gamma(1/(k-1))^{2k-1}}{\Gamma((2k-1)/(k-1))} \frac{1}{(k-1)\varepsilon} + O(g^3)\quad (A.14)$$

Similarly from the normalization (A.8) of the $2k$ -point function

$$Z_g = 1 + \frac{g}{(4\pi)^k} \frac{(2k)!}{(k!)^3} \frac{\Gamma(1/(k-1))^k}{\Gamma(k/(k-1))} \frac{1}{(k-1)\varepsilon} + O(g^2)\quad (A.15)$$

The bare functions expressed in terms of the cutoff and the bare coupling $g_0 = g\mu^{\epsilon(k-1)}Z_g$ are independent of μ . Translating the equations

$$\mu \frac{\partial}{\partial \mu} \Gamma_{\text{bare}}^{(2k)} = 0 \quad (\text{A.16})$$

in terms of the renormalized functions according to equation (A.9), and defining

$$\begin{aligned} \beta(g) &= \mu \frac{\partial}{\partial \mu} g \Big|_{g_0, \Lambda} \\ \gamma(g) &= \mu \frac{\partial}{\partial \mu} \ln Z \Big|_{g_0, \Lambda} \end{aligned} \quad (\text{A.17})$$

one finds the Callan-Symanzik equations in the form

$$\left[\mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} - \ell \gamma(g) \right] \Gamma^{(2\ell)} = 0 \quad (\text{A.18})$$

with

$$\begin{aligned} \beta(g) &= -\epsilon(k-1)g + \frac{g^2}{(4\pi)^{2k}} \frac{(2k)!}{(k!)^3} \frac{\Gamma(1/(k-1))^k}{\Gamma(k/(k-1))} + O(g^3) \\ \gamma(g) &= 2 \frac{g^2}{(4\pi)^{2k}} \frac{1}{(2k-1)!} \frac{\Gamma(1/(k-1))^{2k-1}}{\Gamma((2k-1)/(k-1))} + O(g^3) \end{aligned} \quad (\text{A.19})$$

The β -function exhibits an infrared stable fixed point g_c of order ϵ . Upon integration of the renormalization flow equations, one obtains

$$\begin{aligned} \Gamma^{(2\ell)}(\{p\}, g_2, \mu \exp \int_{g_1}^{g_2} \frac{du}{\beta(u)}) \\ = \exp \left\{ \ell \int_{g_1}^{g_2} \frac{\gamma(u) du}{\beta(u)} \right\} \Gamma^{(2\ell)}(\{p\}, g_1, \mu) \end{aligned} \quad (\text{A.20})$$

Combining with a rescaling of p and μ , this reads

$$\begin{aligned} \Gamma^{(2\ell)}(\{\lambda p\}, g_2, \lambda \mu \exp \int_{g_1}^{g_2} \frac{du}{\beta(u)}) \\ = \lambda^{d-\ell(d-2)} \exp \left\{ \ell \int_{g_1}^{g_2} \frac{\gamma(u) du}{\beta(u)} \right\} \Gamma^{(2\ell)}(\{p\}, g_1, \mu) \end{aligned} \quad (\text{A.21})$$

Choose g_1 as a function of g_2 and λ through

$$\lambda \exp \int_{g_1}^{g_2} \frac{du}{\beta(u)} = 1 \quad (\text{A.22})$$

and let $\lambda \rightarrow 0$, which drives g_1 to the fixed point g_c . We get therefore the scaling laws

$$\lim_{\lambda \rightarrow 0} \Gamma^{(2\ell)}(\{\lambda p\}, g_2, \mu) = \lambda^{d-\ell(d-2+\gamma(g_c))} \Gamma^{(2\ell)}(\{p\}, g_c, \mu) \quad (\text{A.23})$$

The anomalous dimension of the field φ is therefore $\frac{1}{2}\gamma(g_c)$ and the exponent $\eta^{[k]}$ defined as in the φ^4 theory as giving the departure of the two-point correlation from a $r^{-(d-2)}$ behaviour (the appended index $[k]$ recalls the order of multicriticality) is given by

$$\eta^{[k]} = \gamma(g_c) \quad (\text{A.24})$$

To lowest order in ϵ

$$\frac{g_c}{(4\pi)^k} = \epsilon(k-1) \frac{(k!)^3}{(2k)!} \frac{\Gamma(k/(k-1))}{\Gamma(1/(k-1))^k} + \dots \quad (\text{A.25})$$

and

$$\eta^{[k]} = 2 \left[\frac{g_c}{(4\pi)^k} \right]^2 \frac{1}{(2k-1)!} \frac{\Gamma(1/(k-1))^{2k-1}}{\Gamma((2k-1)/(k-1))} + \dots \quad (\text{A.26})$$

reads

$$\eta^{[k]} = 4 [\epsilon(k-1)]^2 \left[\frac{k!^2}{(2k)!} \right]^3 + \dots \quad (\text{A.27})$$

a result which agrees to order ϵ^2 with equation (262) in the case $k=2$ when the number n of field components is set equal to one. In both cases, $\eta^{[2]} = \frac{1}{54}\epsilon^2 + \dots$. One could of course compute other exponents such as the anomalous dimensions of the fields φ^2 , φ^3 , ...

As will be shown in chapter 9 (volume 2), when $d=2$, the anomalous dimension of the field φ for the tricritical model ($k=3$) is believed to be $3/40$ leading to $\eta^{[3]} = 3/20$. If we estimate the same exponent using the leading term only, given in (A.27), taking $\epsilon=1$, $k=3$, we find $\eta_{\text{app}}^{[3]} = \frac{1}{500}$. The disagreement by a factor 75 is quite striking, and increases exponentially with the order of multicriticality. It is suggested that in two dimensions

$$d=2 \quad \eta^{[k]} = \frac{3}{(k+1)(k+2)} \quad (\text{A.28})$$

while the estimate from equation (A.27) obtained by taking $(k-1)\varepsilon = 2$, reads

$$d = 2 \quad \eta_{\text{app}}^{[k]} = 16 \left[\frac{k!^2}{(2k)!} \right]^3 \quad (\text{A.29})$$

The reason for this discrepancy may be understood from the fact that the true expansion parameter is not the small number ε , but $\varepsilon(k-1)$ which takes the finite value two (the dimension of the bare coupling) when one attempts to reach two dimensions starting from the upper critical dimension. Thus the ε -expansion is useless under these circumstances, unless one attempts a resummation based on the computation of several other terms in the expansion.

Notes

The renormalization group methods as an effective tool in the computation of critical properties are due to K.G. Wilson. Some relevant works have already been quoted in the previous chapter. Among early contributions are those by C. Di Castro and G. Jona Lasinio, *Phys. Lett.* **29A**, 322 (1969); K.G. Wilson and M.E. Fisher, *Phys. Rev. Lett.* **28**, 240 (1972); K.G. Wilson, *Phys. Rev. Lett.* **28**, 548 (1972). Rather than giving an exhaustive list of articles which would require a full volume, we refer the reader to the following original reviews, K.G. Wilson and J. Kogut, *Physics Report* **12C**, 75 (1974); M.E. Fisher, *Rev. Mod. Phys.* **46**, 597 (1974); K.G. Wilson, *Rev. Mod. Phys.* **47**, 773 (1975); E. Brezin, J.C. Le Guillou and J. Zinn-Justin in volume VI of the Domb and Green series (1976); S.K. Ma, *Modern Theory of Critical Phenomena*, Benjamin, New York (1976); G. Parisi's contribution to the 1973 Cargèse school which appeared in *J. Stat. Phys.*, **23**, 49 (1980). The book by D.J. Amit, *Field Theory, the Renormalization Group and Critical Phenomena*, 2nd edition, World Scientific, Singapore (1984) reflects also in part the influence of the Saclay group of our friends E. Brézin, J.-C. Le Guillou and J. Zinn-Justin, who have been, together with their collaborators, among the most active in developing a systematic field theoretic approach. We have based most of our presentation on their impressive work as well as G. Parisi's contribution.

Forthcoming books by G. Parisi and by J. Zinn-Justin should give an exhaustive and authoritative view.

The three-dimensional computations quoted in the text are from G.A. Baker, B.G. Nickel, M.S. Green and D.I. Meiron, *Phys. Rev. Lett.* **36**, 1351 (1976); G.A. Baker, B.G. Nickel and D.I. Meiron, *Phys. Rev.* **B17**, 1365 (1978); J. Zinn-Justin and J.C. Le Guillou, *Phys. Rev.* **B21**, 3976 (1980). No attempt has been made to incorporate the latest progress in this domain.

A lattice model for the tricritical singularity is due to J. Blume, V.J. Emery, R.B. Griffiths, *Phys. Rev.* **A4**, 1071 (1971). For the general scaling theory, see E.K. Riedel and F.J. Wegner, *Phys. Rev.* **B7**, 248 (1973), R.B. Griffiths, *Phys. Rev.* **B7**, 545 (1973), E. Abrahams, M.J. Stephen and J.P. Straley, *Phys. Rev.* **B12**, 256 (1975) and M. Wohrer, Thesis, Saclay (1976). The computations in appendix A are due to E. Brézin and M. Bauer.

Some of the works dealing with the continuous aspects in statistical mechanics are based on earlier investigations in quantum field theory. The renormalization group appears in E.C.G. Stueckelberg and A. Peterman, *Helv. Phys. Acta* **26**, 499 (1953); M. Gell-Mann and F.E. Low, *Phys. Rev.* **95**, 1300 (1954). A thorough discussion is found in N.N. Bogoliubov and D.V. Shirkov, *Introduction to the Theory of Quantized Fields*, Interscience, New York (1959). The modern approach with its emphasis on broken scale invariance is due to K. Symanzik, *Comm. Math. Phys.* **18**, 227 (1970) and C.G. Callan, *Phys. Rev.* **D2**, 1541 (1970).

This short summary of some essential references does not do justice to the impressive amount of contributions to the subject and to the impact of renormalization group ideas in many different areas.

LATTICE GAUGE FIELDS

Up to now, continuous field theory has appeared as a tool in the study of critical phenomena. Conversely, techniques from statistical mechanics can be useful in field theory. In 1973, Wilson proposed a lattice analog of the Yang-Mills gauge model. Its major aim was to explain the confinement of quarks in quantum chromodynamics. The lattice implementation of a local symmetry yields a transparent geometric interpretation of the gauge potential degrees of freedom, the latter being replaced by group elements assigned to links. Strong coupling expansions predict a linearly rising potential energy between static sources. Complex phase diagrams emerge when gauge fields are coupled to matter fields, and new phenomena appear, such as the absence of local order parameters. The discretization of fermions leads also to interesting relations with topology. This chapter is devoted to the theoretical developments of these ideas.

6.1 Generalities

6.1.1 Presentation

Schematic as they are, statistical models have directly a physical background at any temperature. Lattices may represent the crystalline structure of solids. They play an important role at short distance, but become irrelevant in the critical region, except as a regulator for the field theory describing the approach to critical points. The opposite point of view can also be considered. A lattice is artificially introduced as a regulator for a continuous field theory. The lattice system has no physical meaning, but can be studied at any "temperature", so that one can get information about its critical region, hopefully described by the initial continuous theory.

In particle physics, an important class of models can be built using the idea of minimal interactions with a gauge field, enabling the conservation of local internal symmetries (Yang-Mills, 1956). One of their striking properties is to be asymptotically free in the ultraviolet regime in four dimensions, at least for nonabelian continuous gauge groups. At short distances, the effective coupling between fields – describing "quarks" and "gluons", assumed to be the fundamental constituents of hadrons – decreases indefinitely, and this fact allows a perturbative treatment. Hence theoretical predictions can be compared to experimental data in this regime. *Quantum chromodynamics*, based on an internal *color* $SU(3)$ symmetry, provides results in remarkable agreement with experimental observations. Examples are deep inelastic scattering of leptons on nuclear targets, hadronic jet production in very high energy collisions, or the spectroscopy of bound states involving heavy quarks. Yang-Mills fields are also used in the unified electroweak model of Glashow, Weinberg and Salam, which is now considered, together with chromodynamics, as the standard model of particle physics.

However, the large-distance behaviour of chromodynamics cannot be reached in the perturbative approach. This method does not explain the permanent confinement of quarks, which cannot be experimentally isolated as elementary particles, at least under present laboratory conditions. The determination of the hadronic mass spectrum – bound states of quarks and gluons – is also a typical strong coupling problem which can be studied by perturbative field theory only within a phenomenological framework. The discretization of space-time using a lattice, has provided a partial answer to these questions. Lattice gauge field systems are studied using statistical techniques; high temperature series yield strong coupling expansions. Mean field approximation methods, as well as numerical simulations, also provide interesting results. The study of the critical region should now be refined for strong interaction physics. Ideas developed in gauge field theory for particle physics have applications in disordered or amorphous systems. Finally, there is some connection between gauge fields and general relativity (which can also be discretized following Regge (1960, see chapter 11 in volume 2)).

The fundamental concept is gauge invariance. Let us consider transformations $g(\mathbf{x})$, belonging to some group G , and depending

on the space point x . "Charged" matter fields $\varphi(x)$ are transformed under a finite, generally linear representation of $g(x)$ at the same point $\varphi(x) \rightarrow g(x)\varphi(x)$. This is a shorthand notation for the action of the group elements on the fields. In a continuous model, derivative terms of the Lagrangian, which are responsible for propagation, induce variations $\delta g(x)$. For continuous groups, if one wishes to construct a local gauge invariant model, these terms should be compensated by the introduction of a new gauge field $A_\mu(x)$, taking its values in the Lie algebra of G . This leads to the concept of minimal coupling. In a discretized model, derivative terms are replaced by finite differences, so that the action contains bilocal terms, for instance the nearest neighbour interaction

$$S_{\text{mat}} = \beta_m \sum_{(i,j)} \varphi_i \cdot \varphi_j \quad (1)$$

The above notation has been shortened for convenience. Fields $\varphi_i = \{\varphi_i^\alpha, \alpha = 1, 2, \dots, d_r\}$ are transformed through the linear representation r of the symmetry group $G : \varphi_i^\alpha \rightarrow (g\varphi_i)^\alpha = \mathcal{D}_{\alpha\beta}^r(g)\varphi_i^\beta$. The dot product $\varphi_i \cdot \varphi_j$ is the quadratic or sesquilinear form which is preserved by the group. Up to now, we have mostly studied models invariant under the orthogonal group $SO(n)$; we are also interested in this chapter by the unitary group $SU(n)$, which leaves invariant the sesquilinear form $\sum_\alpha \varphi_i^{\alpha*} \varphi_j^\alpha$. Simpler groups (Z_2 , Z_n , $U(1)$, ...) will also be considered.

The action in equation (1) is only globally invariant. In order to make it locally invariant, one introduces gauge fields $U_{ij} \in G$ associated to any ordered pair of nearest neighbours (Wilson, 1974). Thus, in going from a continuous to a discretized form, the novelty is the use of the group instead of the Lie algebra. This has the virtue of making local invariance explicit and simple. It also affords the opportunity to use discrete gauge groups which have no direct interpretation in the continuous limit. The modified action

$$S_{\text{mat}} = \beta_m \sum_{(i,j)} \varphi_i U_{ij} \varphi_j = \beta_m \sum_{(i,j)} \varphi_i^{\alpha*} \mathcal{D}_{\alpha\beta}^r(U_{ij}) \varphi_j^\beta \quad (2)$$

is invariant if the gauge fields transform as

$$U_{ij} \rightarrow g_i U_{ij} g_j^{-1} \quad (3)$$

with the constraint

$$U_{ij} \equiv U_{ji}^{-1} \quad (4)$$

The field U_{ij} is not associated to a lattice site, but to an *oriented* lattice link ij . The action has to be completed with an additional term S_J depending only on these new fields. Let us discuss the general form of such a term, which should of course be locally invariant. The product of fields along a closed curve $C = i_1 i_2 \dots i_n i_1$ drawn on the lattice

$$U_C = U_{i_1 i_2} U_{i_2 i_3} \dots U_{i_n i_1} \quad (5)$$

transforms as

$$U_C \rightarrow g_{i_1} U_C g_{i_1}^{-1} \quad (6)$$

i.e. it remains in the same conjugacy class of the group. Let us recall briefly that the group G is partitioned into classes by the following relation. Two elements h and h' are equivalent (conjugate) if there exists $g \in G$ such that $gh = h'g$. A function $\chi(U_C)$ depending only on the class of U_C is gauge invariant, and can therefore be used in the action S_J . The cyclic invariance property of such a function implies that $\chi(U_C)$ is moreover independent from the origin chosen on the closed curve C . Finally, χ should also be a real function (so is the action), and hence the orientation of C has no influence.

Among all possible choices, the simplest curve C (the shortest one) is the boundary of a *plaquette*, an elementary square on the hypercubical lattice. The path forth and back along a link leads to a trivial term, due to the constraint (4). From group theory, any class function can be decomposed on the basis of characters χ_r , which are the traces of the corresponding matrices in the irreducible representation $r : \chi(U_C) = \sum_r c_r \chi_r(U_C)$. The "natural" choice of the fundamental representation for the matrix groups $SU(n)$ and $SO(n)$ leads to the Wilson's action

$$S_J = \beta_J \sum_p \frac{1}{n} \text{Re Tr } U_p \quad (7)$$

with $U_p = U_{ij} U_{jk} U_{kl} U_{li}$ for the plaquette $p \equiv ijkl$. The dimension of the fundamental representation of G in which the trace Tr is taken is n .

Geometrical actions Another choice follows from the following considerations. When β_J is large, U_p fluctuates around unity. In this case (which, as will be seen later, corresponds to the continuous

limit), the most important terms are quadratic in the fluctuations of the field. A geometrical choice uses the geodesic distance from U_p to unity to measure these fluctuations. The metric is induced on G by the invariant quadratic form on the Lie algebra. For $G = U(1)$, $U_p = \exp(i\theta_p)$ and the action $\beta_J \sum_p \theta_p^2$ seems natural. However, it is multivalued. This defect may be corrected using a generalization of an idea due to Villain for the abelian case. For any Lie group, the *heat kernel* Boltzmann factor is the response to a local heating at the unit element on the group manifold after a time $t \approx 1/\beta_J$. This manifold is a circle for the group $U(1)$, a 3-sphere for $SU(2)$,... Hence, the Boltzmann factor is the solution of the heat equation

$$\Delta_g f(g, t) = \frac{\partial f(g, t)}{\partial t}, \quad \text{with } f(g, 0) = \delta(g, 1) \quad (8)$$

where Δ_g is the Laplace-Beltrami operator on the group. The time is proportional to the inverse of β_J . This leads to the generalized Villain action

$$\exp S_J = \prod_p \exp S_J(p),$$

$$\exp S_J(p) = \sum_r d_r \chi_r(U_p) \exp \left(-\frac{C_r^{(2)}}{n\beta_J} \right) \quad (9)$$

where $C_r^{(2)}$ is the value of the quadratic Casimir operator in the representation r , i.e. the corresponding eigenvalue of $-\Delta_g$. Summation runs over all inequivalent irreducible representations of dimension d_r (harmonic analysis on the group). In the $SU(2)$ case, if U is a generic 2×2 matrix conjugate to the diagonal matrix $(e^{i\theta/2}, e^{-i\theta/2})$, the plaquette contribution is the function

$$\exp S_J(p) = \sum_{2j=0}^{\infty} (2j+1) \frac{\sin(j + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta} \exp -\frac{j(j+1)}{2\beta_J} \quad (10)$$

Any other reasonable action is asymptotically equivalent to this one in the large β_J limit. The corresponding Boltzmann weight avoids in particular spurious singularities which may appear in some special limits such as the $n \rightarrow \infty$ limit of Wilson's action.

6.1.2 The continuous limit

Before embarking on a detailed analysis of gauge systems at any coupling β_J , we consider first its formal continuous limit. For the time being, we assume that fluctuations around this limit are small and without essential qualitative effects. The reasoning is only

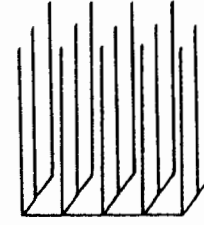


Fig. 1 Maximal tree on a cubic lattice, defining the axial gauge

valid for continuous gauge groups. As in the previous chapters, we develop the action in local terms depending only on fields and their derivatives at a point. The resulting local, regularized Lagrangian is then studied using renormalization group techniques. This procedure, however, needs some modifications to take into account specific properties of the gauge field U_{ij} .

The first step is to characterize the configurations $\{U_{ij}^0\}$ for which the action is maximal and around which the perturbative expansion will be done. They satisfy

$$U_p^0 = U_{ij}^0 U_{jk}^0 U_{kl}^0 U_{li}^0 = 1 \quad (11)$$

for any plaquette (all characters are simultaneously maximal for the group identity). Let us search a gauge transformed configuration $U_{ij}^{0'} = g_i^{-1} U_{ij}^0 g_j$, such that $U_{ij}^{0'} = 1$ on a maximal tree of the lattice, i.e. on a maximal set of links without any loop. Figure 1 displays such a set, and more precisely defines the so-called *axial gauge*, where all gauge fields are set equal to unity in a "timelike" direction. The reader can easily convince himself that the gauge transformation g_i is defined from site to site. It then follows that $U_{ij}^{0'}$ is equal to unity everywhere due to relation (11). Hence the only configurations maximizing the action are pure gauge configurations

$$U_{ij}^0 = g_i g_j^{-1} \quad (12)$$

Generalizing the previous property, one might think that any plaquette configuration $\{U_p\}$ is sufficient to define a link configuration $\{U_{ij}\}$, up to a gauge transformation. Show that this is the case if the gauge group is abelian and if the six plaquette variables

U_p bounding any tridimensional cube satisfy the constraint

$$\prod_{p \in \partial c} U_p = 1 \quad (13)$$

This constraint disappears in two dimensions, and the result remains valid for nonabelian groups. In this case, the variables $\{U_p\}$ are independent and do not interact in the action. As a consequence, a two-dimensional pure gauge model is trivial.

Around the classical solution (12), fields are parametrized using a local gauge field $A_\mu(\mathbf{x})$, taking its values in the Lie algebra of the group G , in the form

$$U_{ij} = g_i \exp(ia A_\mu(\mathbf{x})) g_j^{-1} \quad (14)$$

with $\mathbf{x}_j = \mathbf{x}_i + a\hat{\mu}$, $\mathbf{x} = \frac{1}{2}(\mathbf{x}_i + \mathbf{x}_j)$, a being the lattice spacing. This parametrization cannot be distinguished from another possible one using the path ordered integral $P \exp\{i \int_{t=0}^1 dx^\mu(t) A_\mu(\mathbf{x}(t))\}$, with $\mathbf{x}(t) = (1-t)\mathbf{x}_i + t\mathbf{x}_j$, if one neglects higher order corrections in a . When evaluating the plaquette term U_p , the generally noncommutative products are computed using Baker-Campbell-Hausdorff formula

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\dots}$$

Recall that the general relation

$$\exp C = \exp A \exp B,$$

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] + \frac{1}{12}[B, [B, A]] - \frac{1}{48}[A, [B, [A, B]]] + \dots$$

is obtained by the following algorithm. The product

$$C = \ln \left[\left(1 + A + \frac{A^2}{2!} + \dots \right) \left(1 + B + \frac{B^2}{2!} + \dots \right) \right]$$

is expanded as a sum of homogeneous monomials. If $z_1 z_2 \dots z_p$ is such a monomial of degree p , it is replaced by the expression $(1/p)[z_1, [z_2, [\dots [z_{p-1}, z_p]] \dots]]$ belonging to the Lie algebra generated by A and B .

The expression in terms of the fields at the center of the plaquette is obtained after some straightforward algebra

$$U_p = g_i \exp[ia^2 F_{\mu\nu} + \mathcal{O}(a^4)] g_j^{-1} \quad (15)$$

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu] \quad (16)$$

Finally, the exponential (15) is expanded as a power series in a , and the sum over plaquettes in the action is replaced by an integral (with corrections of order a , which can be computed using the Euler-MacLaurin formula). Up to an additive constant, which can be removed by a redefinition of the partition function, the final result reads

$$S_J = \frac{\beta_J g_0^2 a^{4-d}}{2n} \int d^d \mathbf{x} \left(-\frac{1}{2g_0^2} \text{Tr}(F_{\mu\nu}^2) + \mathcal{O}(a^2) \right) \quad (17)$$

The leading term in this formula coincides with the continuous action postulated by Yang and Mills when the multiplicative term in front of the integral is set equal to unity by an appropriate choice of g_0 . Requiring that the coefficient in front of the integral be unity yields a relation between the inverse temperature β_J of the gauge model and the bare coupling constant g_0 in the continuous theory of the form

$$\beta_J = \frac{2n}{g_0^2} a^{d-4} \quad (18)$$

Naive dimensional analysis leads to $[A_\mu/g_0] = \frac{1}{2}(d-2)$, $[g_0^2] = 4-d$. Hence the continuous theory is renormalizable in four dimensions provided one can show that local invariance can be preserved under renormalization. The higher order terms in a , which have not been explicitly written in equation (17), involve higher derivatives, and their canonical dimension is greater than four. It is therefore expected that, in the large distance limit, their effect is limited to a finite renormalization, without qualitative change in the critical behaviour.

To obtain the continuous limit, one chooses a physical length ξ , computed from correlation functions, and one keeps it fixed (by tuning the bare coupling constant g_0) while the lattice spacing shrinks indefinitely. In lattice units, ξ/a diverges. In other words, one approaches a critical point $g_0 \rightarrow g_{0c}$. This process follows the renormalization group equation

$$a \frac{d\xi}{da} \equiv \left(a \frac{\partial}{\partial a} - \beta(g_0) \frac{\partial}{\partial g_0} \right) \xi = 0 \quad (19)$$

with

$$\beta(g_0) = -a \left. \frac{\partial g_0}{\partial a} \right|_{\xi} \quad (20)$$

The notation $\beta(g_0)$ for the Callan-Symanzik function is the traditional one and should not be confused with the constants β_1, β_m appearing as coefficients in the action. In integral form, equation (20) yields

$$\xi = a \exp \int_{g_0}^{g_0} \frac{dg}{\beta(g)} \quad (21)$$

Hence g_{0c} is a zero of the function $\beta(g)$, with a negative slope (if it is a simple zero). Relation (21) indicates how g_0 approaches g_{0c} when a vanishes. A direct computation of $\beta(g)$ could be performed using the lattice regularization. In chapter 5, we mentioned however that the slope of $\beta(g)$ at g_{0c} is universal (this assertion should be appropriately modified in the case of a multiple zero), while the value of g_{0c} itself depends on the regularization scheme. Therefore, in practice, it is better to perform easier calculations using for instance dimensional regularization. As it turns out, zero coupling is such a critical point with a triple zero in $\beta(g)$, and the two universal coefficients in the expansion

$$\beta(g) = -b_0 g^3 - b_1 g^5 + \mathcal{O}(g^7) \quad (22a)$$

are given, in the case of the $SU(n)$ gauge group, by

$$b_0 = \frac{11}{3} \frac{n}{(4\pi)^2}, \quad b_1 = \frac{34}{3} \frac{n^2}{(4\pi)^4} \quad (22b)$$

While the principle of the computation is simple and follows the discussion of chapter 5, nontrivial technical details make it rather complex in the present case. This is due to an apparent conflict between the desire to maintain local invariance, and the necessity of introducing a gauge fixing term to avoid integration over a continuum of nondynamical gauge degrees of freedom. The way to overcome this difficulty is the subject of an elaborate theory presented in detail in textbooks on the quantization of continuous gauge fields. We only sketch the main steps. Using a basis t_a of the Lie algebra of G , with totally antisymmetric structure constants

f_{abc} such that

$$i[t_a, t_b] = f_{abc} t_c \quad f_{acd} f_{bcd} = C_f^{(2)} \delta_{ab} \quad \text{Tr } t_a t_b = \frac{1}{2} \delta_{ab} \quad (23)$$

one writes $A_\mu = g A_\mu^a t_a$. A term $(\partial_\mu A_\mu^a)^2 / 2\alpha$ is added to the action in order to fix the gauge. Thus the operator in the quadratic part of the Lagrangian has no more zero modes and becomes invertible. This term is compensated by the introduction of anticommuting fields Φ_a (Faddeev-Popov ghost fields), the role of which under integration is to give a Jacobian necessary to maintain gauge invariance. Writing the action in the form

$$S = - \int d^d x \mathcal{L} \quad (24a)$$

the resulting Lagrangian, written in terms of renormalized quantities, reads

$$\begin{aligned} \mathcal{L} = & \frac{1}{4} Z_3 (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2 + \frac{1}{2\alpha} (\partial_\mu A_\mu^a)^2 \\ & + Z_1 g f_{abc} \partial_\mu A_\nu^a A_\mu^b A_\nu^c + \frac{1}{4} g^2 \frac{Z_1^2}{Z_3} f_{abc} f_{ade} A_\mu^b A_\nu^c A_\mu^d A_\nu^e \\ & + Z_3^\Phi \partial_\mu \Phi_a^* \partial_\mu \Phi_a + Z_1^\Phi g f_{abc} \partial_\mu \Phi_a^* A_\mu^b \Phi_c \end{aligned} \quad (24b)$$

That the coefficient of the quartic term in $\mathcal{A}$ has the indicated form, is a consequence of Ward-Slavnov identities, expressing gauge invariance, which also implies the relation $Z_1/Z_3 = Z_1^\Phi/Z_3^\Phi$. The longitudinal part of the propagator is not renormalized. The calculation is performed in $d = 4 - \epsilon$ dimensions, with a dimensionless coupling constant $u = g\mu^{-\epsilon/2}$. The mass scale is therefore defined by μ . The subtraction terms $(Z_i - 1)$ are functions of u , α and ϵ , and are determined order by order in u in such a way that the limit $\epsilon \rightarrow 0$ of physical quantities is finite. We adopt the minimal subtraction scheme, involving only divergent terms behaving as inverse powers of ϵ , without additional finite renormalization. Feynman rules are summarized in figure 2, where a heavy line indicates a gauge potential propagator, while a broken one is associated to a ghost propagator. We adopt the Feynman gauge $\alpha = 0$, which leads to more tractable calculations.

Subtraction terms are determined from the divergent two- and three-point functions. The corresponding one-loop diagrams are displayed in figure 3. Integration techniques for the dimensional continuation have been described in section 5.2.3. For instance, the *unsubtracted* gauge field propagator at the one-loop level (figure 3a)

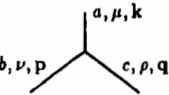
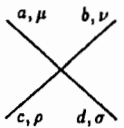
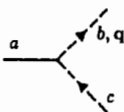
$\frac{a, \mu}{\quad} \quad \frac{b, \nu}{\quad}$	$\frac{1}{(2\pi)^4} \frac{\delta_{ab}}{p^2} \left(\delta_{\mu\nu} - \alpha \frac{p_\mu p_\nu}{p^2} \right)$
$\frac{a}{\quad} \rightarrow \frac{b}{\quad}$	$\frac{1}{(2\pi)^4} \frac{\delta_{ab}}{p^2}$
	$(2\pi)^4 i g f_{abc} [\delta_{\nu\rho}(q-p)_\mu + \delta_{\rho\mu}(k-q)_\nu + \delta_{\mu\nu}(p-k)_\rho]$
	$-(2\pi)^4 g^2 f_{ace} f_{bde} (\delta_{\mu\nu} \delta_{\rho\sigma} - \delta_{\mu\sigma} \delta_{\nu\rho}) - (2 \text{ perm.})$
	$(2\pi)^4 i g f_{abc} q_\alpha$

Fig. 2 Feynman rules for the continuous Yang-Mills Lagrangian.

reads

$$\Pi_{\mu a; \nu b}^{(1)}(p) = \frac{1}{2} g^2 \pi^{d/2} C_f^{(2)} \delta_{ab} \Gamma(2 - \frac{1}{2}d) [(4d-8)B(\frac{1}{2}d, \frac{1}{2}d) + (6-d)B(\frac{1}{2}d-1, \frac{1}{2}d-1)] (p^2)^{d/2-2} (p^2 \delta_{\mu\nu} - p_\mu p_\nu) \quad (25)$$

with

$$B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$$

Expressing this result in terms of the dimensionless variable u and expanding in ϵ , one derives at one-loop order the necessary counterterm leading to a finite result when $\epsilon \rightarrow 0$

$$(Z_3 - 1)^{(1)} = \frac{10}{3} \frac{u^2}{16\pi^2} C_f^{(2)} \frac{1}{\epsilon} \quad (26)$$

The other three Green functions described in figure 3 can be similarly computed. The choice of zero external momenta in the 3-point functions leads to simpler calculations. One obtains the

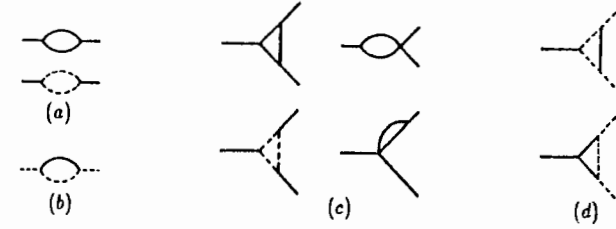


Fig. 3 One loop diagrams for 2- and 3-point functions. (a) Gauge potential propagator. (b) Ghost propagator. (c) Gauge vertex. (d) Mixed gauge and ghost vertex.

following results

$$\begin{cases} (Z_3^\Phi - 1)^{(1)} = \frac{u^2}{16\pi^2} C_f^{(2)} \frac{1}{\epsilon} \\ (Z_1 - 1)^{(1)} = \frac{4}{3} \frac{u^2}{16\pi^2} C_f^{(2)} \frac{1}{\epsilon} \\ (Z_1^\Phi - 1)^{(1)} = -\frac{u^2}{16\pi^2} C_f^{(2)} \frac{1}{\epsilon} \end{cases} \quad (27)$$

The function $\beta(u)$ is defined as

$$\beta(u) = \mu \frac{\partial u}{\partial \mu} \Big|_{g_0, \alpha_0} \quad (28)$$

where in the general case the unrenormalized constants g_0 and α_0 are expressed in terms of the renormalized ones through

$$g_0 = \frac{Z_1(u, \alpha)}{Z_3^{3/2}(u, \alpha)} g, \quad \alpha_0 = Z_3(u, \alpha) \alpha \quad (29)$$

Hence,

$$\beta(u) = \frac{-\frac{1}{2}\epsilon u}{1 + u \frac{d}{du} [\ln Z_1 Z_3^{-3/2}]} = -\frac{1}{2}\epsilon u - \frac{11}{3} u^3 \frac{C_f^{(2)}}{16\pi^2} + \dots \quad (30)$$

which agrees, in the limit $\epsilon \rightarrow 0$, with the leading term in equations (22a)-(22b) (since, for $SU(n)$, the quadratic Casimir invariant is given by $C_f^{(2)} = n$).

The two-loop calculations are performed in a similar way. The diagrams are much more numerous and include the one-loop counterterms. The Slavnov identity $Z_1/Z_3 = Z_1^\Phi/Z_3^\Phi$ which expresses gauge invariance, is imposed term by term at any order.

The subtraction constant Z_1 can then be determined from the other ones.

According to equation (22), the pure gauge theory is asymptotically free in the ultraviolet regime ($g_{0c} = 0$). Starting from a lattice version, the physical continuous region corresponds to large β_j according to equation (18). We introduce a standard mass scale for the lattice regularization (identified by the index L) through

$$\xi_L^{-1} = \Lambda_L \equiv \frac{1}{a} (b_0 g_0^2)^{-b_1/2b_0^2} \exp\left(\frac{-1}{2b_0 g_0^2}\right) \quad (31)$$

This relation is obtained from the integral in equation (21) in the vicinity of $g_0 \sim 0$ using the two leading terms in the β -function. Any quantity expressed in this scale should have a finite limit when letting the lattice spacing a tend to 0. In other words, any physical length ξ will be computed using its finite limit $\xi \Lambda_L$ when β_j grows indefinitely.

It is important in practical applications of the lattice technique to be able to express the relation between Λ_L and similar quantities corresponding to other regularization schemes. Let us consider for instance another regularization characterized by a mass regulator m which will be sent to infinity. In the two schemes, the same correlation lengths $\xi_i \Lambda_L$ and $\xi_i \Lambda_m$ (expressed in the convenient scale) have finite limits. As the two models describe the same physics, the ratio of these limits is independent of the quantity ξ_i considered. The knowledge of the ratio Λ_L/Λ_m is not of purely academic interest. Indeed, one wants to relate perturbative predictions – achieved using the most convenient regularization, such as the dimensional regularization – to nonperturbative quantities, such as particle masses, computed numerically on a lattice.

This ratio, which involves a comparison of the ultraviolet cutoffs, would be known if the β function was explicitly computed in the lattice regularization. However, we did not perform the very complex two-loop calculations. There exists a trick which yields the required ratio by computing only one-loop diagrams. Fields are expanded in the neighborhood of a configuration $\{U_0\}$, which satisfies the classical motion equations. This is called the *background gauge*. Performing this expansion, the action is written as $S_j = S_0 + S_2 + \dots$, S_2 being quadratic in the fluctuations, and $S_0 = S_j(U_0) = (-1/2g_0^2) \int d^4x \text{Tr}(F_{\mu\nu}^0)^2$. Further terms will not be

needed. The partition function reads

$$Z_L(a) = e^{-S_0} \int dU e^{-S_2} (1 + \mathcal{O}(g_0^2(a))) \quad (32)$$

Comparing with the similar formula obtained with the other regularization, one obtains

$$-\ln \frac{Z_L(a)}{Z(m)} = \left[\frac{1}{2g_0^2(a)} - \frac{1}{2g_0^2(m)} + f(ma) \right] \int d^4x \text{Tr}(F_{\mu\nu}^0)^2 \quad (33)$$

The function $f(ma)$, resulting from the calculation of the two path integrals (32), corresponds to a finite renormalization. Its explicit form can be predicted from the (logarithmic) divergences of the one-loop diagrams

$$f(ma) = \frac{11n}{48\pi^2} [\ln ma + C] \quad (34)$$

Formula (31) relates Λ_L to a and a similar relation can be written between Λ_m and m . Equation (33) is thus rewritten as

$$\ln \frac{Z_L(a)}{Z(m)} = \frac{11n}{48\pi^2} \left(\ln \frac{\Lambda_L}{\Lambda_m} - C \right) \int d^4x \text{Tr}(F_{\mu\nu}^0)^2$$

As the two regularizations lead to the same physical results, the finite renormalization has to disappear, requiring that

$$\Lambda_L = \Lambda_m e^C \quad (35)$$

A detailed and rather lengthy calculation enables one to extract the value of the constant C from the comparison of the two-point function for the field $F_{\mu\nu}^0$ in the two regularizations. A choice of gauge is necessary, introducing Faddeev-Popov ghosts. We only indicate here one final result, comparing the Wilson action to the minimal dimensional regularization (index MS) for the gauge group $SU(n)$

$$\ln \frac{\Lambda_{MS}}{\Lambda_L} = \frac{1}{2} (\ln 4\pi - \gamma) - 0.0224780n \quad (36)$$

The numerical constant is obtained from estimates of lattice integrals, and $\gamma \approx 0.57721$ is the Euler's constant. Other formulae, corresponding to more general lattice actions, can be found in the references.

6.1.3 Order parameter and Elitzur's theorem

In the various types of models presented in previous chapters, finitely many local order parameters discriminate among the

possible phases of a statistical system. For instance, the existence of a nonvanishing magnetization signals a spontaneous breaking of global symmetry in a spin model. One might imagine a similar mechanism in gauge theories, and study for instance the mean value $\langle U_{ij} \rangle$ of the field. This simple idea does not work. Elitzur has proved very generally that any *local* quantity (i.e. using only a finite number of fields) which is not gauge invariant (as $\langle U_{ij} \rangle$) has a vanishing mean value *at any temperature*, and therefore cannot be used as an order parameter. In other words, *local* observables cannot exhibit spontaneous breaking of *local* gauge symmetry.

Before proceeding to the proof of this theorem, let us discuss once more the significance of a nonvanishing magnetization. In the absence of an external field, such a phenomenon might at first seem strange, since the two opposite configurations $\{\sigma_i\}$ and $\{-\sigma_i\}$ contribute with the same Boltzmann weight. However, at low temperature, the two regions in configuration space representing opposite magnetizations can only be connected by a dynamical path involving the creation of an infinite interface at the price of an infinite energy cost. Hence the process cannot occur spontaneously. Another description uses a small external field h . The two fundamental configurations $\{s\sigma_i = +1\}$ and $\{\sigma_i = -1\}$ are now energetically separated by a gap $2Nh$, with N the total number of sites. One of these is thus eliminated in the thermodynamic limit $N \rightarrow \infty$, and when the symmetry breaking field is then set equal to zero, one is left with a nonvanishing magnetization. Very generally, the mean value of a quantity $f(\varphi)$ should be computed *in the presence of a vanishingly small external field* J and the thermodynamic limit $N \rightarrow \infty$ should be taken *before* taking the limit $J \rightarrow 0$; the two limits *do not commute*. We denote generically by φ the gauge field taking its values in a compact group manifold and define

$$\langle f(\varphi) \rangle = \lim_{J \rightarrow 0} \lim_{N \rightarrow \infty} \langle f(\varphi) \rangle_{N,J} \quad (37)$$

with

$$\langle f(\varphi) \rangle_{N,J} = Z_{N,J}^{-1} \int \exp(S(\varphi) + J \cdot \varphi) f(\varphi) d\mu(\varphi) \quad (38)$$

In the context of a local symmetry, the proof of Elitzur's theorem applies to a quantity $f(\varphi)$, not invariant under the gauge group $\varphi \rightarrow {}^g\varphi$ which leaves both the action and the field measure

invariant. Non invariant means here that this quantity has no invariant component, i.e.

$$\int f({}^g\varphi) dg = 0 \quad (39)$$

The gauge group is local and the above integral is understood on the lattice as performed over each $g(\mathbf{x})$ independently. Thus it is possible to consider the subgroup of those transformations which act only on the finite set $\{\varphi'\}$ on which the local function $f(\varphi)$ depends and which leaves invariant the complementary subset $\{\varphi''\} : \varphi'' = {}^g\varphi''$. Performing the change of variable $\varphi \rightarrow {}^g\varphi$ in the integral (38) and taking into account the invariance of the action and of the measure, the mean value reads (after averaging over this gauge subgroup)

$$\langle f(\varphi) \rangle_{N,J} = Z_{N,J}^{-1} \int \exp(S(\varphi) + J' \cdot {}^g\varphi' + J'' \cdot \varphi'') f({}^g\varphi) d\mu(\varphi) dg \quad (40)$$

For a source term J bounded by a small enough quantity ϵ , the crucial point is the inequality

$$|\exp(J' \cdot \varphi') - 1| \leq \eta(\epsilon) \quad (41)$$

with $\eta(\epsilon)$ vanishing uniformly with ϵ , whatever the field configuration φ' and the total number of sites N . This inequality holds for gauge systems, because the subset $\{\varphi'\}$ is finite and independent of N , and because the fields φ are assumed to take their values in a compact domain. It does not hold for spin systems, because the partition $\{\varphi\} = \{\varphi'\} \cup \{\varphi''\}$ cannot be performed in the case of a global symmetry, so that $J \cdot \varphi$ is an extensive quantity proportional to N . The integral (40) is split into two terms, writing $\exp J' \cdot \varphi' = 1 + (\exp J' \cdot \varphi' - 1)$. The first one vanishes identically when integrating over g , due to condition (39). The second one is bounded using the inequality (41) and one gets

$$|\langle f(\varphi) \rangle_{N,J}| \leq 2\eta(\epsilon) \text{Sup} f$$

Now the two limits are taken in the correct order and this leads to the expected assertion

$$\langle f(\varphi) \rangle \equiv 0 \quad (42)$$

when the source term vanishes. As a result, any order parameter should be a nonlocal quantity. The physical interpretation of

the gauge model has led Wilson to introduce such a quantity, henceforth called a *Wilson loop*. Let us consider a closed loop C and define

$$W(C) = \langle \text{Tr } U_C \rangle \quad (43)$$

The trace refers here to the fundamental representation (although other choices might also be interesting). With the following assumptions

- β_j small (strong coupling region),
- pure gauge system, with *no coupled matter fields*,
- compact gauge group with a nontrivial center,

we will prove in section 3 that $-\ln W(C)$ is asymptotically proportional to the minimal area $A(C)$ enclosed by the curve C , when C grows indefinitely (area law). The proof is based on the convergence of the strong coupling expansion for this quantity and uses the fact that the leading contribution of this series arises from diagrams involving plaquettes covering the minimal surface bounded by C .

In the weak coupling region, this assertion might be false. For discrete abelian groups such as Z_2 , the area law fails in this regime, as will be proved below using duality arguments. We give here a heuristic perturbative argument, valid for a continuous symmetry group (it becomes, however, inexact for nonabelian, four-dimensional systems which are asymptotically free and where the strong coupling behaviour is expected to extend qualitatively down to small coupling). The parametrization (13) of the fields leads to the representation

$$W(C) = \left\langle \text{P exp } i \oint_C A_\mu dx^\mu \right\rangle = \exp \left(-\frac{1}{2} \oint_C \oint_C \Delta(x-y) dx dy + \dots \right) \quad (44)$$

The propagator $\Delta(x-y)$ of the gauge field A_μ favours small separations because of its short-distance singularity. Hence the exponential argument in (44) is proportional to the length of the curve C .

The *string tension* is defined as the limit

$$K = \lim_{C \text{ large}} \left(-\frac{\ln W(C)}{A(C)} \right) \quad (45)$$

It enables one to distinguish between a *confined* phase ($K \neq 0$) and a nonconfined one ($K \equiv 0$). The quantity K plays the role of an order parameter, albeit through a very nonlocal procedure, and, using this criterion, two phases are always seen to coexist in sufficiently high dimension. The above argument suggests that the confined phase is out of reach of continuous perturbative theory (in spite of many ingenuous attempts). This provides strong motivations for the introduction of lattice gauge theory.

What is the physical interpretation of K ? Let us imagine the following process. At some time, a quark-antiquark pair is produced and separated up to a distance R . Its potential energy is then $V(R)$. The two quarks propagate during a time T , at a fixed separation R , and then annihilate. The corresponding probability amplitude is $\exp(-V(R)T)$ (since we are in Euclidean space, the usual expression $\exp(iEt)$ is replaced by $\exp(-Et)$). This amplitude is precisely the mean value of the Wilson loop operator, for the rectangular curve with area $R \times T$ bounded by the trajectory of the pair. At large distance and strong coupling, the potential is asymptotically linear $V(R) = KR$ and prevents any separation of the two particles which remain in a bound state. In the unconfined phase, where a perimeter law holds, $V(R)$ is bounded and it is possible to separate the two particles with only a finite energy. This interpretation explains why such an order parameter cannot be used in the presence of dynamical matter fields. Indeed, if the separation increases, an additional pair can be created and tends to restore a perimeter law for the Wilson loop parameter, at any coupling.

Compute the string tension K for a two-dimensional lattice gauge system.

6.1.4 Duality

We have already encountered duality when considering the two-dimensional Ising model. It enabled one to obtain relations which determined the critical point and played an important role in the exact solution. Duality can be generalized to arbitrary *abelian* gauge systems defined on an abstract *complex* of dimension d , with the associated notions of neighborhoods and boundaries. In particular, a lattice is a complex, the *cells* of which are (in

increasing dimension) the sites, the links, the plaquettes, the cubes, etc. In general, a complex is a set of cells $c_k^{(i)}$ with the following properties.

- 1) With each cell $c_k^{(i)}$ is associated a nonnegative integral *dimension* k , and the upper bound of cell dimensions is the *dimension* d of the complex.
- 2) With each cell $c_k^{(i)}$ is associated a cell of *opposite orientation* denoted $-c_k^{(i)}$. This mapping is involutive: $-(-c_k^{(i)}) = c_k^{(i)}$.
- 3) With each pair of cells, the dimension of which differs by unity, is associated an integer $(c_k^{(i)} : c_{k-1}^{(j)})$ called the *incidence number*. If it does not vanish, $c_{k-1}^{(j)}$ is called a *face* of $c_k^{(i)}$. The incidence number changes sign whenever the orientation of either cell is reversed

$$(c_k^{(i)} : -c_{k-1}^{(j)}) = (-c_k^{(i)} : c_{k-1}^{(j)}) = -(c_k^{(i)} : c_{k-1}^{(j)}).$$

- 4) For any pair of cells c_k and c_{k-2} , one requires that

$$\sum_i (c_k : c_{k-1}^{(i)}) (c_{k-1}^{(i)} : c_{k-2}) = 0$$

This relation is the basis of homology and cohomology theory.

Show that any lattice is a complex. We shall return to such constructions in chapter 11 (volume 2) when studying random geometry.

A k -chain with values in an abelian group G (denoted additively in this section, in contradistinction with our usual notation) is an odd function defined on the cells of dimension k in G

$$c_k^{(i)} \rightarrow \varphi_i \in G \quad -c_k^{(i)} \rightarrow -\varphi_i$$

These functions have themselves the structure of an abelian group, and we denote a k -chain by $\sum_i \varphi_i c_k^{(i)}$. The *null chain*, denoted by 0, is the chain which assigns the zero in G to all cells. Finally, the boundary and coboundary operators denoted Δ and ∇ (this notation should not be confused with Laplacian and gradient) map the k -chains into the sets of $(k-1)$ -chains and $(k+1)$ -chains respectively, according to

$$\Delta \sum_i \varphi_i c_k^{(i)} = \sum_{i,j} \varphi_i (c_k^{(i)} : c_{k-1}^{(j)}) c_{k-1}^{(j)} \quad (46)$$

$$\nabla \sum_i \varphi_i c_k^{(i)} = \sum_{i,j} \varphi_i (c_{k+1}^{(j)} : c_k^{(i)}) c_{k+1}^{(j)} \quad (47)$$

The coefficients φ_i of a $(k-1)$ -chain are now considered as the fields of a statistical model. A configuration of this system is thus characterized by a $(k-1)$ -chain. We study the action

$$S \left(\sum_i \varphi_i c_{k-1}^{(i)} \right) = \sum_i \chi \left(\sum_j \varphi_j (c_k^{(i)} : c_{k-1}^{(j)}) \right) + \sum_i V(\varphi_i) \quad (48)$$

The real functions V and χ on the group G describe respectively potential and interaction terms.

The case $k=1$ and $G = Z_2 = \{0, 1\}$ (additive group of integers modulo 2) corresponds to the Ising model. In this unusual notation the interaction term in equation (48) is the sum over all links ij of $\chi(\sigma_i + \sigma_j)$ where the function χ satisfies $\chi(0) = \beta$, $\chi(1) = -\beta$. Verify that gauge models correspond to $k=2$.

The exponentiated action is expanded on group characters, which form a basis for the functions defined on an abelian group. The partition function reads

$$Z = \|G\|^{-n_{k-1}} \sum_{(k-1)\text{-chains}} \left\{ \prod_i \sum_r \beta(r) \chi_r \left(\sum_j \varphi_j (c_k^{(i)} : c_{k-1}^{(j)}) \right) \right\} \left\{ \prod_i \sum_r \gamma(r) \chi_r(\varphi_i) \right\} \quad (49)$$

where the Fourier coefficients are

$$\begin{cases} \beta(r) = \|G\|^{-1} \sum_{\varphi \in G} \chi_r(\varphi) \exp \chi(\varphi) \\ \gamma(r) = \|G\|^{-1} \sum_{\varphi \in G} \chi_r(\varphi) \exp V(\varphi) \end{cases} \quad (50)$$

In these formulae, the number of $(k-1)$ -cells is denoted n_{k-1} and $\|G\|$ is the number of group elements. An arbitrary term in the expansion of the products enclosed within braces in (49) can be obtained by choosing a representation r for any index i . Hence this term can be considered as a chain with values in the *dual group* G^* , i.e. the set of irreducible representations of G . The dual G^* of an abelian group G has also the structure of an abelian group,

with its group law defined as the product of representations of G . Such a group structure does not exist in the nonabelian case and this prevents a straightforward generalization of duality. The terms enclosed by the first braces in (49) are k -chains $g = \sum r_i c_k^{(i)}$ with values in G^* , and those enclosed by the second braces are $(k-1)$ -chains $h = \sum -s_j c_{k-1}^{(j)}$. If we group all terms depending on a given field φ_j , we obtain

$$\left[\prod_i \chi_{r_i} \left(\varphi_j (c_k^{(i)} : c_{k-1}^{(j)}) \right) \right] \chi_{-s_j} (\varphi_j) = \chi_{-s_j + \sum_i r_i (c_k^{(i)} : c_{k-1}^{(j)})} (\varphi_j)$$

The summation over φ_j (i.e. over $(k-1)$ -chains in equation (49)) can now be performed. The result vanishes unless the character index is 0 (0 denoting the trivial representation), and gives a factor $\|G\|$ in this case. This representation index is nothing but the component corresponding to the cell $c_{k-1}^{(j)}$ in the $(k-1)$ -chain $-h + \Delta g$ with values in G^* (also written additively). Hence the partition function (49) takes the form

$$Z = \sum_{\substack{k\text{-chains} \\ \text{on } G^*}} \exp \left\{ \sum_i \ln \beta(r_i) + \sum_i \ln \gamma \left(\sum_j r_j (c_k^{(j)} : c_{k-1}^{(i)}) \right) \right\} \quad (51)$$

Comparing the argument of the exponential to the original action (48), one observes a close analogy. The fields r_i are associated to cells with dimension k (instead of $k-1$). They interact if their supporting cells have a common face (boundary interaction), while the fields in (48) interact if they bound the same cell (coboundary interaction). This difference is inessential and disappears if formula (51) is interpreted on the *dual complex*.

To define this new complex, take the same cells and just change their dimensions from k to $d-k$. The incidence numbers are unchanged. Let us denote by c_{d-k}^* the cell dual to c_k obtained using this substitution. As $(c_k : c_{k-1}) = (c_{d-k+1}^* : c_{d-k}^*)$, the characteristic properties for a *dual complex* are fulfilled.

Geometrical interpretation. On a hypercubic lattice, the dual sites are identified with the centers of the hypercubes of the original lattice. A link joining two neighbouring dual sites is dual to the face separating the two corresponding hypercubes, and so on. Up

to boundary effects, the dual lattice is also a hypercubic lattice, shifted with respect to the original one. Analogously the two-dimensional triangular and hexagonal lattices are dual to each other. Similarly, the face centered cubic and body centered cubic lattices are also dual in three dimensions.

Using this interpretation, the partition function (51) is rewritten on the dual lattice with an action

$$S^* \left(\sum_i r_i c_{d-k}^{(i)*} \right) = \sum_i \ln \gamma \left(\sum_j r_j (c_{d-k+1}^{(i)*} : c_{d-k}^{(j)*}) \right) + \sum_i \ln \beta(r_i) + n_{d-k}^* \ln \|G^*\| \quad (52)$$

Up to an inessential additive term, this action is exactly of the original type (48).

Let us proceed further in the case of a vanishing external field ($V \equiv 0$). This means that $\gamma(r)$ is a δ -function on the dual group, forcing its argument r to be the trivial representation 0 of the group G . Hence, the chain h is 0, which means that the chain $g = \sum r_i c_{d-k}^{(i)*}$, interpreted on the dual complex, has a zero coboundary $\nabla g = 0$. Any coboundary of a $(d-k-1)$ -chain is a solution to this equation, since the operators $\nabla \nabla$ and $\Delta \Delta$ identically vanish. Other solutions might, however, exist depending on the topology of the complex. For instance, a circle drawn on a torus is not necessarily the boundary of a surface on this torus. The homology (cohomology) group of order k is defined as the quotient of the image by Δ (∇) of the $(k+1)$ - ($(k-1)$ -) chains by the corresponding kernel. Saying that the cohomology group of order $(d-k)$ is trivial means that $g = \nabla g'$ is the general solution of $\nabla g = 0$.

Show that this is the case for the complex formed by a hypercube and its faces. Verify that the (co)homology groups are stable for any subdivision of the cells and deduce that any lattice with free boundary conditions has trivial (co)homology groups.

The chain g' is defined up to the coboundary of a $(d-k-2)$ -chain. From these considerations, the expression of the dual action

(in the absence of external fields) can be recast in the form

$$S^* \left(\sum_i r_i c_{d-k-1}^{(i)*} \right) = \sum_i \ln \beta \left(\sum_j r_j (c_{d-k}^{(i)*} : c_{d-k-1}^{(j)*}) \right) + n_{k-1} \ln \|G\| - \ln x_{k+1} \quad (53)$$

where x_{k+1} is the total number of $(k+1)$ -chains without boundary (or, equivalently, the number of dual $(d-k-1)$ -chains without coboundary). This number can be extracted from the following recurrence relations

$$x_k = \|G\|^{n_{k+1}} / x_{k+1}, \quad x_d = \|G\| \quad (54)$$

The dual action (53) is of a form similar to the initial one.

Check that in the case $k=1$, $d=2$, $G=Z_2$, one recovers the duality relations for the two-dimensional Ising model.

As an example, a model without an external field defined on cells of dimension $k=d$ is dual to a model with variables defined on sites, interacting with a constant external field. There are no collective effects, and we recover here the triviality of one-dimensional spin models and two-dimensional gauge models.

The four-dimensional Z_2 gauge theory is self-dual. It corresponds to the case $d=4$, $k=2$ without an external field, with the further property that Z_2 is self-dual. Hence the free energy per site satisfies the relation

$$F(\beta) = F(\beta^* = -\frac{1}{2} \ln \tanh \beta) + 3 \ln \sinh 2\beta \quad (55)$$

High and low temperature regions are related by duality, and formula (55) is similar to the corresponding one for the two-dimensional Ising model. If there is one, and only one, transition, it should occur at the fixed point of the transformation $\beta \rightarrow \beta^*$, i.e. at

$$\beta_c = \frac{1}{2} \ln(1 + \sqrt{2}) \quad (56)$$

Numerical simulations of this system confirm this result, the observed transition being here of first order.

Finally, let us consider the three-dimensional Z_2 gauge system with bosonic matter (Higgs) fields, the action of which reads

$$(U_{ij} = \pm 1, \sigma_i = \pm 1)$$

$$S = \beta \sum_p U_{ij} U_{jk} U_{kl} U_{li} + \gamma \sum_{(i,j)} \sigma_i U_{ij} \sigma_j \quad (57)$$

For a given matter field configuration, we choose the *unitary* gauge corresponding to the transformation $U'_{ij} = \sigma_i U_{ij} \sigma_j$. With this choice, all reference to the matter fields disappears and the model described by equation (57) can be identified to the case $k=2$ of the general expression (48). In three dimensions, the dual action reads

$$S^* = -\frac{1}{2} \ln \tanh \gamma \sum_p U_{ij}^* U_{jk}^* U_{kl}^* U_{li}^* - \frac{1}{2} \ln \tanh \beta \sum_{(i,j)} \sigma_i^* U_{ij}^* \sigma_j^* \quad (58)$$

and the system is again self-dual. The exchange of gauge and matter couplings should be noticed, as well as the interchange of high and low temperature regions. In the convenient variables $\tilde{\beta} = \ln \sinh 2\beta$ and $\tilde{\gamma} = \ln \sinh 2\gamma$, the phase diagram is symmetric with respect to the line $\tilde{\beta} = \tilde{\gamma}$. In the particular case $\gamma=0$, the three-dimensional pure gauge model is dual to the Ising model. It therefore admits a *second order* transition, and the numerical results for the Ising model allow to locate it at $\beta_c \approx 0.7613$.

(1) One can establish directly this equivalence between the two systems using a reasoning similar to the one presented in chapter 2. One starts from the high temperature diagrammatic expansion; graphs have the topology of closed surfaces made with p plaquettes and yield a contribution $\tanh^p \beta$. These surfaces can be characterized by the set of cubes filling their interior with variables σ_i assigned to the center of each cube: $\sigma_i = -1$ inside the surface, $+1$ outside. A plaquette separating two adjacent cubes i and j does or does not belong to the surface according to the value of $\sigma_i \sigma_j$. Hence it is possible to re-express the sum over all high temperature diagrams as the partition function of an Ising model.

(2) Apply duality to a cyclic group Z_n . This model depends on $\lfloor \frac{1}{2}n \rfloor$ parameters and duality yields a relation between these variables, leading to a symmetry of the phase diagram. In particular, deduce the position of the transition $\beta_c = \frac{2}{3} \ln(1 + \sqrt{3})$ of self-dual models with Z_3 symmetry. Three phases arise in systems with $n \geq 5$, and it is no more possible to find the exact location of the transitions. Duality only relates their respective positions.

(3) Consider the $U(1)$ case. The dual group G^* is the additive group of integers. The corresponding duality has been exploited in chapter 4 for the XY-model.

(4) What is the observable dual to the Wilson loop, and, more generally, to any other observable in an abelian gauge system? Hint: the main difference in the diagrammatic expansion arising from the expression (51) consists in changing the representations r_i of the plaquettes on a given arbitrary surface with boundary C to $r_i \otimes f$ (f being the fundamental representation used to define the Wilson operator). Verify that the dual observable appears as the ratio of a frustrated partition function where some coupling constants have been modified, to the usual partition function. For instance, in three dimensions, for the Z_2 gauge model, the coupling of all links crossing a surface bounded by C is reversed. The high temperature diagrams of the dual Ising model should interlace the curve C in order to have different contributions in the two partition functions. Deduce from this fact that the Wilson loop parameter has a perimeter law at low temperature of the gauge model.

't Hooft has introduced an interesting gauge invariant observable for an arbitrary gauge group provided it has a nontrivial center. This includes of course nonabelian systems. Let us describe this observable in a four dimensional model (the case of interest) leaving the reader to give its generalization. Take a closed loop C^* on the dual lattice bounding a surface Σ^* . The 't Hooft observable is defined as the ratio of a frustrated partition function to an unfrustrated one. The frustration amounts to multiply all plaquette variables dual to those of Σ^* by the same element of the center. In general, this loop observable has a behaviour dual to the Wilson loop (i.e. a perimeter law at large coupling and an area law at small coupling). There are interesting cases, such as the Z_n models mentioned above, where both observables have a perimeter law in an intermediate massless phase. Wilson and 't Hooft loops have nontrivial relative topological properties, analogous to order and disorder variables.

6.2 Structure of the phase diagram

6.2.1 Mean field approximation

In spite of several *a priori* difficulties, such as Elitzur's theorem, the mean field method, as described in chapter 3, allows a first exploration of the phase diagram for a pure gauge system. A trial

action is written as

$$S_H = \frac{H}{n} \sum_{(ij)} \text{Re Tr } U_{ij} \quad (59)$$

The corresponding partition function per link is

$$z(H) = \exp u(H) = \int \exp \left(\frac{H}{n} \text{Re Tr } U \right) dU \quad (60)$$

and hence each field has a nonvanishing average. In matrix notation, the mean value is proportional to the unit matrix

$$\langle U \rangle_H = z(H)^{-1} \int U \exp \left(\frac{H}{n} \text{Re Tr } U \right) dU = u'(H) \quad (61)$$

With such a gauge noninvariant trial action, it is clear that the conditions for the validity of Elitzur's theorem are not fulfilled. This defect will be cured later. It is nevertheless interesting to use this trial action to derive inequalities, since the results obtained in this way for invariant quantities are correct, as we shall see. The calculation scheme is the same as in chapter 3. The inequality (3.14) leads to

$$\frac{\ln Z}{Nd} \geq \sup_H \{ u(H) - H u'(H) + \frac{1}{2} \beta_J (d-1) u'(H)^4 \} \quad (62)$$

The right-hand side is the *mean field approximation* for the free energy. The function $u(H)$, defined through the integral (60) of a convex function (an exponential) over a compact domain, is also a convex function, which is minimal at $H = 0$ where it behaves as H^2 . Hence the maximization of the r.h.s. of (62) leads to the solution $H = 0$, which is unique at $\beta_J = 0$ and in the vicinity of this point. In contradistinction with spin systems, this maximum always remains a local maximum, whatever the temperature. Indeed the term $u'(H)^4$ behaves as H^4 , and therefore cannot change the curvature at $H = 0$. Nevertheless, this maximum does not remain the highest one as β_J increases. The behaviour of the function to be maximized is displayed in figure 4 for several values of $\beta^* = 2\beta_J(d-1)$. For a critical value $\beta^* = \beta_c^*$, the highest maximum suddenly changes its location. Hence we observe a first order transition, with discontinuities in the values of some physical observables. The corresponding characteristic shape of the free energy is displayed in figure 5. The dotted curve corresponds to the (unphysical) position of the minimum. Dashed parts display

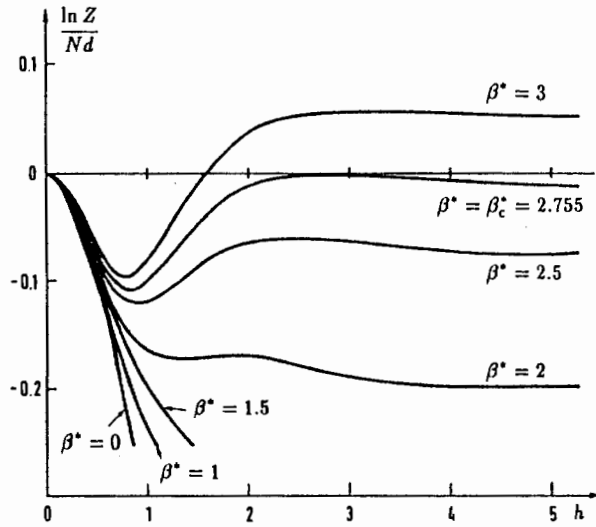


Fig. 4 Trial free energies for a Z_2 gauge group, to be maximized in H for different values of $\beta^* = 2\beta_1(d-1)$.

the position of secondary maxima, which are physically observable as metastable phases. Correlation lengths remain finite at the critical point T . Physical quantities admit analytic continuation at this point which represents metastable states. Therefore, there is *no continuous limit* in the present approximation. The end points A (located at infinity) and B of the metastable regions correspond to singularities of the observables in the variable β_j . The transition at T is a deconfining one. The Wilson loop observable, corresponding to a curve with a perimeter P , has a mean value

$$W(C) = \left\langle \text{Tr} \prod_c U_{ij} \right\rangle = n[u'(H)]^P \quad (63)$$

In the low temperature region ($H \neq 0$), it decreases according to a perimeter law, which shows the deconfined character of this phase. The area law of the confined high temperature region cannot

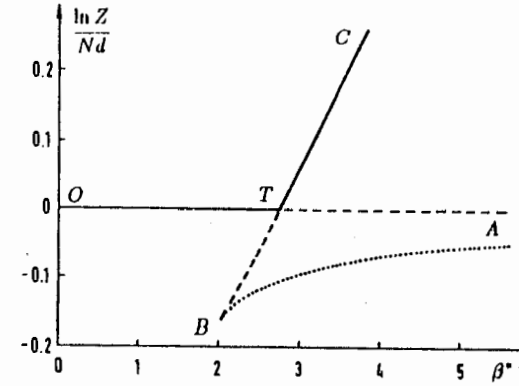


Fig. 5 The free energy in the mean field approximation (Z_2 group).

be obtained in this approximation without further corrections ($W(C) = 0$ in this region in the present context).

The simple pattern given by the mean field approximation is confirmed by numerical simulations. Let us discuss first the Z_2 case. The observed transition is indeed a first order one as soon as the dimension is greater or equal to four. Its location is reasonably predicted. In dimension four, mean field yields $\beta_c = 0.459$ to be compared with the position $\beta_c = 0.4407$, exactly known from duality. In dimension three, although the predicted order of the transition is incorrect – it should be a second order one, as the system is dual to the Ising model –, there is still some similarity between the position $\beta_c = 0.7613$ and the mean field prediction $\beta_c = 0.689$.

The upper critical dimension for nonabelian continuous gauge systems is expected to be $d = 4$, where the theory becomes asymptotically free. In higher dimensions, mean field predictions are confirmed by numerical simulations. For instance, the first order transition of the $SU(2)$ system for $d = 5$ is predicted to be at $\beta_c = 2.12$ to leading order and has been observed at $\beta_c = 1.64$. At $d = 4$, there is no evidence for any transition in the numerical simulations. However, observable quantities show rapid changes in a small region. If one explains this fact by a remembrance of the transition occurring in larger dimensions, one observes again

some similarity between the observed and predicted positions (for $SU(2)$ e.g., 2.826 and 2.2 respectively). All these facts seem to indicate that this approximation provides a reasonable description of the system and is therefore a starting point for a systematic computation of corrections.

(1) Perform the same calculations using the axial gauge. The result reads

$$\frac{\ln Z}{N(d-1)} \geq \sup_H \left\{ u(H) - Hu'(H) + \frac{1}{2}\beta_J(d-2)u'(H)^4 + \beta_J u'(H)^2 \right\} \quad (64)$$

The last term in this formula might now change the sign of the curvature at $H = 0$. In sufficiently high dimension, it is too weak to produce such a change before the first order transition occurs. Its effect is just to bring back the end point A of the metastable strong coupling phase at a finite distance

$$\beta_A = [2u''(0)]^{-1} \quad (65)$$

As the dimension decreases, the swallowtail shape $TABT$ shrinks and disappears at a critical dimension

$$d_c = 2 - \frac{1}{6} \frac{u'''(0)}{u''(0)^3} \quad (66)$$

This choice of gauge seems therefore to be of interest in the implementation of the mean field approximation. Moreover, having made such a choice of gauge, the use of a noninvariant trial action (59) no longer presents theoretical problems. We will also see that, for continuous groups, one cannot escape an explicit choice of gauge for a determination of the corrections. Nevertheless, other problems appear. For instance, let us consider a Wilson loop of dimensions $R \times T$ using the time direction along which the gauge field has been set equal to unity. The approximation leads to a behaviour in R rather than $R + T$ since $W(C) = n[u'(H)]^{2R}$. This result prevents in particular an exponential decrease for fixed R as $T \rightarrow \infty$ which can be predicted from rigorous bounds.

This difficulty disappears when corrections are added in the case of a continuous gauge group, but remains when dealing with discrete groups. The mean field approximation appears in this case as a dilute gas approximation, each gauge field fluctuating independently in the mean field without interacting with its neighbours. This technique can be seen as a partial resummation of the low temperature expansion. The flip of a gauge field in a temporal direction is replaced, in the axial gauge, by the flips

of an infinity of spatial fields (figure 6), but nevertheless leads to a finite energy change. The approximation does not take into account this phenomenon, and finite order corrections will not change this feature. In particular, the reader can check that the low temperature behaviour is incorrect. Using the above observation, it is now possible to correct this omission by adding the contribution of a dilute gas of such defects. The correct low temperature behaviour and the perimeter law for temporal Wilson loops are restored. However, the computations of corrections is much more complex, and one generally prefers to use the method which is described below and which does not require the gauge to be fixed.

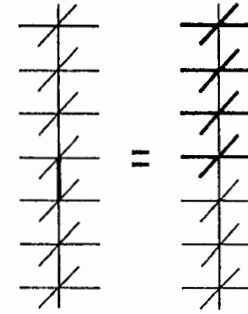


Fig. 6 The equivalent of flipping a temporal gauge field in the axial gauge.

(2) Using the mean field approximation, study the generalized action for a $SU(n)$ gauge group

$$S = \sum_p \left(\beta_J \frac{1}{n} \text{Re Tr } U_p + \beta_A \frac{1}{n^2 - 1} |\text{Tr } U_p|^2 \right) \quad (67)$$

This extension is interesting, because it allows to probe the role of the center of the group. The additional term is invariant under transformations by an element belonging to the center Z_n : $U_{ij} \rightarrow \exp(2ik\pi/n)U_{ij}$, and is the fundamental Wilson action for the group $SU(n)/Z_n$. The two-dimensional phase diagram provides an interpolation between several limits

$\beta_A = 0$ Wilson action for $SU(n)$.

$\beta_J = 0$ Wilson action for $SU(n)/Z_n$.

$\beta_A = \infty$ action for the discrete gauge group Z_n , since this limit selects the configurations in which all plaquette variables U_p belong to the center. Hence, using a gauge transformation, all gauge fields U_{ij} may be chosen in the center Z_n .

$\beta_J = \infty$ forces all plaquette variables to be equal to unity and yields a trivial model, with only pure gauge configurations and without any transition.

Some care in the elaboration of the mean field approximation is needed in order to treat central elements correctly. We write $U_{ij} = V_{ij}\epsilon_{ij}$, and treat independently V_{ij} and ϵ_{ij} with two associated mean fields H and h . One obtains a phase diagram which is displayed in figure 7, with three different regions

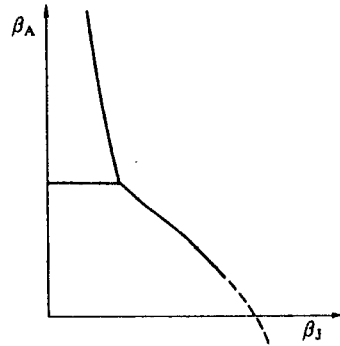


Fig. 7 Phase diagram predicted by the mean field approximation for extended theories with the action (67).

- a) strong coupling phase for β_J, β_A small ($H = 0, h = 0$),
- b) weak coupling phase of $SU(n)/Z_n$ ($H = 0, h \neq 0$) obtained as β_A is large and β_J is small (fields fluctuate around central values),
- c) weak coupling phase ($H \neq 0, h \neq 0$).

All transition lines are of first order. Numerical simulations confirm this prediction. However, for $d = 4$, a part of the line (dashed in the figure) disappears. Phases a and c are continuously related, while the weak coupling region $SU(n)/Z_n$ remains isolated.

The end point is not very far from the axis $\beta_A = 0$. The crossover zone observed in $SU(n)$ systems ($n = 2, 3$) in which the properties of the system vary rapidly, can be explained by the presence of this neighbouring singular point. Note also that the model with $SU(n)/Z_n$ symmetry has no order parameter of the Wilson type (since this simple group is centerless). Nevertheless it exhibits a first order transition, even at $d = 4$.

6.2.2 Corrections to mean field and restoration of gauge invariance

We study now the corrections due to the fluctuations around mean field. This is done using the technique developed in chapter 3, which uses a Laplace transform of the measure on the fields. However, in the present case, it is not possible to use the identity (3.142). The calculation can nevertheless be performed, using a double steepest descent method that we now describe. The exponential representation of the δ -function

$$\delta(U_{\alpha\beta} - V_{\alpha\beta}) = \frac{1}{2i\pi} \int_{-i\infty}^{i\infty} \exp H_{\beta\alpha}^* (U_{\alpha\beta} - V_{\alpha\beta}) dH_{\beta\alpha} \quad (68)$$

is inserted in the definition of the partition function for every matrix element $(U_{ij})_{\alpha\beta}$ of the gauge fields. Here V_{ij} and H_{ij} are $n \times n$ complex matrices on which we integrate, using flat measures

$$[dV] = \prod_{(ij)\alpha\beta} d(V_{ij})_{\alpha\beta} \quad \text{and} \quad [dH] = \prod_{(ij)\alpha\beta} d(H_{ij})_{\alpha\beta}$$

not to be confused with the Haar measure dU on the group (which appears as a submanifold of these $n \times n$ matrices). For each link (ij) , only one pair (H, V) of these matrices is introduced, regardless of the link orientation, according to $V_{ji} \equiv V_{ij}^\dagger, H_{ji} \equiv H_{ij}^\dagger$. The partition function reads

$$Z(J) = \int \exp \left\{ S(V_{ij}) + \sum_{(ij)} \text{Re Tr } J_{ji}^\dagger U_{ij} \right\} \prod_{ij} \delta^{(n^2)}(U_{ij} - V_{ij}) [dV] \prod_{ij} dU_{ij}$$

$$= \int [dH][dV] \prod_{ij} dU_{ij} \exp \left\{ S(V_{ij}) + \sum_{ij} \left(\text{Re Tr}(J_{ji} + H_{ji})^\dagger U_{ij} - \text{Re Tr} H_{ji}^\dagger V_{ij} \right) \right\} \quad (69)$$

Generalizing the function (60) to a matrix external field

$$u(H) = \int \exp \left(\frac{1}{n} \text{Re Tr} H^\dagger U \right) dU \quad (70)$$

the integration over gauge fields can be performed, and one gets

$$Z(J) = \int \exp \left\{ S(V_{ij}) + \sum_{ij} u(H_{ij} + J_{ij}) - \text{Re Tr} H_{ij}^\dagger V_{ji} \right\} [dV][dH] \quad (71)$$

This integral is now computed using the steepest descent method. For spin systems, the integration over V is Gaussian and can be done explicitly, without further corrections. This is no longer true in the present case. To leading order, we find

$$\ln Z = \sum_{ij} \left[u(\mathcal{H}_{ij} + J_{ij}) - \text{Re Tr} \mathcal{H}_{ij}^\dagger \mathcal{V}_{ij} \right] + S(\mathcal{V}_{ij}) \quad (72)$$

This expression is stationary for $H = \mathcal{H}$, $V = \mathcal{V}$ given by

$$\mathcal{V}_{ij} = \frac{\partial u(\mathcal{H}_{ij} + J_{ij})}{\partial H_{ij}^\dagger} \quad (73)$$

$$\mathcal{H}_{ij} = \frac{\partial S(\mathcal{V}_{ij})}{\partial V_{ji}} \quad (74)$$

In the particular case of a uniform solution with constant matrices proportional to unity, these expressions can be identified with those in equation (62), so that

$$\mathcal{H}_{ij} = \frac{H}{n} I \quad (75)$$

We are now in a position to compute systematic corrections to this approximation. But first, let us solve the paradoxes concerning gauge invariance. Saddle point equations (73)–(74), although gauge invariant, may have noninvariant solutions $\{\mathcal{V}_{ij}, \mathcal{H}_{ij}\}$; but then any gauge transformed configuration $\{g_i \mathcal{V}_{ij} g_j^{-1}, g_i \mathcal{H}_{ij} g_j^{-1}\}$ is also a solution. When computing the integral (71), one should

separate the degrees of freedom corresponding to gauge transformations from other “transverse” degrees of freedom. The former yield a multiplicative factor (to a power equal to the number of sites) since the integrand is independent from these transformations. The steepest descent method can only be applied to the latter. Note the transformation law of the sources $J_{ij} \rightarrow g_i J_{ij} g_j^{-1}$, which induces corresponding transformations on the observables. Starting from the solution (75), a set of other saddle points can thus be constructed by gauge transformations. For instance, the value $\langle U_{ij} \rangle = u'(H)$ for the solution (75) becomes $g_i \langle U_{ij} \rangle g_j^{-1}$. Integrating now on gauge degrees of freedom gives

$$\langle U_{ij} \rangle = \int g_i u'(H) g_j^{-1} \prod dg_i \equiv 0 \quad (76)$$

As a consequence, Elitzur’s theorem is satisfied. We also see that the mean field approximation as presented in the previous subsection yields correct results for gauge invariant quantities. The virtue of the present formalism is to allow the systematic computation of corrections.

In the case of a discrete group, the saddle points generated by gauge transformations are well separated, and there are no zero modes. The problem is to know whether the contributions of the saddle points should be *summed* or *averaged*. The question makes sense when the number of these saddle points can vary. The answer is that they should be summed over. Let us prove this for the example of an integral in one variable $\int \exp f(x) dx$, where $f(x)$ is analytic with several maxima (figure 8). It is convenient to introduce the variable $y = -f(x)$. The integral becomes $\oint_C e^{-y} f'(x)^{-1} dy$ along a curve C in the complex y plane displayed in figure 8b. It surrounds the singularities of the multivalued function $x(y)$. Since we deal with analytic functions, it is possible to distort the contour into two independent contours, as shown in figure 8c. The contribution of each one is evaluated using a Taylor expansion around the singular end point, and is found to coincide with the saddle point contribution. This proves the summation prescription over all equivalent saddle point contributions, as long as the analyticity domain of the function is sufficient to allow the distortion of contours. If the two maxima have different heights, the higher one dominates exponentially over the lower. The main effect of the minimum is to limit the radius of convergence of the Taylor

series of $1/f'(x(y))$. Hence, after integration, the series obtained by the steepest descent method are generally not convergent, but only asymptotic.

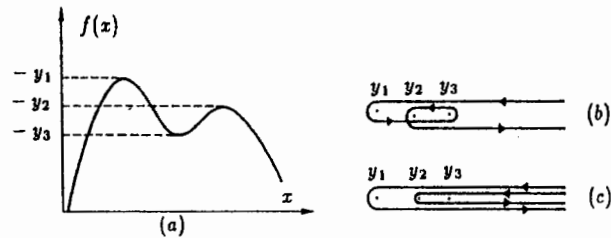


Fig. 8 Various integration contours (b) and their deformations (c) when using the steepest descent method for a function with several extrema (a).

In the application of this method to gauge systems, one deals with integrals in several variables. The analyticity properties required to distort the contours are harder to establish. Nevertheless, by analogy, we presume that the summation prescription is the correct one.

In particular, let us discuss the validity of the method near the critical dimension where a first order transition turns into a second order one. The transition point and the two end points of the metastable zones coalesce. Near the end point of the strong coupling region, the unique saddle point splits into gauge equivalent saddle points (2^N for Z_2). In this case, the minimum separating two neighbouring saddle points is very close, and the convergence of the method will be particularly poor. Note the additive factor $N \ln 2$ which appears in the expression for the free energy in the weak coupling phase, which is the effect of the summation prescription and is required to obtain the zero coupling entropy.

6.2.3 Discrete groups: $1/d$ expansion

It is expected that the mean field approximation gives good results in high dimension. For spin models, we have seen that corrections can be rearranged as a $1/d$ expansion. Let us examine now the case of gauge systems.

We will show, in chapter 7 (volume 2), that diagrammatic series can be ordered into successive corrections to the mean field



Fig. 9 A typical diagram occurring in the computation of corrections to the mean field approximation.

approximation, and we will describe the diagrams which are to be computed. The reader will find there all the technical details, and we only give here a quick description of the method. Looking at formula (71), one remarks that each term generated in the expansion of $\exp S(V_{ij})$ can be described as a set of plaquettes. The term $\text{Re Tr } H_{ij}^\dagger V_{ji}$ couples the fields V lying on the boundaries of these plaquettes to $u(H_{ij})$. Hence, if k plaquettes meet along a link ij , one gets a contribution $\partial^k u(H_{ij}) / \partial H_{ij}^k$. The reader can also verify – this will be proved in chapter 7 – that *reducible* diagrams, i.e. those which can be disconnected by the omission of one link, should not be considered. Indeed, the self-consistent field H_{ij} is such that reducible contributions are automatically generated by formula (74).

In figure 9 we show a typical diagram in which each plaquette is represented by a black point located at its center and joined to the four midpoints of adjacent links (open dots). The irreducibility condition forbids the possibility of arborescences, the resulting diagram being strongly connected. This is the case displayed in figure 9, showing the general aspect of a one-loop diagram.

Let us construct such a loop with n plaquettes, that is with $2n$ segments of length $a/2$. As this loop should be closed, these segments can be associated in parallel pairs of opposite direction. Therefore, the loop uses at most n different directions chosen among the d possible ones. The choice of these directions yields a factor $d^n (1 + \mathcal{O}(1/d))$ in the contribution of the diagram. Moreover, each plaquette comes with a factor β_j , which is of order $1/d$ (because it appears only in a combination $2\beta_j(d-1)$ in the

mean field approximation, so that this expression is kept fixed in the limit $d \rightarrow \infty$). The loop yields therefore at most a contribution of order $1 + \mathcal{O}(1/d)$, which is subdominant with respect to the leading term (62), of order d . Any additional loop in a more complex diagram will provide additional $1/d$ factors. Indeed, due to the irreducibility condition, it should share at least one link with previously constructed loops, the orientations of which have already been chosen.

Thus the diagrams contributing to the first $1/d$ correction have been characterized. They have only one loop and are all similar to the one drawn in figure 9. Their number is infinite, in contradistinction with spin models. In the latter case, a loop with n links is of order $\beta^n d^{n/2}$, behaving as $(1/d)^{n/2}$ as a function of its size n (for fixed βd). In the gauge case, despite the infinite number of diagrams, it is possible to re-sum their contributions. The diagram of figure 9 with n plaquettes contributes a term

$$\frac{d! [2\beta_j u'(H) u''(H)]^n}{2n(d-n)!} \quad (77)$$

Prove this result.

For large d , these terms have to be summed for $n \geq 3$. The two plaquette term should also be added and the final correction reads

$$\begin{aligned} \frac{\ln Z}{Nd} &= (\text{mean field contribution}) \\ &+ \left[-\frac{1}{2} \ln(1-x) - \frac{1}{2}x + \frac{1}{8}x^2 + \frac{1}{4}xy + \frac{1}{16}y^2 \right] \frac{1}{d} + \mathcal{O}\left(\frac{1}{d^2}\right) \end{aligned} \quad (78)$$

with

$$x = 2\beta_j(d-1)u'^2(H)u''(H) \quad y = 2\beta_j(d-1)u''^2(H) \quad (79)$$

An additional contribution $d^{-1} \ln \|G\|$ which takes into account the multiplicity of the saddle points, as discussed in the previous subsection, should be added when H does not vanish.

Compute the next correction.

Figure 10 displays the free energy per link for the Z_2 gauge group, taking into account the first two corrections. The low

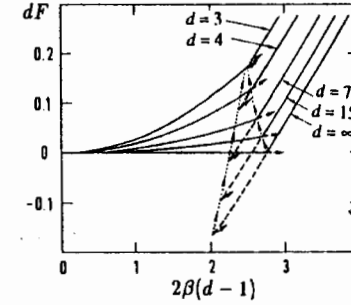


Fig. 10 Free energy of the Z_2 gauge system computed using the first two corrections in $1/d$.

temperature metastable region shrinks as the dimension decreases and disappears for $d_c \simeq 2.9$, very close to $d_c = 3$ which is expected to be the critical dimension for this model. However, the second order transition occurring in dimension three cannot be seriously analyzed by this method. Note also that the end point of the high temperature metastable region remains at infinity. We shall see later that its position behaves as $d^{3/4}$, and is thus out of reach of this $1/d$ expansion technique. From our previous discussion, these features are expected.

6.2.4 Continuous groups: computation of corrections

We have already outlined the necessity of gauge fixing for continuous gauge groups. The axial gauge choice leads to simpler calculations. We treat in detail the Wilson action for a d -dimensional model with $SU(n)$ or $U(n)$ gauge group. We use the same method as above and write the partition function as

$$Z = \int \prod_{\text{spatial links}} dU_l \exp \left\{ \frac{\beta_j}{n} \text{Re} \left[\sum_{\text{spatial plaquettes}} \text{Tr} U_p + \sum_{\text{timelike plaquettes}} \text{Tr} U_l U_l^\dagger \right] \right\} \quad (80)$$

We insert

$$1 = \prod_{\text{spatial links}} \int d^{2n^2} V_l \delta^{(2n^2)}(V_l - U_l) \quad (81)$$

$$= \prod_l \int dV_l dH_l (2i\pi)^{-2n^2} \exp \left\{ \frac{1}{n} \text{Re Tr}(U - V) \right\}$$

It is possible to integrate over the U variables. In the axial gauge, formulae (72)–(74) are slightly modified. To lowest order, the result becomes

$$\frac{\ln Z}{N} = (d-1)[u(\mathcal{H}) - \mathcal{H}\mathcal{V} + \frac{1}{2}\beta_J(d-2)\mathcal{V}^4 + \beta_J\mathcal{V}^2] \quad (82)$$

with the following saddle point conditions

$$\begin{cases} \mathcal{V} = u'(\mathcal{H}) \\ \mathcal{H} = \beta_J \mathcal{V} [2 + 2(d-2)\mathcal{V}^2] \end{cases} \quad (83)$$

The action is then expanded up to quadratic terms around the saddle point. In order to simplify the calculations of the second derivatives needed for the determination of the first correction, we introduce projectors on the trivial and adjoint representations of $SU(n)$

$$\begin{cases} P_{\alpha\beta;\gamma\delta}^{(0)} = \frac{1}{n} \delta_{\alpha\beta} \delta_{\gamma\delta} \\ P_{\alpha\beta;\gamma\delta}^{(1)} = \delta_{\alpha\delta} \delta_{\beta\gamma} - \frac{1}{n} \delta_{\alpha\beta} \delta_{\gamma\delta} \end{cases} \quad (84)$$

The required derivatives can be parametrized as

$$\begin{cases} \left. \frac{\partial^2 u(H)}{\partial H_{\beta\alpha} \partial H_{\delta\gamma}} \right|_{H=\mathcal{H}1} = \left. \frac{\partial^2 u(H)}{\partial H_{\beta\alpha}^* \partial H_{\delta\gamma}^*} \right|_{H=\mathcal{H}1} = A P_{\alpha\beta;\gamma\delta}^{(0)} + B P_{\alpha\beta;\gamma\delta}^{(1)} \\ \left. \frac{\partial^2 u(H)}{\partial H_{\beta\alpha} \partial H_{\delta\gamma}^*} \right|_{H=\mathcal{H}1} = A' P_{\alpha\beta;\gamma\delta}^{(0)} + B' P_{\alpha\beta;\gamma\delta}^{(1)} \end{cases} \quad (85)$$

We split the operators according to their action on the antihermitian and hermitian parts of V . After a Fourier transformation,

one finds respectively the following $(d-1) \times (d-1)$ -matrices

$$\Delta_{\mu\nu}(k) = \left[2 \cos k_d + 2\mathcal{V}^2 \sum_{\rho \neq \mu, d} \cos k_\rho \right] \delta_{\mu\nu} + 4\mathcal{V}^2 (1 - \delta_{\mu\nu}) \exp(\frac{1}{2}i(k_\mu - k_\nu)) \sin \frac{1}{2}k_\mu \sin \frac{1}{2}k_\nu$$

$$\bar{\Delta}_{\mu\nu}(k) = \left[2 \cos k_d + 2\mathcal{V}^2 \sum_{\rho \neq \mu, d} \cos k_\rho \right] \delta_{\mu\nu} + 4\mathcal{V}^2 (1 - \delta_{\mu\nu}) \exp(\frac{1}{2}i(k_\mu - k_\nu)) \cos \frac{1}{2}k_\mu \cos \frac{1}{2}k_\nu \quad (86)$$

Functional integration yields to order one loop

$$Z = Z_{\text{mean field}} \times \left[\det^{-\frac{1}{2}}(1 - 2\beta_J n(A' - A)\Delta) \det^{-\frac{1}{2}(n^2-1)}(1 - 2\beta_J n(B' - B)\Delta) \right]^N$$

$$\times \left[\det^{-\frac{1}{2}}(1 - 2\beta_J n(A' + A)\bar{\Delta}) \det^{-\frac{1}{2}(n^2-1)}(1 - 2\beta_J n(B' + B)\bar{\Delta}) \right]^N \quad (87)$$

Establish this formula and determine the next correction.

Without gauge fixing, the propagator $\Delta_{\mu\nu}$ takes a slightly different form. In both brackets in equation (86), the $\cos k_d$ term is multiplied by $\mathcal{V}^2$, and the indices μ and ν can take the value d . Hence the matrix Δ has a zero eigenvalue, which corresponds to the eigenvector $\sin \frac{1}{2}k_\mu \exp \frac{1}{2}ik_\mu$ and which reflects the existence of local gauge transformations. The mass spectrum can be read on formula (87) as corresponding to the zero eigenvalues of the operators appearing in the determinants.

It is instructive to look at the weak coupling limit (H large). Taking into account the invariance of $u(H)$ with respect to unitary transformations, one finds

$$\begin{cases} B' - B = \frac{u'(\mathcal{H})}{2n\mathcal{H}} = \frac{\mathcal{V}}{2n\mathcal{H}} \\ A' - A = \begin{cases} B' - B & \text{for } U(n) \\ \mathcal{O}(1/H^2) & \text{for } SU(n) \end{cases} \end{cases} \quad (88)$$

Using the asymptotic forms of $u(H)$ for H large

$$u(H) = \begin{cases} H + \left(\ln \frac{\prod_{k=1}^{n-1} k!}{(2\pi)^{\frac{1}{2}n}} + \frac{1}{2}n^2 \ln n \right) - \frac{1}{2}n^2 \ln H \\ \quad + \frac{n^2}{8H} + \frac{n^4}{16H^2} + \dots & \text{for } U(n) \\ H + \left(\ln \frac{\prod_{k=1}^{n-1} k!}{(2\pi)^{\frac{1}{2}(n-1)}} + \frac{1}{2}(n^2 - 2) \ln n \right) - \frac{1}{2}(n^2 - 1) \ln H \\ \quad - \frac{n^2 - 1}{8H} - \frac{(2n^2 - 5)(n^2 - 1)}{48H^2} + \dots & \text{for } SU(n) \end{cases} \quad (89)$$

one finds

$$\mathcal{H} \sim 2\beta_J(d-1) \quad \mathcal{V} \sim 1 - \frac{A_0}{4\beta_J(d-1)} \quad (90)$$

A_0 being the coefficient of $-\ln H$ in the expansions (89). Similarly

$$\begin{cases} B \sim -B' \sim -\frac{1}{4nH} \\ A \sim -A' \sim \begin{cases} -\frac{1}{4nH} & \text{for } U(n) \\ \mathcal{O}(1/H^2) & \text{for } SU(n) \end{cases} \end{cases} \quad (91)$$

Hence

$$\frac{\ln Z}{N} = \frac{\ln Z_{\text{mean field}}}{N} - \frac{1}{2} \left[\frac{n^2}{n^2 - 1} \right] \text{Tr} \ln \left(1 - \frac{\beta_J \mathcal{V}}{\mathcal{H}} \Delta \right) + \mathcal{O}\left(\frac{1}{\beta_J}\right) \quad (92)$$

Between the square brackets, the upper value n^2 refers to the $U(n)$ gauge group, and the lower one $n^2 - 1$ to $SU(n)$. Only one operator remains to be studied. At fixed momentum, this $(d-1) \times (d-1)$ matrix has two distinct eigenvalues. The first one

$$\lambda_1 = 4\beta_J(\mathcal{V}/\mathcal{H}) \sin^2 \frac{1}{2} k_d$$

vanishes at $k_d = 0$, and corresponds to the massless propagation of the gauge degrees of freedom which have not been fixed by the axial gauge choice (these are the time independent gauge transformations). The second one

$$\lambda_2 = 4\beta_J(\mathcal{V}/\mathcal{H}) \sin^2 \frac{1}{2} k_d + \mathcal{V}^2 \sum_{\mu=1}^{d-1} \sin^2 \frac{1}{2} k_\mu$$

is $d-2$ times degenerate and vanishes only at $\mathbf{k} = 0$. Other hermitian channels are related to the operator $\bar{\Delta}$ and have no massless excitations. Indeed, it was expected that in this large β_J limit, the antihermitian channels represent the most important contribution to the corrections, since they generate the gauge potential in the continuous limit.

For large β_J , the energy per plaquette behaves as

$$E = \frac{2}{d(d-1)} \frac{\partial}{\partial \beta_J} \frac{\ln Z}{N} = \frac{1}{d} [(d-2)\mathcal{V}^4 + 2\mathcal{V}^2] \sim 1 - \frac{A_0}{d\beta_J} + \dots \quad (93)$$

which is the expected result. Without gauge fixing, we would have obtained $1 - A_0/\beta_J(d-1)$ and this incorrect result would then have to be corrected to all orders. This phenomenon is general with continuous groups. At any order in $1/\beta_J$, only a finite number of terms contributes to the corrections to the mean field approximation when the gauge has been fixed. This was not the case for the discrete gauge groups, and we have seen that it was important not to fix the gauge in order to avoid defects of infinite order.

Treat the case of $SU(2)$. Some differences appear, because the representations of this group are real (equivalent to their complex conjugate representations). Using the parametrization $H = h_0 - ih \cdot \sigma$ in terms of Pauli matrices, one sees in particular that $\text{Im}h$ plays no role. The integration over this variable yields a δ function for $\text{Im}V$ which decouples to all orders.

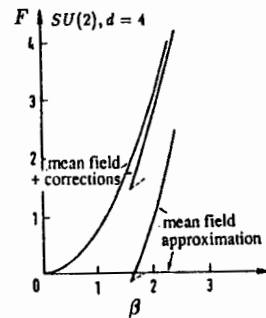
Let us now look at the phase diagram, taking into account the corrections. The free energy of the strong coupling solution ($\mathcal{H} = \mathcal{V} = 0$) is compared to the weak coupling one (approximation (93) is often sufficient for this purpose). Predictions of first order transitions are very close to results from numerical simulations, as can be seen on table I. The most striking results concern $SU(2)$ and $SU(3)$ groups in four dimensions. As shown on figure 11, weak and strong coupling arcs do not meet, but remain parallel and very close to each other in a large domain. It is of course not possible to say that the corrected mean field predicts the absence of a transition, as is suggested in numerical simulations, but this is very encouraging for this technique.

Table I. Corrected mean field predictions versus numerical results for the location of the transition

d	group	mean field	Monte-Carlo
4	$U(1)$	1.03	1.01
	$U(2)$	3.45	3.30 ± 0.05
	$U(3)$	7.30	6.88 ± 0.05
	$U(4)$	12.5	12.14 ± 0.7
	$U(5)$	19.5	18.8 ± 1.1
	$U(6)$	28.0	27.0 ± 2.6
	$U(\infty)$	$0.38 \times 2\pi^2$	$0.375 \times 2\pi^2^*$
	$SU(2)$	†	†
	$SU(3)$	†	†
	$SU(4)$	11.6	10.2
	$SU(5)$	18.0	16.4 ± 0.2
	$SU(6)$	26.5	24.0 ± 1.0
5	$SU(2)$	1.77	1.64

*extrapolation

†unobserved transition.

Fig. 11 Corrections to the mean field approximation for $SU(2)$ in four dimensions, according to H. Flyvbjerg, B. Lautrup and J.-B. Zuber, *Phys. Lett.* 110B, 179 (1982).

6.3 Strong coupling expansions

6.3.1 Convergence

In the strong coupling region, it is natural to expand the Boltzmann weight as a power series in the small parameter β_J

$$\exp S = \prod_p \exp \beta_J \chi(U_p) = \prod_p (1 + \beta_J \chi(U_p) + \frac{1}{2} \beta_J^2 \chi^2(U_p) + \dots) \quad (94)$$

One then further expands the product over the plaquettes, and each term is integrated over gauge fields. It is useful to associate a geometrical representation to each term – which we call a *diagram* or *graph* – as the set of plaquettes used in its contribution. A given plaquette may appear several times, according to the power of $\chi(U_p)$ selected in the expansion of the r.h.s. of (94). In the next subsection, we give a technique for resumming the contributions of plaquettes lying at the same location. This method makes full use of the structure of the gauge group and leads to easier integrations over gauge fields.

A given link cannot bound one and only one plaquette of the diagram since the integration over the corresponding gauge field leads to a zero contribution. Hence links are shared by at least two plaquettes. In particular, if one only considers diagrams such that any link bounds exactly two plaquettes, such graphs have the topology of simple closed surfaces made with plaquettes. A more complex diagram has also singular lines along which several open simple surfaces share their boundaries.

The general techniques needed for determining these diagrammatic series are described in detail in chapter 7 (volume 2). We only mention in this section the particular features of gauge theories and analyze the corresponding results. First, we want to examine the convergence of the expansions. The method consists of establishing an upper bound for the term of order n in the strong coupling expansion for the average $\langle X \rangle$ of a local observable X . Local means here that only a finite number of gauge fields appear in the definition of X . Let $\mathcal{L}$ be the finite, non-empty set of corresponding links. The convergence theorem will also be valid for the free energy per link, since its first derivative is the internal energy $E = \langle \chi(U_p) \rangle$ which is local.

We assume that the plaquette action $\chi(U_p)$ is bounded; this is the case for discrete or compact Lie groups, on which we focus our

attention. We also assume that it is positive. This is not really a restriction, since this condition can be fulfilled by adding an inessential constant to the action. Hence, in some strong coupling region, the following inequality holds

$$0 \leq \Omega(U_p) \equiv \exp \beta_j \chi(U_p) - 1 \leq C_1 \beta_j \quad (95)$$

The lattice Λ is finite, but arbitrarily large. We take the thermodynamic limit at the end of the reasoning, after obtaining a radius of convergence independent of the lattice size. Using these conventions, the average of the observable reads

$$\langle X \rangle = Z_\Lambda^{-1} \int \prod_l dU_l \prod_{p \in \Lambda} [1 + \Omega(U_p)] X \quad (96)$$

with

$$Z_\Lambda = \int \prod_l dU_l \prod_{p \in \Lambda} [1 + \Omega(U_p)] \quad (97)$$

Let us expand the product in the integrand (96). Each term is interpreted as a diagram (set of plaquettes) $\mathcal{D}$ which can be split into two parts. The first, $\mathcal{D}_1$, is the union of all connected parts of $\mathcal{D}$ such that at least one link of $\mathcal{L}$ is adjacent to one plaquette of $\mathcal{D}_1$. The second is the complementary set $\mathcal{D}_2 = \mathcal{D} \setminus \mathcal{D}_1$. The integral (96) is factorized into the product of the two corresponding contributions. If $\mathcal{D}_1$ is fixed, all corresponding diagrams $\mathcal{D}$ are obtained by adding any diagram of the partition function of the lattice $\Lambda \setminus \overline{\mathcal{D}_1 \cup \mathcal{L}}$. This lattice is obtained from Λ by suppressing the set of plaquettes, (abusively) denoted as $\overline{\mathcal{D}_1 \cup \mathcal{L}}$, which are adjacent to a link of $\mathcal{L}$ or to a link bounding a plaquette of $\mathcal{D}_1$. Hence

$$\langle X \rangle = \sum_{\mathcal{D}_1} \left\{ \int \prod_{l \in \mathcal{D}_1 \cup \mathcal{L}} dU_l \prod_{p \in \mathcal{D}_1} \Omega(U_p) X \right\} Z_{\Lambda \setminus \overline{\mathcal{D}_1 \cup \mathcal{L}}} Z_\Lambda^{-1} \quad (98)$$

It is now possible to find an upper bound for the various terms involved in this formula. As Ω is positive and since there are more plaquettes in Λ than in $\Lambda \setminus \overline{\mathcal{D}_1 \cup \mathcal{L}}$, the ratio $Z_{\Lambda \setminus \overline{\mathcal{D}_1 \cup \mathcal{L}}} Z_\Lambda^{-1}$ is less than unity. The function X is local and depends on a finite number of gauge fields varying in a compact or discrete group; its absolute value is therefore bounded by a constant C_2 . Finally, the inequality (95) can be used to majorize the products of Ω over the

plaquettes of $\mathcal{D}_1$, and one gets

$$|\langle X \rangle| \leq \sum_k n(k) C_2 (C_1 \beta_j)^k \quad (99)$$

where k is the number of plaquettes of $\mathcal{D}_1$ and $n(k)$ is the number of diagrams $\mathcal{D}_1$ with k plaquettes. We only need now an estimate of $n(k)$. Let us define an iterative process, adding one plaquette at each step. At a given stage, we have an incomplete diagram with a set of unsaturated links onto which new plaquettes will be tied in the next steps. At the beginning, the set of unsaturated links is $\mathcal{L}$. Let us number the links of Λ arbitrarily; at each step, we decide to add the plaquette along the unsaturated link of lowest rank. There are at most $2(d-1)$ possibilities to choose this plaquette. Next, we have to see whether the four links bounding the added plaquette are unsaturated or not. This yields at most $2^4 = 16$ possibilities. To order k , the diagrams are those for which the set of unsaturated links is empty. Hence we get a generous upper bound

$$n(k) \leq [32(d-1)]^k \quad (100)$$

This bound ensures that the r.h.s. of (99) is a convergent series provided $\beta_j < [32C_1(d-1)]^{-1}$. This completes the proof that the strong coupling expansion for $\langle X \rangle$ has a nonvanishing radius of convergence.

This upper bound is very crude, although it is sufficient for our purpose. Establish more restrictive bounds.

This convergence theorem has two important consequences. First, correlation functions behave exponentially. Let $X^{(\mathbf{x})}$ be the observable X translated by a vector $\mathbf{x}$ on the lattice. One gets the following bound on the correlation function

$$\left| \langle X X^{(\mathbf{x})} \rangle_c \right| = \left| \langle X X^{(\mathbf{x})} \rangle - \langle X \rangle \langle X^{(\mathbf{x})} \rangle \right| \leq \text{Cst } e^{-m|\mathbf{x}|} \quad (101)$$

with

$$am \geq -4 \ln \beta_j + \text{Cst}$$

Indeed, only diagrams connecting the links of X and $X^{(\mathbf{x})}$ contribute to this quantity. They have at least $4|\mathbf{x}|/a$ plaquettes. This number corresponds to a minimal tube made of plaquettes and joining the two distant sets of links X and $X^{(\mathbf{x})}$. In the r.h.s.

series of (99), the summation starts at this value, and the result follows.

The second important consequence is the proof of the exponential decrease of the Wilson loop in terms of the minimal area $\mathcal{A}(C)$ bounded by the loop. Under hypotheses which have been given in section 1.3 and which will be justified below, the minimal diagram is made of the plaquettes covering the surface of minimal area bounded by the curve and contains $\mathcal{A}(C)$ plaquettes. The upper bound (99) yields a lower bound for the string tension (45)

$$K \geq -\ln \beta_j + \text{Cst} \quad (102)$$

6.3.2 Character expansions

Our goal here is to perform more effectively some of the group integrations. The diagrammatic interpretation of the expansion should however be modified in order to use the corresponding techniques. This modification yields an expansion analogous to the $\tanh \beta$ series (rather than β) for the Ising model. We recall first some useful results of the theory of compact groups. Let us consider the irreducible (unitary) representations r , of dimension d_r , of the group. Their matrix elements are a basis for all functions defined on the group. Using the invariant measure dU normalized to unity ($\int dU = 1$), they satisfy the following orthogonality and completeness relations

$$\int dU \mathcal{D}_{\alpha\beta}^r(U) \mathcal{D}_{\gamma\delta}^{s*}(U) = \frac{\delta_{rs} \delta_{\alpha\gamma} \delta_{\beta\delta}}{d_r} \quad (103)$$

$$\sum_{r, \alpha, \beta} d_r \mathcal{D}_{\alpha\beta}^r(U) \mathcal{D}_{\alpha\beta}^{s*}(V) = \delta(U, V) \quad (104)$$

On the r.h.s. of equation (104), the δ -function has the appropriate definition on the group manifold. The characters $\chi_r(U) = \sum_{\alpha} \mathcal{D}_{\alpha\alpha}^r(U)$, defined as the traces of irreducible representations, are thus an orthonormal basis for the *class* functions on the group. This statement is expressed by the relations

$$\int dU \chi_r(U) \chi_s^*(U) = \delta_{rs} \quad (105)$$

$$\sum_r d_r \chi_r(UV^{-1}) = \delta(U, V) \quad (106)$$

In particular, we have the following useful consequence

$$d_r \int dU \chi_r(U) \chi_s^*(U^{-1}V) = \delta_{rs} \chi_r(V) \quad (107)$$

The decomposition in terms of characters generalizes Fourier series. Let us apply it to the exponentiated action

$$\exp \beta_j \chi(U_p) = \sum_r \tilde{\beta}_r \chi_r(U_p) \quad (108)$$

with

$$\tilde{\beta}_r = \int dU \chi_r^*(U) \exp \beta_j \chi(U) \quad (109)$$

We define

$$\tilde{\beta}_r = \tilde{\beta}_0 d_r \beta_r \quad \text{for } r \neq 0 \quad (110)$$

so that the partition function may be written as

$$Z = \tilde{\beta}_0^{n_p} \int \prod_p \left[1 + \sum_{r \neq 0} d_r \beta_r \chi_r(U_p) \right] \prod_{ij} dU_{ij} \quad (111)$$

As usual, each term of the expansion of the product in the integrand is interpreted as giving rise to a diagram. Each one is thus a set of *distinct* plaquettes, to each of which a nontrivial representation has been assigned. We emphasize that, in contradistinction with the expansion (94), two plaquettes should not share the same location on the lattice. Note that the plaquette action χ is real, but that this need not be the case for χ_r . It is necessary to choose for each plaquette a conventional orientation. Changing this convention replaces U_p by U_p^{-1} , that is replaces the representation r by its conjugate $\bar{r}$, ($\mathcal{D}_r(U^{-1}) = \mathcal{D}_r^\dagger(U) \sim \mathcal{D}_{\bar{r}}(U)$), which is generally not equivalent to r . Nevertheless, as the original χ is real, one has $\beta_r = \beta_{\bar{r}}$ and the choice of a conventional orientation does not affect the results (as was of course to be expected).

(1) Give the explicit form for the Fourier components β_r for the two groups $SU(2)$ and $SO(3)$. Matrices U can be parametrized as

$$U = \cos \frac{1}{2} \theta + i \sigma \cdot \hat{n} \sin \frac{1}{2} \theta, \quad 0 \leq \theta \leq 4\pi \quad (112)$$

in terms of Pauli matrices. The Haar measure is normalized to unity and reads

$$dU = \sin^2 \frac{1}{2} \theta \frac{d\theta}{2\pi} \frac{d^2 \hat{n}}{4\pi} \quad (113)$$

The corresponding irreducible representations are classified by a non-negative integer or half-integer spin j and the characters are given by

$$\chi_j(U) = \frac{\sin(j + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta} \quad (114)$$

A direct application of (109) yields the coefficients $\tilde{\beta}_j$ in terms of modified Bessel functions, and therefore

$$\exp(\frac{1}{2}\beta_j \text{Tr } U) = \exp \beta_j \cos \frac{1}{2}\theta = \sum_j 2(2j+1) \frac{I_{2j+1}(\beta_j)}{\beta_j} \chi_j(U) \quad (115)$$

The group $SO(3)$ leads to analogous expressions, discarding the representations with noninteger spin j . Table II displays the values of these coefficients for some usual gauge groups, computed for Wilson's action. Finally, the geometrical action derived from the heat kernel (9) is directly written in the Fourier representation. For example, the values of the quadratic Casimir operator $C_j^{(2)} = j(j+1)$ for the $SU(2)$ group lead to formula (10).

(2) Derive the asymptotic form of the coefficients β_r for the Wilson action in the two limiting cases, β_j small or large.

In the expansion (109) of the exponentiated action, the orthogonality relation (105) gives a vanishing result as long as the representation r does not occur in the expansion of the corresponding power of χ , the character for the sum of the fundamental representation and of its conjugate. Therefore β_r vanishes as an integral power of β_j . This yields the small β_j behaviour. For instance, for the $SU(2)$ group, at least $2j$ spins $\frac{1}{2}$ are needed to construct a spin j , and hence $\beta_j \sim \beta_j^{2j}$. Note also that β_r increases monotonically from 0 to 1 as β_j varies from 0 to infinity. For a continuous gauge group, the large β_j behaviour is obtained through a steepest descent method around $U = I$

$$\tilde{\beta}_r \sim \text{Cst } d_r \frac{\exp(\beta_j \chi(1))}{\beta_j^{X/2}} (1 + \mathcal{O}(\beta_j^{-1})) \quad (116)$$

where X denotes the dimension of the Lie algebra (i.e. the number of infinitesimal independent generators of the group). Up to a shift in the action, amounting to subtracting the value of the character

Table II. Coefficients in the character expansion of the Wilson action

group	β_0	r and $\chi_r(U)$	d_r	β_r
Z_2	$\cosh \beta_j$	$\sim \frac{Z_2 U}{n e^{i n \varphi}}$	1	$\sinh \beta_j$
$U(1)$	$I_0(\beta_j)$	$j \frac{1}{2} \text{ integer } \frac{\sin(j + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta}$	1	$I_n(\beta_j)$
$SU(2)$	$2I_1(\beta_j)/\beta_j$	$j \text{ integer } \frac{\sin(j + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta}$	$2j+1$	$I_{2j}(\beta_j) - I_{2j+2}(\beta_j)$
$SO(3)$	$e^{\beta_j/3} (I_0(\frac{2}{3}\beta_j) - I_1(\frac{2}{3}\beta_j))$	$l_1 \geq \dots \geq l_n U = \text{Diag}(e^{i\alpha_l})$	$2j+1$	$e^{\beta_j/3} (I_j(\frac{2}{3}\beta_j) - I_{j+1}(\frac{2}{3}\beta_j))$
$U(n)$	$\det I_{i-j}(\beta_j/n)$	$\chi = \frac{\det(\exp i(l_j + n - j)\alpha_i)}{\det(\exp i(n-j)\alpha_i)}$	$\prod_{i < j} \frac{l_i - l_j + j - i}{j - i}$	$\det I_{l_i - j + i}(\beta_j/n)$
$SU(n)$	$(1 \leq i, j \leq n)$ $\sum_{k=-\infty}^{\infty} \det I_{i-j+k}(\frac{\beta_j}{n})$	as $U(n)$ with $l_n \equiv 0$	as $U(n)$	$\sum_{k=-\infty}^{\infty} \det I_{l_i - j + i + k}(\frac{\beta_j}{n})$
$U(\infty)$ or $SU(\infty)$	$\exp(\frac{\beta_j}{2n})^2$ in the limit $n \rightarrow \infty$ with fixed $\beta_j n^2$			

at the identity, the coefficient $\tilde{\beta}_r$ behaves according to a power of β_J for a continuous group. On the other hand, for a discrete group, this coefficient decreases exponentially.

The diagrammatic strong coupling expansion may be analyzed from a topological point of view. A given subset of plaquettes is homeomorphic to a simple surface if any link bounds at most two plaquettes of this subset. The boundary of this surface is made of the links bounding exactly one plaquette, and is homeomorphic to a set of simple closed curves. Any diagram can be decomposed as a set of maximal simple surfaces by cutting it along the links bounding more than two plaquettes. Each of the boundary curves is

- either a true free boundary, bounding only one simple surface,
- or a singular branch line along which n ($n \geq 3$) simple surfaces meet.

The classification of simple surfaces is made according to the following criteria.

- Orientability.** Two neighbouring oriented plaquettes are said to be *coherently oriented* if they induce opposite orientations on the common adjacent link. A surface is *orientable* if there exists an orientation convention for all its plaquettes such that any pair of neighbouring plaquettes is coherently oriented.
- The number b of holes, i.e. the number of simple closed curves bounding the surface.
- The *genus* $g = 2 - (n_2 - n_1 + n_0 + b) \geq 0$, where n_2 is the number of plaquettes, n_1 the number of distinct links and n_0 the number of sites located on the surface. Any orientable surface of genus g (which is always even in this case) is homeomorphic to a sphere with $\frac{1}{2}g$ handles and b holes. The torus corresponds to the case $g = 2$, $b = 0$, and the disk to $g = 0$, $b = 1$. Conversely, if the surface is not orientable, it is homeomorphic to a sphere drilled with $b + g$ holes. Among them, g (≥ 1) are closed by a Möbius band. For example, the Möbius band corresponds to $b = 1$, $g = 1$; the projective plane to $g = 1$, $b = 0$; and the Klein bottle to $g = 2$, $b = 0$. In later chapters, when dealing exclusively with orientable surfaces, we shall denote by g , and still call genus, half of the above (even) quantity.

Now, let us perform the integration over the gauge fields. Relation (103) forbids the existence of free boundaries for diagrams pertaining to the partition function. For a given observable, the free boundary coincides with the set of links occurring in the definition of this observable. Relation (107) allows the integration over the internal links of the simple surface components corresponding to a given diagram. In particular, the plaquettes of a simple surface component should carry the same representation when they are coherently oriented. As a consequence, non orientable surfaces carrying complex representations (representations which are not equivalent to their conjugate) should be excluded. The integration over inner links of a simple component yields

$$\beta_r^{n_2} d_r^{2-g-b} \prod_{\text{boundaries}} \chi_r(U_{\text{boundary}}) \quad (117)$$

Check this result.

The final step consists of integrating gauge fields along the singular branch lines. This is the most delicate part, and it depends on the Clebsch–Gordan coefficients coupling the representations of the group.

Show that, if there is only one singular closed line which is the common boundary shared by p simple surfaces with representations $\tau_1, \tau_2, \dots, \tau_p$, the integration over gauge fields yields an integral factor $N_{\tau_1, \dots, \tau_p}$, which counts the number of times the trivial representation is contained in the product $\tau_1 \otimes \tau_2 \otimes \dots \otimes \tau_p$.

Let us consider in more detail the possible topologies of diagrams contributing to the Wilson loop. The simplest topology is a disk, the boundary of which is the loop. As we have already seen, this kind of diagram leads to an exponential decrease of the Wilson loop parameter with respect to the minimal area enclosed by the loop. However, other topologies may occur, e.g. the toric diagram displayed in figure 12. This kind of diagram obviously leads to a perimeter law, unless its contribution vanishes identically. It describes a process in which the charges are screened and can thus be separated, without confinement. The leading behaviour of the terms associated with the disk and the toric topologies are respectively $\beta_J^{L^2}$ and $\beta_J^{16(L-1)}$, so that the transition between an area law and a perimeter law is expected to occur

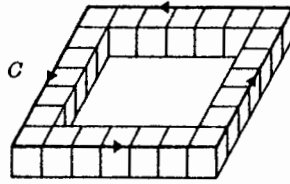


Fig. 12 A diagram leading to a perimeter law for the Wilson loop.

for a loop side $L \simeq 15$. This size exceeds the possibilities of present computers and, from numerical simulations, one may hastily conclude that the asymptotic behaviour follows an area law. Fortunately, the toric diagrams vanish when the center of the group is nontrivial. Indeed, using the result from the last exercise, one sees that the contribution vanishes when $r \otimes \bar{r}$ does not contain the representation used in the definition of the Wilson loop. Since on the one hand $r \otimes \bar{r}$ induces a trivial representation on the center (which is an invariant subgroup), while on the other the fundamental representation is non trivial on this subgroup, the afore mentioned contribution has to vanish. For instance, with the $SU(2)$ group, one cannot screen a spin $\frac{1}{2}$ with $j \otimes j$ combinations, which are always of integral spin. With the $SO(3)$ gauge group, the fundamental representation of which is of spin 1, this screening does occur and the Wilson loop operator cannot be used to distinguish the phases of the system.

6.3.3 Free energy

We have now all the prerequisites to derive various expansions. The problem of enumerating and counting all graphs is treated in chapter 7 (volume 2). We consider first the series for the free energy per site, for a d -dimensional hypercubic lattice. The results given below are valid for any gauge group. Let us introduce first some coefficients corresponding to the various possible topologies. These are

- The sphere

$$S_n = \sum_r \beta_r^n d_r^2$$

- The torus

$$T_n = \sum_r \beta_r^n$$

- p disks sharing the same closed curve as boundary

$$\Theta_{n_1 \dots n_p} = \sum_{r_1 \dots r_p} \beta_{r_1}^{n_1} \dots \beta_{r_p}^{n_p} d_{r_1} \dots d_{r_p}$$

- A cylinder closed at both extremities by duplicated disks

$$\zeta_{n_1 \dots n_5} = \sum_{r_1 \dots r_5} \beta_{r_1}^{n_1} \dots \beta_{r_5}^{n_5} d_{r_1} d_{r_2} d_{r_4} d_{r_5} N_{r_1 r_2 r_3} N_{r_3 r_4 r_5}$$

- Four simple arcs sharing the same endpoints; taken by pairs, they bound six different disks

$$\Phi_{n_{12} n_{13} n_{14} n_{23}}^{n_{34} n_{24} n_{25}} = \sum_{i < j} \prod d_{r_{ij}} \beta_{r_{ij}}^{n_{ij}} \int dU_i \prod_{i < j} \chi_{r_{ij}}(U_i U_j^{-1})$$

The free energy, up to and including diagrams with 16 plaquettes, is then given by

$$\begin{aligned} \frac{2 \ln Z}{d(d-1)} = & \ln \tilde{\beta}_0 + (d-2) \left[\frac{1}{3} S_6 + (2d-5) S_{10} + \frac{10}{3} (d-3) S_{12} \right. \\ & + (20d^2 - 106d + 143) S_{14} + (84d^2 - 504d + 757) S_{16} \\ & + \frac{1}{4} (d-3) T_{16} - (2d - \frac{29}{6}) S_6^2 - (44d^2 - 226d + 294) S_6 S_{10} \\ & + (2d-5) \Theta_{551} + \frac{2}{3} (d-3) (2d-5) (\Theta_{555} + \Theta_{5551}) \\ & + \frac{8}{3} (d-3) (3\Theta_{842} + 3\Theta_{10,4,2} + 2\Theta_{933} + \frac{3}{2} \Theta_{664} + \Phi_{444}^{111}) \\ & \left. + (20d^2 - 108d + 149) (2\Theta_{951} + \zeta_{51415}) + o(\beta_J^{16}) \right] \end{aligned} \quad (118)$$

This general formula can now be applied to any given gauge group or any variant one-plaquette action. This is just a straightforward, although tedious, exercise in algebra. At this level of complexity, the help of a computer algebra system is useful. Once the series with numerical coefficients has been obtained, it can be analyzed using the various general techniques reported in chapter 7 (volume 2).

We recall that pure gauge theories have a first order transition above some critical dimension. The free energy is not singular at this point, as was suggested by the mean field approximation. In this case, the two branches corresponding to the weak and strong coupling phases are two determinations of the same multivalued

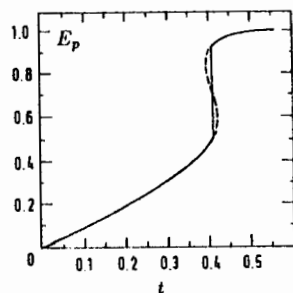


Fig. 13 The plaquette energy E as a function of $t = \tanh \beta_J$ for the Z_2 gauge system.

analytic function. It is *theoretically* possible, from the series, to perform an analytic continuation of the function and, following the adequate path on Riemann sheets, to compute the weak coupling determination and therefore to have access to the transition point. *In practice*, available series are much too short to allow for this process.

Let us analyze in more detail the four-dimensional Z_2 model. A direct analysis of the series in β_J shows evidence for a singularity located around $\tanh \beta_J = 0.48 - 0.49$, which is beyond the transition point $\tanh \beta_c = \sqrt{2} - 1$, exactly known by duality. We interpret this singularity as the end point of the metastable strong coupling phase. As suggested by the mean field approximation, the coupling constant might be a regular, uniform function of the internal energy per plaquette E . This can be checked by inverting the series $E(\beta_J)$ into a new series $\beta_J(E)$. The analysis confirms the expectation; no evidence for a singularity is found, confirming the idea of a first order transition which admits an analytic unfolding. Figure 13 displays the resummed series and allows the transition to be located, compatible within errors with the duality prediction.

Four-dimensional nonabelian gauge systems are more difficult to analyze. Numerical simulations seem to imply no transition. Nevertheless one finds a peak in the specific heat (figure 14). Various resummation techniques of the series diverge in this region, and it is rather difficult to reproduce the structure seen in figure 14. Additional hypotheses are necessary to improve the fit. For instance, one may conjecture the existence of a pair of

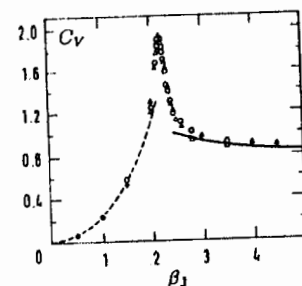


Fig. 14 The specific heat of the four-dimensional $SU(2)$ system (from B. Lautrup and M. Nauenberg, *Phys. Rev. Lett.* 45 (1980) 1755).

conjugate complex poles in this region. For practical purposes, the best technique consists in matching the series into the weak coupling region by using results from asymptotic freedom. An example will be given for the string tension.

In the limit of a large dimension, strong coupling expansions can be reordered in a way different from the $1/d$ series of successive corrections to mean field. Looking at formula (118), one sees that the coefficient of each term is a polynomial in d with a degree increasing by unity every fourth term. It is thus suggested to keep the quantity $\beta_J^4 d$ constant as d tends to infinity and to reorder the series as an expansion in $d^{-1/4}$. Let us consider the following reduction process on a connected diagram. Suppress a slab of the lattice limited by two consecutive hyperplanes $k < x_\mu < k+1$ and containing at least one plaquette of the diagram. As the diagram is closed, at least four plaquettes disappear. Moreover, other plaquettes may be duplicated when the two half spaces are stuck together and must be replaced (combining their group representations) either by a single plaquette, or no plaquette at all. In any case, this operation decreases the number of plaquettes by at least four, while the *dimensionality* of the diagram (i.e. the number of different directions effectively used by the plaquettes of the diagram, or equivalently the degree of the polynomial in d entering in its contribution) decreases at most by one unit. It is also clear that the dominant terms in the above limit are obtained whenever exactly four plaquettes disappear at each step in the reduction, until one is left with a simple diagram having

the shape of a cube. This characterizes the dominant diagrams as the boundary of connected trees made with three-dimensional cubes, with all plaquettes sharing the same group representation.

Since each cube has six faces, it has at most six neighbours in the tree diagram. In a step-by-step reconstruction, the addition of a new cube on a given plaquette of a cube leads to $2d-5$ ($\approx 2d$) possibilities. Some of these may be forbidden by excluded volume effects, but their number is negligible with respect to $2d$. The expectation value $p_r = \langle \chi_r(U_p) \rangle / d_r$ of a single plaquette in the representation r , satisfies in this limit the self-consistency equation

$$p_r = \beta_r + 2dp_r^5 \quad (119)$$

This is parametrized as

$$\begin{cases} f_r \equiv \frac{p_r}{\beta_r} = (1 - u_r)^{-1} \\ x_r \equiv 2d\beta_r^4 = u_r(1 - u_r)^4 \end{cases} \quad (120)$$

The free energy is recovered by integration of the internal energy p_r per plaquette

$$\begin{aligned} \tilde{F} &\equiv \frac{\ln Z}{N} - \frac{d(d-1)}{2} \ln \beta_0 \\ &= \frac{d^{3/2}}{12\sqrt{2}} \sum_{r \neq 0} d_r^2 x_r^{1/2} \frac{u_r(1-3u_r)}{(1-u_r)^2} (1 + \mathcal{O}(d^{-1/4})) \end{aligned} \quad (121)$$

With the Wilson action, only the fundamental representation and its conjugate give a dominant contribution to (121), and the subtracted free energy $\tilde{F}$ takes a universal form displayed on figure 15. It presents a cusp at $u = 1/5$, i.e. for $2d\beta_j^4 = 4^4/5^5$. As the rearrangement of the series is only possible within the radius of convergence, only the arc OA is physically acceptable. This is the whole strong coupling phase, including the metastable region. The second derivative of the free energy, which could be interpreted as a plaquette-plaquette susceptibility, behaves as $(\beta_A - \beta_j)^{-1/2}$ in the vicinity of the point A , as d goes to infinity. This is reminiscent of a continuous phase transition with a corresponding correlation length diverging as $(\beta_A - \beta_j)^{-\nu}$ for an exponent $\nu = \frac{1}{4}$ (see chapter 11 in volume 2). The weak coupling branch is not seen here. Indeed, from the analysis done in the mean field framework, this phase lies along the vertical axis, using the present variables.

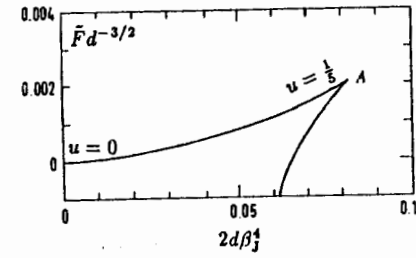


Fig. 15 The leading order in a $d^{-1/4}$ expansion for the rescaled (and subtracted) free energy in large dimension plotted as a function of $2d\beta_j^4$.

Compute the next order correction to this approximation.

6.3.4 String tension and roughening transition

In nonabelian gauge theories without transitions in four dimensions, one wants to extrapolate strong coupling series for the string tension to the weak coupling regime. This would provide a link with the scaling behaviour and allow useful information on confinement in the continuous limit to be obtained. However, two difficulties arise. The first one has already been discussed and is related to the delicate resummation in the intermediate region. The second one is specific to this nonlocal observable and corresponds to the *roughening transition* which does not affect bulk quantities. Its name originates from crystal growth theory where a similar phenomenon occurs.

As a simple example, let us consider a three-dimensional Ising model, dual to a Z_2 gauge theory, with boundary conditions such that there exists at low temperature an interface separating two regions of opposite orientation. Boundary spins are constrained to point up above a given plane, and down under this plane. The surface tension is related to the difference in free energy as compared with a homogeneous system (chapter 2). If A denotes the area of the minimal interface, and ΔF the difference in free energy,

$$\Delta F = -KA \quad (122)$$

with K the string tension of the corresponding Z_2 gauge model in its strong coupling phase.

At zero temperature, the interface is a crystallographic plane. As T increases, it fluctuates more and more. When a temperature T_R , generally smaller than T_c , is reached, these properties change drastically. The fluctuations are so important that it is no longer possible to say that the interface is localized. Let us be more precise and analyze quantitatively the phenomenon. Let h_i be the height of the interface above its equilibrium position $\mathbf{r}_i$ in the reference plane, and consider the quantity

$$p(h, \mathbf{r}_i - \mathbf{r}_j) = \langle \delta(h_i - h_j - h) \rangle \quad (123)$$

which expresses the probability that two points separated by a distance $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ have a height difference of h . The width of the interfacial region can be defined as the second momentum Δh of this distribution in the large distance limit

$$\sum h^2 p(h, r_{ij}) = \langle (h_i - h_j)^2 \rangle \rightarrow (\Delta h)^2 \text{ as } r_{ij} \rightarrow \infty \quad (124)$$

We say that the system undergoes a *roughening transition* if this quantity diverges as the temperature reaches (from below) a *roughening temperature* T_R . An equivalent description is to consider the free energy fL needed for the construction of a step of length L on the interface. The quantity f vanishes for $T \geq T_R$, so that the large-scale deformations of the surface are energetically costless. The surface tension K has no reason to vanish at T_R , but is singular at this temperature, as will be seen below.

So far this description has been qualitative. It is based on converging evidence coming from the study of exact models, Monte Carlo simulations and experiments on crystal growth. A convenient description is the *solid on solid (SOS)* model where one considers only a restricted class of possible deformations, excluding overhangs and disconnected inclusions of one phase into the other. It leads to an action of the type $\beta\mathcal{H} = \beta \sum_{(i,j)} |h_i - h_j|$,

which can be also approximated by $\beta\mathcal{H} = \beta \sum_{(i,j)} (h_i - h_j)^2$, a model belonging to the same universality class. One recognizes the discrete Gaussian model, dual to the XY -model which has been studied in chapter 4. Its transition is the roughening transition. The surface tension, which is the free energy of this system, admits an essential singularity of the form

$$K = A(\beta) + B(\beta) \exp\left(-\frac{C}{\sqrt{T_R - T}}\right) \quad (T < T_R) \quad (125)$$

The step free energy f is identified with the inverse correlation length, hence

$$f \sim \xi_{XY}^{-1} \sim \exp\left(-\frac{C}{2\sqrt{T_R - T}}\right) \quad (126)$$

The scale governing the phenomena is ξ_{XY} when $T < T_R$ (this corresponds to the high temperature region of the XY -model). For $T > T_R$, ξ_{XY}^{-1} vanishes and the scale reduces to the size L of the sample. Therefore, we introduce the quantity λ defined by $\lambda = \ln(\xi_{XY}/a)$ for $T < T_R$, $\lambda = \ln(L/a)$ for $T > T_R$. At large distances, the probability $p(h, r)$ which has been introduced in (123) reads

$$p(h, L) \sim \frac{1}{\sqrt{4\pi\lambda}} \exp\left(-\frac{h^2}{4\lambda}\right) \quad (127)$$

Hence the interface width is

$$\Delta h = \sqrt{2\lambda} \quad (128)$$

while the probability for a point of the interface to be at its equilibrium position is

$$p(0) = \frac{1}{\sqrt{4\pi\lambda}} \sim \frac{1}{\Delta h} \sim \begin{cases} (T_R - T)^{1/4} & \text{for } T < T_R \\ \frac{1}{\sqrt{\ln L}} & \text{for } T > T_R \end{cases} \quad (129)$$

Such behaviour is presumably universal, and should also apply to the interface in the original Ising model. According to the standard lore, the roughening transition does not affect bulk quantities like the free energy or its derivatives, nor the spin-spin correlation functions.

From duality, these conclusions remain valid for the Z_2 gauge model in three dimensions, where the interface free energy per unit area is interpreted as the string tension. The discussion can also be extended to higher dimensions. A rigorous proof of the existence of this transition is still lacking. To get clear evidence, we need a better indicator than the Wilson loop which only presents a very weak singularity of the type (125). Other quantities diverge at T_R and yield a sharper signal. For instance,

$$\eta^2 = \frac{\sum_{\mathbf{x}_\perp} E(\mathbf{x}) x_\perp^2}{\sum_{\mathbf{x}_\perp} E(\mathbf{x})} \quad (130)$$

with

$$E(\mathbf{x}) = \frac{\langle W(C)\chi(U_p) \rangle - \langle W(C) \rangle \langle \chi(U_p) \rangle}{\langle W(C) \rangle} \quad (131)$$

Intuitively, $E(\mathbf{x})$ measures, using a test plaquette p parallel to the Wilson loop plane, the density of chromoelectric energy at point $\mathbf{x}$, and $\mathbf{x}_\perp$ is the distance from p to the minimal surface bounded by the Wilson loop. As a consequence, η^2 generalizes the expression (124), and one expects a divergence according to a power law. Corresponding series can be computed and analyzed, and one gets a clear indication for the expected singularity. Remarkably, the position of the roughening transition, expressed in terms of the gauge coupling β_f relative to the fundamental character, seems to be largely independent of the gauge group (up to uncertainties in the numerical analysis), which confirms its geometrical nature.

$$\begin{cases} d = 3 & \beta_{fR} \simeq 0.46 \\ d = 4 & \beta_{fR} \simeq 0.40 \\ d = 5 & \beta_{fR} \simeq 0.37 \end{cases} \quad (132)$$

The respective positions of the roughening and of the bulk transition are not fixed. A heuristic argument claims that the deconfining phase transition will be of first order if the roughening transition occurs at a lower temperature. Indeed, if the deconfining transition is of second order, the interquark potential behaves as $1/R$ at the critical point, whereas it is natural to expect an exponential fall-off of the corrections to the leading behaviour up to the roughening point.

One may wonder what is the critical theory at T_R . In three dimensions, we saw that it belongs to the universality class of the 2-dimensional XY -model. In d dimensions, the same reasoning suggests a set of $(d-2)$ two-dimensional decoupled XY -models. The indicator η would thus have a critical exponent $\frac{1}{4}$, while the step free energy vanishes with an essential singularity. This vanishing plays an important role; the surface can then freely move in the directions orthogonal to its minimal plane and behaves thus as a continuous surface. Translation and rotation symmetries are then restored. This has been checked both in

strong coupling calculations of the interquark potential, and in numerical simulations measuring the potential.

The analysis of strong coupling series for the string tension is of course complicated by the presence of the roughening transition. Generally, the effect of the nonanalyticity at β_R is to give the fallacious impression that the string tension vanishes prematurely. Figure 16 displays a comparison of strong coupling expansions to numerical simulations. In the transition region between strong and weak coupling regimes, the curve can be tangentially connected to the predictions of the renormalization group and asymptotic freedom. The string tension has a dimension of $length^{-2}$. Expressed in physical units, the parameter $\sigma = K\alpha^2$ can be interpreted in the framework of the old relativistic string model and is related to the Regge slope by $\sigma = (2\pi\alpha')^{-1}$. The Regge relation of particle physics is the almost linear expression of spin versus mass square for low lying particles and resonances. Experimentally, $\alpha' \simeq 0.90\text{GeV}^{-2}$ so that $\sqrt{\sigma} \simeq 420\text{MeV}$. Formula (31) provides the weak coupling behaviour

$$K \sim \sigma \Lambda_L^{-2} (b_0 g_0^2)^{-b_1/b_0^2} \exp(-1/b_0 g_0^2) \quad (133)$$

The corresponding curve looks almost like a straight line in the logarithmic scales used in figure 16. Relating this curve to the series extrapolations in the transition region allows the parameter Λ_L to be determined. It is in fair agreement with the values obtained by other techniques.

The string tension gives only the coefficient of the linearly rising part of the quark-antiquark static potential. Much effort has been devoted to obtaining its overall shape for a description of heavy quark-antiquark meson bound states, accessible experimentally. It is not our purpose here to review this important aspect of quantum chromodynamics, directly tied to particle phenomenology. Nor will we discuss a possible finite temperature deconfining transition of importance in astrophysics and heavy ion collisions. We will rather limit ourselves with a short description of the excitation spectrum of pure gauge theories.

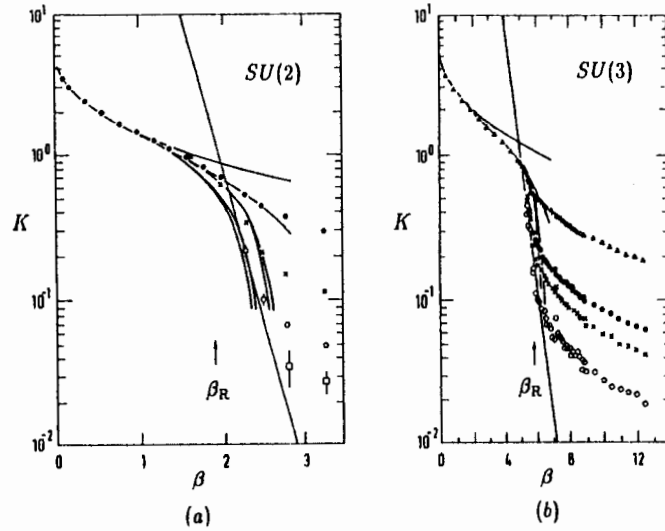


Fig. 16 The string tension for $SU(2)$ (a) and $SU(3)$ (b) systems. Data are taken from M. Creutz, *Phys. Rev. D* **21** 2308 (1980) and M. Creutz and K.J.M. Moriarty, *Phys. Rev. D* **26** 2166 (1982).

6.3.5 Mass spectrum

The mass spectrum of pure gauge theories seems to be the best candidate for applications of strong coupling series. Effects such as the roughening transition will not be present, and it is therefore expected that the predictions of strong coupling and asymptotic freedom can be easily matched in the transition region. The corresponding states have no quark components and are searched for experimentally under the name of *glueballs*.

The derivation of strong coupling series for the quantities of interest is, however, somewhat delicate. Correlation functions are expected to behave exponentially at large distance according to

$$G(r) = \langle \mathcal{O}_1(0) \mathcal{O}_2(r) \rangle_c \sim A_1(r) e^{-m_1(\hat{r})r} + A_2(r) e^{-m_2(\hat{r})r} + \dots \quad (134)$$

This formula involves several masses $m_1 \leq m_2 \leq \dots$ which depend on the direction $\hat{r}$ of observation. The simplest idea is to compute the correlation function between two distant plaquettes, using diagrams connecting their boundaries, and to derive $m_1(\hat{r}) =$

$\lim_{r \rightarrow \infty} \ln G(r)/r$. However, a degeneracy problem appears and prevents a direct exponentiation of the series.

In order to explain this phenomenon, let us compute the correlation length of the Ising model in a direction parallel to a principal axis.

Let n be the distance (in lattice spacing units) of the two spins. The lowest order diagram is drawn as the straight line joining the two locations. The next correction adds opposite kinks located at sites i and j ($0 \leq i < j \leq n$). Therefore

$$\langle \sigma_0 \sigma_n \rangle = (\tanh \beta)^n + (d-1)(\tanh \beta)^{n+2}n(n+1) + \dots \quad (135)$$

The correlation length is obtained by taking the logarithm, which is linear in n to lowest order

$$\xi = -\frac{1}{\ln \tanh \beta} (1 + \mathcal{O}(\tanh \beta)) \quad (136)$$

The subsequent correction to the logarithm of the correlation function is unfortunately quadratic in n and the series for the correlation length seems to have infinite terms. This paradox is due to the fact that the transfer matrix have several equal eigenvalues. Taking a direction parallel to a principal axis amounts to combining the corresponding eigenvectors in an unfortunate way which leads to this disaster. From translational invariance, a correct diagonalization can however be carried out using a Fourier transformation

$$G(p_\perp, x_0) = \sum_{x_\perp} \langle \sigma_0 \sigma_{x_0, x_\perp} \rangle e^{-ip_\perp x_\perp} \sim e^{-x_0/\xi} + \dots \quad (137)$$

Now, the first correction to the lowest order diagram is just a single kink (spins are shifted by one unit in perpendicular directions)

$$G(p, x_0) = (\tanh \beta)^n + (\tanh \beta)^{n+1} 2d(n+1) \sum_\mu \cos p_\mu + \dots \quad (138)$$

There are no more difficulties in computing the logarithm, even including further corrections.

The previous exercise shows that it is necessary to diagonalize the transfer matrix, taking into account the various symmetries. This can be partly done by going over to momentum space. One has also to take into account the discrete rotational invariance of the lattice and “ C -parity”, related to charge conjugation. We



Fig. 17 Random walk of a plaquette.

introduce the following Green function

$$G_{\mu\nu,\rho\sigma}^{\pm}(\mathbf{k}) = \sum_{\mathbf{x}} \left\langle \text{Tr} \frac{1}{2} \left(U_{\mu\nu}(0) \pm U_{\mu\nu}^{\dagger}(0) \right) \right. \\ \left. \times \text{Tr} \frac{1}{2} \left(U_{\rho\sigma}(\mathbf{x}) \pm U_{\rho\sigma}^{\dagger}(\mathbf{x}) \right) \right\rangle_c e^{i\mathbf{k}\cdot\mathbf{x}} \quad (139)$$

where $U_{\mu\nu}(\mathbf{x})$ is the plaquette variable U_p with directions μ and ν , the center of which is located at $\mathbf{x}$. Instead of computing G , let us, as usual, discuss its inverse. The lowest order contributions to this quantity are fourth order terms connecting two nearby plaquettes and completing the boundary of a cube. This is described in momentum space by a matrix $M_{\mu\nu,\rho\sigma}(\mathbf{k})$ obeying the symmetry relations

$$M_{\mu\nu,\rho\sigma} = M_{\rho\sigma,\mu\nu} = C M_{\nu\mu,\rho\sigma} \quad (C = \pm 1) \quad (140)$$

Explicitly,

$$\begin{cases} M_{\mu\nu,\mu\nu} = 2 \sum_{\rho \neq \mu, \nu} \cos k_{\rho} \\ M_{\mu\nu,\mu\rho} = \begin{cases} 4 \cos \frac{1}{2} k_{\nu} \cos \frac{1}{2} k_{\rho} & \text{when } C = +1 \\ 4 \sin \frac{1}{2} k_{\nu} \sin \frac{1}{2} k_{\rho} & \text{when } C = -1 \end{cases} \end{cases} \quad (141)$$

Other components vanish. Returning to the propagator G expressed as

$$[G^{-1}(\mathbf{k})]_{\mu\nu,\rho\sigma} = (1 - \beta_r^4 M_{\mu\nu,\rho\sigma}(\mathbf{k})) \quad (142)$$

one obtains a plaquette random walk description depicted on figure 17. In configuration space, this amounts to approximate G by retaining only diagrams in the shape of tubes connecting an initial to a final plaquette locations, in complete analogy with Brownian motion. Further corrections are obtained by adding various "decorations" to the propagator tube.

We have not yet taken into account the discrete rotational symmetry of the lattice. In four dimensions and for a propagation

along the fourth direction, we have to decompose the propagator according to the representations of the discrete cubic symmetry (rotations keeping invariant the fourth dimension). The following quantities are to be considered

- A singlet $\text{Re Tr}(U_{12} + U_{23} + U_{31})$
- A doublet $\{\text{Re Tr}(U_{13} - U_{23}); \text{Re Tr}(U_{13} + U_{23} - 2U_{12})\}$
- A triplet $\{\text{Im Tr } U_{12}; \text{Im Tr } U_{23}; \text{Im Tr } U_{31}\}$ which exists only for complex representations of gauge groups.

Check this decomposition.

We refer to the literature for details on the explicit computation of the glueball mass spectrum. Let us just quote a result for the fundamental singlet state 0^{++} for a $SU(3)$ colour gauge group. Series have been derived up to eighth order. To obtain stable extrapolations, it is better to re-express the scalar mass in terms of the plaquette energy before attempting to estimate the series. One finds

$$\frac{m_s}{\Lambda_L} \approx 340 \pm 40 \quad (143)$$

This value agrees with numerical simulations and would lead to a glueball mass of approximately 800 MeV. This result should be taken with a grain of salt, due to several uncertainties, including the effects of quarks neglected in this calculation. There seems also to exist a 2^{++} state with a mass close to the fundamental one. Higher excited states might also be present. It is still an open problem to identify the glueballs experimentally. An extensive numerical effort is made to ascertain the hadronic mass spectrum involving quarks. There are however a number of technical difficulties, described below, in the case where the latter have a small (and in the limit vanishing) mass.

6.4 Lattice fermions

6.4.1 The doubling problem

The main goal of lattice gauge theories is to yield a regulated model for quantum chromodynamics. Hence, one is specially

interested in fermionic fields. We have described in section 1 the technique which transforms a globally invariant action (as the bosonic action (1)) into a gauge invariant one. This method remains valid for fermions, but serious problems appear when one wants to write down a satisfactory global action for physical fermions.

Starting from the continuous action for free fermionic fields $\bar{\psi}$ and ψ of spin $\frac{1}{2}$

$$S_{\text{cont}} = \int d^d x \left(\frac{1}{2} i \bar{\psi} \gamma_\mu \partial_\mu \psi - m \bar{\psi} \psi \right) \quad (144)$$

the first idea is to replace derivatives by finite differences involving only nearest neighbours $\partial_\mu f \rightarrow (f_{i+a\bar{\mu}} - f_i)/a$. This leads to the so-called *naive fermionic action*

$$S_{\text{nf}} = \frac{1}{2} i a^{d-2} \sum_{(ij)} \bar{\psi}_i \gamma_{ij} \psi_j - m a^d \sum_i \bar{\psi}_i \psi_i \quad (145)$$

with

$$\gamma_{ij} = \gamma_\mu (x_i - x_j)^\mu$$

For short, we do not write explicitly all the indices of the Grassmannian fields $\bar{\psi}_i, \psi_i$. Besides the spatial location i referring to a lattice site, there are a spinor index (saturated by the Dirac matrices γ), a colour index (related to the internal symmetry which will become a local one and on which the gauge fields will act), and possibly flavour indices (corresponding to other global internal symmetries which we ignore here; the mass m may depend on them). The Euclidean Dirac matrices obey the anticommutation rule

$$\{\gamma_\mu, \gamma_\nu\} = -2\delta_{\mu\nu} \quad (146)$$

They will be written in some faithful matricial representation of minimal dimension $2^{\lfloor d/2 \rfloor}$. The corresponding matrices generate the algebra of all $\lfloor d/2 \rfloor \times \lfloor d/2 \rfloor$ matrices. In even dimension, the matrix

$$\hat{\gamma}_5 = i^{d/2} \gamma_1 \gamma_2 \cdots \gamma_d \quad \hat{\gamma}_5^2 = 1 \quad (147)$$

is linearly independent of, and anticommutes with, the γ_μ 's. Added to the d previous matrices (and multiplied by i), it will be the $(d+1)$ -th Dirac matrix for the odd dimension $d+1$. In the case where d is even, the γ_μ 's can be chosen antihermitian and

$\hat{\gamma}_5$ hermitian. The notation $\hat{\gamma}_5$ is borrowed from four-dimensional space-time physics.

Let us compute the propagator corresponding to the free field action (145). In the momentum representation, it reads

$$\Delta(\mathbf{p}) = \left(\frac{1}{a} \gamma_\mu \sin p_\mu a + m \right)^{-1} \quad (148)$$

The masses of the particles are determined by the position of the poles. In the limit $a \rightarrow 0$, poles lie in the vicinity of points $\bar{\mathbf{p}}$ with components equal to either 0 or π/a (recall that the components of the momenta are defined on the lattice modulo $2\pi/a$). Therefore, there are 2^d such poles and not a unique one as expected. This species multiplication reflects a discrete symmetry of the naive fermionic action. Let us introduce in even dimension the d operators $T_\mu = \gamma_\mu \hat{\gamma}_5 (-1)^{\sum_{\nu \neq \mu} z_\nu / a}$, $\mu = 1, \dots, d$. They generate a discrete group of order 2^{d+1} , and it can be easily checked that the action (145) is invariant under their action ($\psi \rightarrow T_\mu \psi$, $\bar{\psi} \rightarrow \bar{\psi} T_\mu^\dagger$). Each of these transformations T translates the momentum by one of the quantities $\bar{\mathbf{p}}$ defined above and acts on spinor indices through a matrix $S(\bar{\mathbf{p}})$. In the vicinity of each of these poles, the propagator (148) reads

$$\Delta(\mathbf{p} = \bar{\mathbf{p}} + \mathbf{k}) = S(\bar{\mathbf{p}}) \frac{m - \gamma_\mu k_\mu}{k^2 + m^2} S(\bar{\mathbf{p}})^{-1} + \mathcal{O}(a) \quad (149)$$

The effect of the transformations S is just to change the representation of the Dirac matrices into equivalent ones. This formula shows that the 2^d Dirac particles play exactly the same role. At this level of free fields, it is possible to introduce in the action (145) a projector on only one of these states and therefore to eliminate the unwanted partners. This is not possible any more, using local operators, when the fields interact, since the unwanted particles can be produced by pairs and contribute to intermediate states.

Study the degeneracy problem in odd dimensions.

Several solutions have been proposed to eliminate this problem. However, every one presents some disadvantages. We will show in the next subsection that it is in fact impossible to define a fermionic action on the lattice with all desired properties. A first attempt at a cure is to partly delocalize the mass term on

neighbouring sites in an amount proportional to a parameter r . The continuous limit $a \rightarrow 0$ is formally not changed by this operation. One obtains in this way the fermionic *Wilson action*

$$S_{\text{wf}} = \frac{1}{2} \sum_{(ij)} \bar{\psi}_i (i\gamma_{ij} a^{d-2} + r a^{d-1}) \psi_j - \left(m + \frac{dr}{a}\right) \sum_i \bar{\psi}_i \psi_i \quad (150)$$

The corresponding free propagator reads

$$\Delta(p) = \left(\frac{1}{a} \gamma_\mu \sin p_\mu a + m + \frac{r}{a} \sum_\mu (1 - \cos p_\mu a) \right)^{-1} \quad (151)$$

In the neighborhood of momentum $\bar{p}$ (with k components equal to π/a , and $d-k$ vanishing components), the particle has, in the limit $a \rightarrow 0$, a mass equivalent to $m + 2kr/a$. In the continuous limit, only one state near $p = 0$ remains with a finite mass and other particles decouple, their mass becoming infinite.

The disadvantage of this formulation is to break explicitly the chiral invariance of the massless limit

$$\psi \rightarrow e^{\lambda \gamma_5} \psi, \quad \bar{\psi} \rightarrow \bar{\psi} e^{\lambda \gamma_5} \quad (152)$$

which was a (noncompact) symmetry of the continuous action and of the naive lattice fermionic action as the mass m vanishes. In particle physics, this symmetry plays a very important role in the theory of weak interactions. Indeed, it then follows that the two states with opposite chirality ± 1 , obtained from the action of the projectors $\frac{1}{2}(1 \pm \gamma_5)$, are decoupled in a massless theory and remain so, even when turning on the interaction. It is thus possible to write an action for particles with a definite chirality, without introducing partners with the opposite one (Weyl fermions). The standard model of Glashow-Weinberg-Salam for electroweak interactions uses this property in a crucial way. It would therefore be desirable to construct a chiral invariant regularized lattice model, without the unwanted feature of particle doubling.

6.4.2 The Nielsen-Ninomiya theorem

A theorem due to Nielsen and Ninomiya states that this lattice construction is impossible under some general hypotheses. It was first proved in a Hamiltonian formalism. We give here a proof

adapted to the Euclidean formulation. The ingredients of this no-go theorem are the following. Consider a *discrete* fermionic system on a *regular* lattice, with a *local*, *hermitian* and *translational invariant* action. Then, there are as many states with chirality $+1$ than states with chirality -1 . We will make more precise the hypotheses as we proceed with the proof.

Assume to the contrary that it is possible to write an action for fermions with chirality $+1$. The most general quadratic term is of the form

$$\sum_{i,j,\mu} \bar{\psi}_i \gamma_\mu \frac{1}{2} (1 + \hat{\gamma}_5) i \mathcal{F}_\mu(\mathbf{x}_i - \mathbf{x}_j) \psi_j \quad (153)$$

Translational invariance requires the function $\mathcal{F}_\mu$ to depend only on the difference of arguments, and hermiticity that $\mathcal{F}_\mu^*(\mathbf{x}) = \mathcal{F}_\mu(-\mathbf{x})$ (assuming that the lattice has a symmetry $\mathbf{x} \rightarrow -\mathbf{x}$). We shall now show that this action automatically contains opposite chirality partners. The Fourier transform

$$\tilde{\mathcal{F}}_\mu(\mathbf{k}) = \sum_i \mathcal{F}_\mu(\mathbf{x}_i) e^{i\mathbf{k} \cdot \mathbf{x}_i} \quad (154)$$

is analytic in $\mathbf{k}$, periodic with the periodicity of the reciprocal lattice, and real. Rather than analyticity, we only need that $\tilde{\mathcal{F}}_\mu(\mathbf{k})$ be smooth, and this is the case as soon as $|\mathbf{x}|^d \mathcal{F}_\mu(\mathbf{x})$ vanishes sufficiently fast at large distance. Hence the theorem applies for a very large class of models using interactions beyond nearest neighbours.

The quadratic action (153) is diagonal in the momentum representation, and each physical state corresponds to a simple isolated zero of $\tilde{\mathcal{F}}_\mu(\mathbf{k})$ (these are usually called "critical points" for the vector field $\tilde{\mathcal{F}}_\mu(\mathbf{k})$). In the neighborhood of such a point, for a once differentiable $\tilde{\mathcal{F}}_\mu$, we have

$$\tilde{\mathcal{F}}_\mu(\mathbf{k}) = \mathcal{M}_{\mu\nu}(\mathbf{k} - \bar{\mathbf{k}})_\nu + \mathcal{O}((\mathbf{k} - \bar{\mathbf{k}})^2) \quad (155)$$

We assume a generic situation with $\partial/\partial k_\nu \tilde{\mathcal{F}}_\mu(\mathbf{k})|_{\mathbf{k}=\bar{\mathbf{k}}}$ nonsingular. The real matrix $\mathcal{M}$ with elements $\mathcal{M}_{\mu\nu}$ can be decomposed as the product $\mathcal{M} = \mathcal{O}S$ of an orthogonal matrix by a symmetric positive definite one. The orthogonal matrix $\mathcal{O}$ can always be absorbed into a redefinition of the fermionic fields. Indeed there always exists a matrix $\mathcal{R}$ such that

$$\gamma_\mu (1 + \hat{\gamma}_5) \mathcal{O}_{\mu\nu} = \mathcal{R}^{-1} \gamma_\nu (1 + \varepsilon \hat{\gamma}_5) \mathcal{R} \quad (156)$$

The sign $\varepsilon = \det \mathcal{O} = \pm 1$ (which is also the sign of $\det \mathcal{M}$) is called the *index* of the vector field $\tilde{F}_\mu(\mathbf{k})$ at the critical point $\bar{\mathbf{k}}$. To prove this relation, we observe that (even) d -dimensional space can be considered as the subspace of a $d+1$ -dimensional one. Extend in this larger space the orthogonal transformation $\mathcal{O}_{\mu\nu}$ to an element $\tilde{\mathcal{O}}_{\mu\nu} \in SO(d+1)$ by acting on the last coordinate axis through multiplication by $\varepsilon = \det \mathcal{O}_{\mu\nu}$. Since $SO(d+1)$ admits a spinor representation, there exists a matrix $\mathcal{R}$ in the Clifford algebra such that

$$\mathcal{R}^{-1} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_d \\ \hat{\gamma}_5 \end{pmatrix} \mathcal{R} = \begin{pmatrix} \gamma_\mu \mathcal{O}_{\mu\nu} \\ \varepsilon \hat{\gamma}_5 \end{pmatrix}$$

thus justifying (156).

Prove equation (156) in details for $d=2$.

The positive definite matrix $\mathcal{S}$ generates a transformation in momentum space around $\bar{\mathbf{k}}$. It amounts to a positive rescaling of each component of $\mathbf{k} - \bar{\mathbf{k}}$ in a rotated frame. This does not affect the momentum space orientation. As a consequence, the critical point (155) can be interpreted as giving rise to a physical state of chirality ε .

The real vector field $\tilde{F}_\mu(\mathbf{k})$ is defined on a compact manifold, here the torus T^d . Indeed, as the lattice is regular, the function $\tilde{F}_\mu(\mathbf{k})$ is periodic. A theorem due to Poincaré and Hopf states that the sum of indices ε of a real vector field defined on a compact manifold is equal to the Euler characteristic of this manifold, here 0 for the torus T^d . This theorem is a generalization of the following simple one-dimensional property: a continuous periodic function in one variable vanishes as many times with a positive derivative as with a negative one. This example suggests how one should deal with degenerate cases, when some critical points coincide. The conclusion is that there are as many physical states with chirality $+1$ as with chirality -1 .

As a consequence of the Nielsen–Ninomiya theorem, it is necessary to violate one of the hypotheses if one wants to construct a chiral invariant lattice model. A possible way out might be the use of a random lattice (chapter 11, volume 2). Another one is

the choice

$$\tilde{F}_\mu(\mathbf{k}) = k_\mu$$

for $-\pi/a < k_\mu < \pi/a$. This function is discontinuous at the limit of the first Brillouin zone $|k_\mu| < \pi/a$, and the corresponding action is not local. In particular, any pair of sites interacts with the same strength, whatever the distance is. This introduces serious difficulties in further developments, such as the introduction of gauge interactions.

The above discussion pertains only to free fermions. It is however presumably related to the anomalies found in the continuum theories of chiral fermions interacting with gauge fields. The latter have to be compensated for in realistic models of electroweak interactions.

6.4.3 Staggered fermions

An alternative technique has also been introduced to reduce the number of species in the naive fermionic action and to keep some discrete aspects of chiral symmetry (Kogut, Susskind). Unfortunately, it still does not allow a discretization of Weyl fermions. The idea is to use Grassmannian fields with only one component (i.e. without a spinor index) and to reconstruct the physical multicomponent spinor using the various poles of its propagator. In other words, the different components of a fermion are distributed on the sites of 2^d “superlattices” with a doubled spacing and interpreted as a unique one component Grassmannian field. A version of this model reads

$$S_{\text{KS}} = -\frac{1}{2} a^{d-1} \sum_{i,\mu} (-)^{x_{i,1} + \dots + x_{i,\mu-1}} (\bar{\chi}_i \chi_{i+\hat{\mu}} + \bar{\chi}_{i+\hat{\mu}} \chi_i) + i m a^d \sum_i \bar{\chi}_i \chi_i \quad (157)$$

where the Grassmannian fields $\bar{\chi}$, χ have only one component. The technique of reconstruction of the physical spinor $\bar{\psi}$, ψ with $2^{[d/2]}$ components is suggested by the following exercise.

Diagonalize the kinetic part of the quadratic naive fermionic action (145) with respect to the spinorial indices and show then that the states can be decoupled. For that purpose, one defines unitary matrices A_i depending on the site i such that the change

of fields

$$\psi_i = A_i \chi_i, \quad \bar{\psi}_i = \bar{\chi}_i A_i^\dagger \quad (158)$$

diagonalizes simultaneously the Dirac matrices

$$A_i^\dagger \gamma_\mu A_{i+\mu} = \Delta_{i\mu} \quad (159)$$

It can be checked that

$$\begin{cases} A_i = \gamma_1^{x_{i,1}} \dots \gamma_d^{x_{i,d}} \\ \Delta_{i,\mu} = -(-1)^{x_{i,1} + \dots + x_{i,\mu-1}} \end{cases} \quad (160)$$

is a possible solution. Using the transformation (158), the action takes the form (157), except that $\bar{\chi}, \chi$ have as many components as $\bar{\psi}, \psi$; however, they are decoupled, and will remain decoupled after the introduction of gauge fields. It is thus possible to keep only one of these components, and this leads to the action (157).

Physical fermions are reconstructed as follows. Each lattice site x_i is characterized by its position y in a lattice with doubled spacing and its location η in the corresponding cluster of 2^d original sites, according to

$$x_{i,\mu} = 2y_\mu + \eta_\mu, \quad \text{with } \eta_\mu = 0 \text{ or } 1.$$

The field χ is considered as a field with 2^d components (indexed by η) on the doubled lattice

$$\begin{cases} \chi_i \equiv (-1)^{y_1 + \dots + y_d} \chi_\eta(y) \\ \bar{\chi}_i \equiv (-1)^{y_1 + \dots + y_d} \bar{\chi}_\eta(y) \end{cases} \quad (161)$$

After some straightforward algebra, the action (157) takes the form

$$S = a^d \sum_{y,\mu,\eta,\eta'} \bar{\chi}_\eta(y) \Gamma_{\eta,\eta'}^\mu \Delta_\mu \chi_{\eta'}(y) + a \bar{\chi}_\eta(y) \tilde{\Gamma}_{\eta,\eta'}^\mu \delta_\mu \chi_{\eta'}(y) \quad (162)$$

In this formula there appear two finite difference symmetric operators Δ_μ and δ_μ which tend respectively towards the first and second derivatives in the continuous limit of a vanishing lattice spacing a

$$\begin{cases} \Delta_\mu f(y) \equiv \frac{1}{4a} [f(y + \hat{\mu}) - f(y - \hat{\mu})] \xrightarrow{a \rightarrow 0} \partial_\mu f(y) \\ \delta_\mu f(y) \equiv \frac{1}{4a^2} [f(y + \hat{\mu}) - 2f(y) + f(y - \hat{\mu})] \xrightarrow{a \rightarrow 0} \partial_\mu^2 f(y) \end{cases} \quad (163)$$

The matrices Γ and $\tilde{\Gamma}$ introduced above are defined through

$$\begin{cases} \Gamma_{\eta,\eta'}^\mu \equiv 2^{-[\frac{1}{2}d]} \text{Tr}(\Gamma_\eta^\dagger \gamma_\mu \Gamma_{\eta'}) \\ \tilde{\Gamma}_{\eta,\eta'}^\mu \equiv 2^{-[\frac{1}{2}d]} \text{Tr}(\Gamma_\eta^\dagger \gamma_\mu \Gamma_{\eta'}) (\delta_{\eta',\eta-\hat{\mu}} - \delta_{\eta',\eta+\hat{\mu}}) \\ \Gamma_\eta \equiv \gamma_1^{\eta_1} \dots \gamma_d^{\eta_d} \end{cases} \quad (164)$$

starting from any given representation of the Dirac matrices. Finally, formula (162) can be rewritten in terms of physical fermions with a usual spinorial index α and an additional flavour index λ , both taking $2^{[\frac{1}{2}d]}$ different values

$$\begin{cases} \psi_y^{\alpha\lambda} \equiv 2^{-3[\frac{1}{2}d]/2} \sum_\eta (\Gamma_\eta)_{\alpha\lambda} \chi_\eta(y) \\ \bar{\psi}_y^{\alpha\lambda} \equiv 2^{-3[\frac{1}{2}d]/2} \sum_\eta (\Gamma_\eta)_{\alpha\lambda}^* \bar{\chi}_\eta(y) \end{cases} \quad (165)$$

The final form of the action is

$$\begin{aligned} S_{KS} = (2a)^d \sum_{y,\mu} [\bar{\psi}_y (\gamma_\mu \otimes 1) \Delta_\mu \psi_y + a \bar{\psi}_y (\hat{\gamma}_s^\dagger \otimes \gamma_\mu^\dagger \hat{\gamma}_s^\dagger) \delta_\mu \psi_y] \\ + (2a)^d \text{im} \sum_y \bar{\psi}_y \psi_y \end{aligned} \quad (166)$$

Its interpretation is in terms of $2^{[\frac{1}{2}d]}$ different species of physical fermions. These different flavours are coupled on the lattice by the second term of the first summation, even if this term is irrelevant in the continuous limit $a \rightarrow 0$.

In spite of all their shortcomings, these various fermionic discretizations have been used efficiently in numerically probing quantum chromodynamics in the confined regime, a problem that is still under active investigation.

Notes

The original idea of lattice gauge theory is due to K.G. Wilson, *Phys. Rev. D* **10** 2445 (1974). The corresponding action in the abelian case had been proposed by F.J. Wegner, *J. Math. Phys.* **12** 2259 (1971), as an example of a model with duality properties and no local order parameter. Some of the original papers, including the work of R. Balian and the authors, are reprinted in

C. Rebbi's book, *Lattice Gauge Theories and Monte-Carlo Simulations* (World Scientific, Singapore 1983). Numerous reviews are also available, such as those by J.B. Kogut, *Rev. Mod. Phys.* **51** 659 (1979), J.-M. Drouffe and J.-B. Zuber, *Phys. Reports* **102** 1 (1983), on theoretical aspects, M. Creutz, L. Jacobs and C. Rebbi, *Phys. Reports* **95** 201 (1983), on numerical simulations. They include a large bibliography. The book by M. Creutz, *Quarks, Gluons and Lattices* (Cambridge University Press, Cambridge 1983) offers a nice introduction to the subject. A number of rigorous properties are studied in E. Seiler's monograph *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics*, Lecture Notes in Physics **159**, Springer Verlag, Berlin (1982).

We have not discussed in the text the Hamiltonian formalism due to J. Kogut and L. Susskind, *Phys. Rev.* **D11** 395 (1975). Among variant actions, we quote the program for an improved lattice action initiated by K. Symanzik, in *Mathematical Problems in Theoretical Physics*, R. Schrader et al eds, Lecture Notes in Physics, **153**, Springer Verlag, Berlin (1982).

The quantization of Yang-Mills fields in the continuum is presented in the books of J.C. Taylor *Gauge Theories of Weak Interactions*, Cambridge University Press, (1976), and L.D. Faddeev and A.A. Slavnov, *Gauge Fields — Introduction to Quantum Theory*, Benjamin Cummings (1980). Asymptotic freedom was discovered by D.J. Gross and F. Wilczek, *Phys. Rev. Lett.* **30** 1343 (1973), and H.D. Politzer, *Phys. Rev. Lett.* **30** 1346 (1973). For the determination of the Λ_L parameter, see A. Hasenfratz and P. Hasenfratz, *Phys. Lett.* **93B** 165 (1980). R. Dashen and D.J. Gross, *Phys. Rev.* **D23** 2340 (1981) have proposed to use the background gauge. See also H. Kawai, R. Nakayama and K. Seo, *Nucl. Phys.* **B189** 40 (1981); P. Weisz, *Phys. Lett.* **100B** 331 (1981); A. Hasenfratz and P. Hasenfratz, *Nucl. Phys.* **B193** 210 (1981); A. Gonzalez-Arroyo and C.P. Korthals Altes, *Nucl. Phys.* **B205** [FS5] 46 (1982). A review on duality can be found in R. Savit, *Rev. Mod. Phys.* **52** 453 (1980). The work of S. Elitzur on the absence of order parameter is in *Phys. Rev.* **D12** 3978 (1975). Dual non local observables were introduced by G. 't Hooft, *Nucl. Phys.* **B138**, 1 (1978).

The steepest descent method is presented in the book of R.B. Dingle, *Asymptotic Expansions — Their Derivation and*

Interpretation, Academic Press, New York (1973). Systematic corrections to mean field are found in E. Brézin and J.-M. Drouffe, *Nucl. Phys.* **B200** [FS4] 93 (1982). For corrections to mean field, see H. Flyvbjerg, B. Lautrup and J.-B. Zuber, *Phys. Lett.* **110B** 279 (1982).

The topological classification of surfaces can be found in P.S. Aleksandrov, *Combinatorial topology*, Graylock 1956.

Numerous strong coupling series are available in the literature. The series given for the free energy is taken from the work of one of the authors *Nucl. Phys.* **B170** [FS1] 91 (1980). For the string tension, see G. Münster, *Nucl. Phys.* **B180** [FS2] 23 (1981); G. Münster and P. Weisz, *Nucl. Phys.* **B180** [FS2] 13 (1981); J.-M. Drouffe and J.-B. Zuber, *Nucl. Phys.* **B180** [FS2] 253, 264 (1981). Series for the glueball mass are found in N. Kimura and A. Ukawa, *Nucl. Phys.* **B205** [FS5] 637 (1982); K. Seo, *Nucl. Phys.* **B209** 200 (1982); G. Münster, *Phys. Lett.* **121B** 53 (1983). The resummation of strong coupling series in large dimension, including a discussion of a lattice model for random surfaces, is studied in J.-M. Drouffe, G. Parisi and N. Sourlas, *Nucl. Phys.* **B161** 397 (1979).

Staggered fermions are introduced in T. Banks, J. Kogut and L. Susskind, *Phys. Rev.* **D13** 1043 (1976), and their interpretation as flavoured fermions follows H. Kluberg-Stern, A. Morel, O. Napoly and B. Petersson, *Nucl. Phys.* **B220** [FS8] 447 (1983). The original proof of the no-go theorem is due to H.B. Nielsen and M. Ninomiya, *Nucl. Phys.* **B185** 20 (1981). We have patterned our presentation on the work of L.H. Karsten, *Phys. Lett.* **104B** 315 (1981). The Poincaré-Hopf theorem is standard in differential geometry; see e.g. J.W. Milnor, *Topology from the Differentiable Viewpoint* (The University Press of Virginia, Charlottesville, 1965). For a review on anomalies in the context of continuous field theory, see S.B. Treiman, R. Jackiw, B. Zumino and E. Witten, *Current Algebra and Anomalies*, World Scientific, Singapore (1985).

INDEX

A

A-D-E classification, 579, 754
 Abe, 157
 Abnormal dimension, 176
 Adjoint representation, 625
 Airy's equation, 653
 Alvarez, 780
 Amplitude ratios, 314
 Amputation, 239
 Anderson, 646
 Andrews, Baxter, Forrester, 588
 Annealed average, 802
 Anomalies, 268
 Anomalous dimension, 241, 274
 Anomaly, 516
 Anticommuting variables, 48
 Area law, 345
 Articulated diagrams, 250
 Ashkin-Teller model, 554, 594
 Asymptotic freedom, 329, 437, 727
 Atiyah-Singer index theorem, 793
 Avron, Seiler, 729
 Axial gauge, 333

B

Background gauge, 340
 Baker, Nickel, Green, Meiron, 317
 Baker-Campbell-Hausdorff
 formula, 334

Bare parameters, 237
 Baxter, 558
 Baxter, Andrews, Forrester, 579
 Becchi-Rouet-Stora (BRS), 647, 657, 717
 Belavin, Polyakov,
 Zamolodchikov, 501, 578
 Beltrami equation, 784
 Bender, Wu, 447
 Berezin, 48
 Berezinskii, 200
 Berlin-Kac, 140
 Bernal, 739
 Berry's phase, 734
 Bethe approximation, 121
 Bethe lattice, 419
 Block spin method, 168
 Bogoliubov, Parasiuk,
 Hepp, Zimmermann
 renormalization scheme,
 264
 Born amplitude, 282
 Bosonization, 97, 722
 Brownian motion, 1, 649, 778, 801
 Brézin, Le Guillou, Zinn-Justin, 311

C

Caianiello, 95

Index

Callan-Symanzik equation, 233, 273, 285, 309
 Canonical degree, 237
 Canonical dimension, 23, 241
 Cardy, 546, 601
 Cartan matrix, 622
 Cartan subalgebra, 619
 Cartan-Killing, 579, 619
 Casimir effect, 519, 564
 Casimir invariants, 583, 625, 635
 Casimir operator, 332
 Cauchy determinant, 84
 Cayley tree, 419
 Central charge, 515
 Character expansion, 374
 Cheeger, Müller, Schrader, 768
 Cheng-Wu, 88
 Chevalley, Serre, 622
 Christ, Friedberg, Lee, 738
 Classical Coulomb gas, 193
 Classical Heisenberg model, 26, 107
 Clausius-Mossotti formula, 109
 Clifford algebra, 52, 63
 Co-roots, 623
 Coleman, 197
 Configuration number, 439
 Confinement, 345
 Conformal anomaly, 773
 Conformal invariance, 502
 Conformal weights, 511
 Connected correlation functions, 239
 Contragredient form, 527
 Coordination number, 1
 Corrections to scaling laws, 301
 Correlation length, 40
 Coulomb gas, 203, 691
 Coulombic partition function, 593
 Coxeter numbers, 585, 624, 626
 Coxeter-Dynkin diagrams, 622
 Creutz, 473
 Critical amplitudes, 189

Critical exponent, 6
 Critical exponents, 125, 280
 Critical point, 40, 274
 Critical slowing down, 460, 471
 Cross ratio, 518
 Crossover exponent, 720
 Cumulant expansion, 77, 239, 412
 Curie transition, 58

D

Darboux-Christoffel kernel, 699
 David, 805
 Decimation, 176
 Deconfining transition, 354
 Dedekind's function, 528, 549, 611
 Descendant, 526
 Di Francesco, Saleur, Zuber, 577
 Diagram, 405
 Diffusion equation, 3
 Dimensional analysis, 236
 Dimensional regularization, 250
 Dimensional transmutation, 716
 Dirac equation, 93
 Dirac field, 721
 Dirac matrices, 394
 Dirichlet-Voronoi construction, 739
 Discrete Laplacian, 3
 Dominant weight, 635
 Dotsenko, 561, 563
 Dotsenko, Dotsenko, 716, 720, 728
 Dual complex, 348
 Dual Coxeter number, 624
 Dual group, 347
 Duality transformation, 180, 204
 Duality, 59, 345
 Dynamical dimension, 242
 Dyson, 310, 647, 665, 690

Index

E

Effective potential, 156, 240, 249
 Eguchi, Ooguri, 578
 Elitzur's theorem, 341
 Elliptic functions, 609
 Energy-momentum tensor, 507
 Enveloping algebra, 583
 ϵ -expansion, 235, 290, 311
 Equation of state, 166
 Equations of motion, 145
 Euclidean fields, 21
 Euler characteristic, 740, 765, 793
 Euler's formula, 705
 Euler's pentagonal identity, 549, 610
 Euler's relation, 409

F

Faddeev-Popov determinant, 785
 Faddeev-Popov ghost fields, 337, 436
 Feigin-Fuchs, 533
 Ferdinand, Fisher, 574
 Fermi levels, 714
 Feynman integrals, 246
 Feynman rules, 154, 244, 337
 Finite size effects, 478
 Fisher circles, 139
 Fisher, 233, 235, 290, 480
 Fisher-Gaunt, 157
 Fock-Bargmann space, 50, 676
 Fokker-Planck equation, 487, 649
 Free fields, 22
 Freudenthal formula, 626
 Friedan, 780, 798
 Friedan, Qiu, Shenker, 543, 592, 618, 638
 Frustrated partition function, 72, 86
 Frustration, 352, 588
 Fujikawa, 776

G

Gaudin, 699
 Gauge fields, 329
 Gauge fixing, 435
 Gauge invariance, 329
 Gauss series, 561
 Gauss sum, 612
 Gauss-Bonnet formula, 766
 Gaussian discrete model, 205
 Gaussian integrals, 22
 Gaussian model, 33, 550, 592
 Gell-Mann-Low formula, 522
 Generating functionals, 236
 Genus, 705
 Gepner, Qiu, 582
 Gepner, Witten, 581
 Glueball mass spectrum, 390
 Goddard, Kent, Olive, 636, 642
 Goldstone modes, 107, 118, 120, 127
 Goldstone's theorem, 298, 300, 435, 495
 Grading, 524
 Graph, 405
 Grassmannian integrals, 48
 Green function, 8
 Green functions, 238
 Gross-Neveu model, 490, 720
 Group characters, 331
 Group theoretical factors, 259

H

Haar measure, 427
 Hadron spectrum, 494
 Haffnian, 95
 Hall conductance, 729
 Halperin, 647
 Hard hexagon model, 558
 Hardy-Ramanujan formula, 548
 Harris criterion, 716, 720
 Hausdorff dimension, 6, 802, 808
 Heat bath algorithm, 460

Heat kernel, 332, 776, 790
 Hierarchical model, 192
 Higgs fields, 350
 High temperature expansion, 33, 421
 Highest weight vector, 526
 Hilbert–Einstein action, 768
 Hioe, Montroll, 715
 Hohenberg, 197, 219
 Homology and cohomology groups, 349
 Hopping parameter, 492
 Hopping probability, 759
 Huse, 588
 Hypergeometric equation, 561
 Hyperscaling hypothesis, 165

I

Importance sampling, 458
 Infrared fixed point, 171
 Interacting fields, 25
 Interface energy, 480
 Intersection probability, 16
 Irrelevant operators, 174, 307
 Ising model, 33, 58, 573, 605

J

Jordan–Wigner representation, 63
 Julia set, 193

K

Kac character formula, 640
 Kac determinant, 528, 532
 Kac table, 555, 559, 589
 Kac–Moody algebra, 526, 581, 619, 638
 Kadanoff, 168
 Kadanoff–Ceva, 87
 Kaufmann, 574
 Kazakov, Kostov, Migdal, 808

Killing field, 786
 Killing vector, 790, 794
 Kirchhoff's theorem, 410, 805
 Kirkwood–Yvon theory, 110
 Kosterlitz–Thouless transition, 193, 200
 Kramers–Wannier duality, 61
 Kronecker's formula, 567
 Kubo's formula, 732

L

Lagrangian, 233
 Landau levels, 675
 Landau–Ginzburg criterion, 158, 166
 Langevin equation, 486
 Laplace transform, 153
 Laplace–Beltrami operator, 332
 Lattice fermions, 393
 Lattice gauge fields, 328
 Lattices, 43
 Laughlin, 688, 729
 Lee–Yang singularity, 125, 129, 131, 547, 566, 588
 Legendre transformation, 155, 159, 239, 415
 Level spacings, 696
 Level, 581
 Liapunov exponent, 666
 Lie algebra, 330, 579, 619
 Lifshitz tail, 654
 Liouville action, 773, 779, 795
 Liouville theory, 798
 Lipatov, 310, 447
 Localization length, 670
 Loop expansion, 246
 Loop group, 626
 Low temperature expansion, 430
 Lower critical dimension, 143

M

Ma, 456
 Majorana field, 93, 554, 721
 Majority rule, 183
 Marginal operators, 174, 218, 307
 Markov process, 5, 458
 McBryan–Spencer, 222
 McCoy–Wu, 78
 Mean field approximation, 108, 352
 Meijering, 743
 Mermin–Wagner theorem, 143, 193, 197, 219
 Metropolis algorithm, 462
 Microcanonical simulations, 465
 Migdal–Kadanoff approximation, 178
 Minimal coupling, 330
 Minimal models, 546
 Minimal subtraction scheme, 337
 Modular group, 567
 Modular invariance, 564
 moduli, 786
 Monodromy, 562
 Monte Carlo renormalization group, 456, 481
 Monte Carlo sweep, 460
 Multicritical points, 317
 Möbius transformation, 510

N

n -vector model, 25, 118, 259, 427, 436, 595
 Nambu–Goto action, 781
 Neveu–Schwarz and Ramond boundary conditions, 577, 616
 Nickel, 311
 Nielsen–Ninomiya theorem, 396
 Niemeijer–Van Leeuwen, 183
 Noether current, 508
 Noether's theorem, 97

nonlinear σ -model, 26, 107, 436, 769
 Number of loops, 247

O

One particle irreducible Green functions, 239
 One-loop correction, 155
 Onsager, 48, 58, 68, 75
 Orbifold, 554, 594
 Order-disorder variables, 87
 Oscillator wavefunctions, 698

P

Padé approximants, 448
 Painlevé equations, 98
 Parisi, Sourlas, 686, 715
 Parisi, Wu Yong Shi, 486
 Pasquier, 579, 588, 595
 Path integral, 12
 Peierls' argument, 432
 Perimeter law, 345
 Perturbation theory, 242
 Perturbative renormalization, 264
 Peter–Weyl theorem, 427
 Pfaffian, 57
 Phenomenological renormalization, 224
 Pippard–Ginsberg, 165
 Planar approximation, 703
 Planar limit, 805
 Plaquette, 60, 331
 Poisson distribution, 665
 Poisson formula, 204, 580
 Polyakov loops, 469
 Polyakov, 738, 779, 780, 798
 Potts model, 176, 557, 595
 Power counting, 261
 Primary divergent diagrams, 262
 Primary field, 512
 Pseudofermion method, 494

Index

Q

Quadratic differential, 518
Quantum Chromodynamics, 329, 490
Quantum Hall effect, 675
Quasiprimary field, 512
Quenched approximation, 491

R

Radial quantization, 521
Radius of gyration, 801, 804
Random geometry, 738
Random lattice, 664, 739
Random matrices, 690
Random number generator, 463
Random potential, 647, 679
Random surfaces, 764
Random walk, 1
Real space renormalization, 162, 168
Regge, 768
Relevant operators, 174
Renormalization flow, 213, 270
Renormalization group, 270
Renormalized Green functions, 265
Replica trick, 658
Reproducing kernel, 684
Resolvent, 680
Ricatti equation, 649
Riemann's ζ function, 569
Riemann-Christoffel symbol, 785
Roots, 620
Roughening transition, 91, 205, 385

S

Scaling dimension, 505
Scaling fields, 174
Scaling hypothesis, 165
Scaling laws, 162

Scaling regime, 89
Schwarzian derivative, 517, 551
Schwinger functions, 238
Schwinger term, 631
Second order phase transition, 70
Self-avoiding walk, 31
Semicircle law, 690, 708
Shankar, 721, 729
Simply laced Lie algebra, 579
Sine-Gordon model, 206
Six vertex model, 554
Slater determinant, 698
Solid on solid (SOS) model, 82, 205, 386
Spherical model, 140
Spin wave approximation, 197
Spontaneous magnetization, 61, 72
Spontaneous symmetry breaking, 112
Staggered Kogut-Susskind fermions, 399
Star-triangle transformation, 183
Stochastic quantization, 486
String tension, 344, 385
Strong coupling expansion, 371
Strong coupling series, 33
Sugawara, 634
Sum over paths, 9, 27
Sum over surfaces, 780
Super-renormalizability, 262
Super-Schwarzian derivative, 615
Superconformal algebra, 613
Superficial convergence, 262
Superfield, 683, 718
Supersymmetry, 682
Surface tension, 85
Swendsen, 456, 482
Szegő's lemma, 74, 76

T

't Hooft loop, 352
't Hooft, 703
Tadpole graph, 244
Teichmüller space, 786
Thermalization time, 460
Thermodynamic limit, 23
 θ -series, 609
Toeplitz matrix, 74
Topological complex, 345
Transfer matrix, 37, 62, 136
Tree diagrams, 154, 248
Triangulated manifold, 765
Tricritical Ising model, 618

U

Universality, 33
Upper critical dimension, 144

V

Vacuum diagrams, 244
Vandermonde determinant, 77
Verma module, 526
Vertex functions, 239
Vertex operator, 554
Villain action, 204, 332
Virasoro algebra, 521, 523
Virasoro characters, 547
Virial theorem, 663
von Neumann-Wigner theorem, 734
Vortices, 197, 201, 553

W

Ward identity, 145, 508, 513
Ward-Slavnov identity, 337
Watson's quintuple product, 611
Wavefunction renormalization, 264
Wegner, 679
Weierstrass function, 612
Weights, 624
Weinberg's theorem, 274
Wess-Zumino-Witten model, 626
Weyl character formula, 625
Weyl group, 622
Weyl invariance, 782
Wick ordering, 247, 536
Wick's theorem, 24, 57, 94, 99, 243, 434
Wigner, 647, 690
Wilson fermions, 396
Wilson loop, 344, 765
Wilson, 233, 235, 290, 311, 328, 456, 482
Wilson's action, 331
Wulff construction, 89

X

XY-model, 193, 594

Y

Yang, 75
Yang, Mills, 328, 335

Z

Z_n clock model, 474
Zamolodchikov, 436, 587, 808
 ζ -regularization, 568

and let e_F stand for the Fermi level, i.e. the highest occupied one. We have

$$\begin{aligned} N^2 E(g) &= \sum_{1 \leq k} e_k \theta(e_F - e_k) \\ N &= \sum_{1 \leq k} \theta(e_F - e_k) \end{aligned} \quad (300)$$

These expressions are exact for any N (it would be more appropriate to write $E_N(g)$ instead of $N^2 E(g)$ for finite N , since the N^2 behaviour is only asymptotic). Of course this involves the values e_k which are not directly known. Luckily, for N large, the dominant contribution to the expressions (300) comes from very excited one-body states, and a semiclassical approximation becomes sufficient. In other words, we substitute to the above discrete sums, integrals over classical phase space (λ, p) according to

$$\begin{aligned} N^2 E(g) &= N e_F \\ &\quad - \int \frac{d\lambda dp}{2\pi} \left(e_F - \frac{1}{2} p^2 - \frac{1}{2} \lambda^2 - \frac{g}{N} \lambda^4 \right) \\ &\quad \times \theta \left(e_F - \frac{1}{2} p^2 - \frac{1}{2} \lambda^2 - \frac{g}{N} \lambda^4 \right) \\ N &= \int \frac{d\lambda dp}{2\pi} \theta \left(e_F - \frac{1}{2} p^2 - \frac{1}{2} \lambda^2 - \frac{g}{N} \lambda^4 \right) \end{aligned} \quad (301)$$

Let us integrate over p and rescale λ and e_F as $\sqrt{N}\lambda$, $e_F \rightarrow N\epsilon$. We obtain a pair of equations giving the solution in parametric form

$$\begin{aligned} E(g) &= \epsilon - \int_{-\infty}^{+\infty} \frac{d\lambda}{3\pi} [2\epsilon - \lambda^2 - 2g\lambda^4]^{\frac{3}{2}} \theta(2\epsilon - \lambda^2 - 2g\lambda^4) \\ 1 &= \int_{-\infty}^{+\infty} \frac{d\lambda}{2\pi} [2\epsilon - \lambda^2 - 2g\lambda^4]^{\frac{1}{2}} \theta(2\epsilon - \lambda^2 - 2g\lambda^4) \end{aligned} \quad (302)$$

which are elliptic integrals. The second equation expresses stationarity of E with respect to ϵ for fixed g . The equations (302) offer a striking example of the simplifications brought about by the planar approximation.

When we eliminate ϵ , we find E as a function of g , analytic in the vicinity of the origin. This is in contradistinction with the properties of the exact solution for which $g = 0$ is an essential

singularity corresponding to a quantum instability by tunnelling for $g < 0$. In the present approximation, the nearest singularity is for a negative value of g such that the Fermi level reaches the maximum of the potential, i.e.

$$g_c = -\frac{\sqrt{2}}{3\pi} \quad \epsilon = -\frac{1}{16g_c} \quad (303)$$

One would not expect that the planar approximation gives very sensible results for finite N and *a fortiori* for $N = 1$. It comes as a surprise to see that the comparison, as shown in the table, with the numerical data (called E_{exact}) obtained for the anharmonic oscillator (Hioe and Montroll (1975)), is quite favourable over a large range of couplings

g	E_{planar}	E_{exact}
0.01	0.505	0.507
0.1	0.547	0.559
0.5	0.651	0.696
1.0	0.740	0.804
50	2.217	2.500
1000	5.915	6.694

(304)

For large g , the agreement is less satisfactory. The behaviour of the planar approximation, as deduced from equation (302), is

$$E_{\text{planar}}(g) \underset{g \rightarrow \infty}{\sim} \frac{3}{7} \left(\frac{3}{2\pi^{\frac{1}{2}}} \Gamma^2\left(\frac{3}{4}\right) \right)^{\frac{4}{3}} g^{\frac{1}{3}} = 0.58993 g^{\frac{1}{3}} \quad (305)$$

while the exact solution behaves as

$$E_{\text{exact}}(g) \underset{g \rightarrow \infty}{\sim} 0.66799 g^{\frac{1}{3}} \quad (306)$$

10.5 Spin systems with random interactions

We present two further applications of anticommuting variables to disordered systems. The first is due to Parisi and Sourlas. Even though, in the final analysis, it does not allow a complete quantitative understanding of the behaviour of the system to be

studied, close to its upper critical dimension, namely a spin model with a random, symmetry breaking, external magnetic field, for reasons to be clarified below, it gives a simple explanation for an observation made in the framework of perturbation theory. This relates the behaviour of the disordered system to a pure one in a lower dimension. In the second subsection, we present the work of Dotsenko and Dotsenko pertaining to another extreme situation. The goal is to look for the effect of weak bond impurities close to criticality, when disorder is expected to have a marginal effect according to a criterion due to Harris. Although the treatment may be open to some criticism, both the method of analysis as well as the results are interesting. These difficulties reflect once more the subtleties of disordered systems.

10.5.1 Random external field and dimensional transmutation

Consider an Ising model in the presence of an external random field, in the continuous field φ^4 version, and assume a Gaussian distribution for the external field with zero-range correlations. Physical quantities are to be averaged over the quenched disorder. The thermal average is denoted $\langle A \rangle$ and its average over disorder $\overline{\langle A \rangle}$. With h the random external field, we write the free energy as

$$F(h) = \ln \int \mathcal{D}\varphi \exp - \int d^d x \{ \mathcal{L}(x) + h(x)\varphi(x) \} \quad (307)$$

$$\bar{F} = \int \mathcal{D}h F(h) \exp - \frac{1}{2} \int d^d x h^2(x)$$

with

$$\mathcal{L}(x) = \frac{1}{2}(\partial\varphi)^2 + V(\varphi) \quad V = \frac{1}{2}m^2\varphi^2 + g\varphi^4 \quad (308)$$

To compute $\bar{F}$ one can use the replica trick, which replaces the computation of $\overline{F(h)}$ by the average $\overline{Z(h)^n}$, where $Z(h)$ is the partition function, as the number of replicas goes to zero. In a perturbative expansion the most divergent diagrams in the infrared limit are those involving the largest possible number of h^2 insertions. Resumming these most divergent diagrams leads to the conclusion that $F(h)$ is given by its saddle point approximation, with $\bar{F}(h)$ obtained by averaging this approximate value over

disorder. This introduces the mean field φ_h , a solution of

$$-\Delta\varphi_h(x) + V'(\varphi_h(x)) + h(x) = 0 \quad (309)$$

In this approximation the two-point function for instance is given by

$$\begin{aligned} \overline{\langle \varphi(x)\varphi(0) \rangle} &= \int \mathcal{D}h \varphi_h(x)\varphi_h(0) \exp - \frac{1}{2} \int d^d x h^2(x) \\ &= \int \mathcal{D}h \mathcal{D}\varphi |\det(-\Delta + V''(\varphi))| \varphi(x)\varphi(0) \\ &\quad \prod_x \delta(-\Delta\varphi(x) + V'(\varphi(x)) + h(x)) \exp - \frac{1}{2} \int d^d x h^2(x) \end{aligned} \quad (310)$$

In the second expression we have used the definition of $\varphi_h(x)$ deduced from the mean field equation (309) in an indirect way short of exhibiting $\varphi_h(x)$ in closed form. The product of δ -functions is accompanied by a Jacobian, the absolute value of a determinant involving the second variation of the Lagrangian with respect to $\varphi(x)$. If the correspondence $h \leftrightarrow \varphi$ were one-to-one, the absolute value sign could be omitted, since $\det(-\Delta + V''(\varphi))$ would never vanish. This is implicitly assumed in a perturbative expansion. We will also use here this assumption, without a serious justification, as a means of resumming the leading logarithmic terms. Many disordered system are such that even the mean field approximation is of great complexity (this is for instance the case of spin glasses) and in such cases a similar assumption can be unwarranted.

After these apologies, we follow Parisi and Sourlas, ignore the absolute value sign in the determinant and proceed to investigate the consequences. The product of δ -functions is replaced by a parametric representation over a (pure imaginary) conjugate field $\alpha(x)$, and the determinant is expressed as a Grassmannian integral over a fermionic field $\psi, \bar{\psi}$. This gives the representation

$$\begin{aligned} \overline{\langle \varphi(x)\varphi(0) \rangle} &= \int \mathcal{D}\varphi \mathcal{D}\alpha \mathcal{D}(\psi\bar{\psi}) \varphi(x)\varphi(0) \exp - \int d^d x \mathcal{L}_{\text{eff}}(x) \\ \mathcal{L}_{\text{eff}}(x) &= \bar{\psi}(-\Delta + V''(\varphi))\psi + \alpha(-\Delta\varphi + V'(\varphi)) - \frac{1}{2}\alpha^2 \end{aligned} \quad (311)$$

We expect that this Lagrangian satisfies a BRS supersymmetry. Let $\bar{a}$ denote an infinitesimal anticommutative parameter, and ϵ a d -dimensional vector, the transformation in question is of the type

$A \rightarrow A + \delta A$ for the set of fields $A \equiv (\varphi, \psi, \bar{\psi}, \alpha)$, with $\delta A = \bar{a}sA$

$$\begin{aligned} s\varphi &= -\varepsilon \cdot x\psi & s\alpha &= 2\varepsilon \cdot \partial\psi \\ s\psi &= 0 & s\bar{\psi} &= \varepsilon \cdot x\alpha + 2\varepsilon \cdot \partial\varphi \end{aligned} \quad (312)$$

The variation of the action is the integral of a total derivative

$$\delta \int d^d x \mathcal{L}_{\text{eff}}(x) = 2\bar{a} \int d^d x \varepsilon \cdot \partial \{ \psi [-\Delta\varphi + V'(\varphi)] \} \quad (313)$$

which can be set equal to zero by requiring that the quantity in brackets vanishes at infinity.

Show that

$$s^2\varphi = s^2\psi = s^2\alpha = 0 \quad s^2\bar{\psi} = -2\varepsilon^2\psi \quad (314)$$

Could one modify the formalism to enforce $s^2 = 0$?

It is convenient to introduce a superfield Φ function of x and of two anticommuting variables θ and $\bar{\theta}$

$$\Phi(x, \theta, \bar{\theta}) = \varphi(x) + \bar{\theta}\psi(x) + \bar{\psi}(x)\theta + \theta\bar{\theta}\alpha(x) \quad (315)$$

We normalize the integration over θ and $\bar{\theta}$ by requiring that

$$\int d\bar{\theta} d\theta \theta\bar{\theta} = 1 \quad (316)$$

In superspace, we define the generalized Laplacian

$$\Delta_S = \Delta + \frac{\partial}{\partial\bar{\theta}} \frac{\partial}{\partial\theta} \quad (317)$$

where derivatives act to the left. One verifies that the effective action in the path integral (311) takes the form

$$\begin{aligned} S_{\text{eff}} &= \int d^d x d\bar{\theta} d\theta \left\{ -\frac{1}{2} \Phi \Delta_S \Phi + V(\Phi) \right\} \\ &= \int d^d x \left\{ \bar{\psi} (-\Delta + V''(\varphi)) \psi + \alpha (-\Delta\varphi + V'(\varphi)) - \frac{1}{2} \alpha^2 \right\} \end{aligned} \quad (318)$$

We leave it as an exercise to the reader to check this rather surprising fact, which implies invariance under rotations in superspace.

The interpretation of the two Grassmann coordinates is that they amount to subtracting (rather than to adding) two dimensions to d -dimensional space, in rotationally invariant integrals. To see this, let us look at an integral over a configuration variable

$(x, \theta, \bar{\theta})$. Let f be an algebraic function with a sufficiently fast decrease at infinity. Consider the integral

$$-\frac{1}{\pi} \int d^d x d\bar{\theta} d\theta f(x^2 + \theta\bar{\theta}) = -\frac{1}{\pi} \int d^d x f'(x^2)$$

Let us use polar coordinate with $x^2 = r^2$, $S_d = 2\pi^{d/2}/\Gamma(\frac{1}{2}d)$ the area of the unit sphere in d -dimensional space. We find by integration by parts for $d > 2$

$$\begin{aligned} -\frac{1}{\pi} \int d^d x f'(r^2) &= -\frac{S_d}{\pi} \int_0^\infty r^{d-1} dr f'(r^2) \\ &= \frac{S_d}{2\pi} \left(\frac{1}{2}d - 1\right) \int_0^\infty dr^2 (r^2)^{\frac{1}{2}d-2} f(r^2) = \int d^{d-2} x f(x^2) \end{aligned} \quad (319)$$

This verifies the claim for integrals over a single configuration variable. It extends to several variables, as the interested reader will find out, and leads to the astonishing dimensional reduction by two units, observed perturbatively and justified by this global analysis. The conclusion of Parisi and Sourlas is that the supersymmetric d -dimensional model is equivalent to a purely bosonic one in $d - 2$ dimensions. For the random field model, it applies at best close to its upper critical dimension (i.e. six for the disordered system, equivalent to a φ^3 theory). The behaviour in lower dimensions is not fully understood, and we shall not explore it further.

10.5.2 The two-dimensional Ising model with random bonds

We turn to another marginal effect of disorder on critical behaviour. Consider a system close to a continuous transition. Let θ denote the deviation from the critical temperature in reduced units. A relevant local operator K of dimension $\Delta \leq d$ contributes to the action an extra term of the form

$$\delta S = \int d^d x g(x) K(x) \quad (320)$$

with a coupling g of dimension $L^{\Delta-d}$. In the absence of perturbation the correlation length ξ behaves as $\theta^{-\nu}$. From scaling arguments, if we turn on a constant g , the singular part in the free energy per unit volume is of the form

$$F(\theta, g) = \theta^{2-\alpha} f(g/\theta^k) \quad (321)$$

with a nonnegative crossover exponent

$$k = \nu(d - \Delta) \quad (322)$$

The scaling function is regular at the origin (this is the purpose of extracting the factor $\theta^{2-\alpha}$). When g is a quenched random external quantity, it is natural to expect that the would-be singular part of the total free energy takes a form generalizing equation (321)

$$\ln Z(\theta, g(\mathbf{x})) = \theta^{2-\alpha} \varphi(g(\mathbf{x})/\theta^k) \quad (323)$$

where φ is a regular functional close to a vanishing argument. Assume that $g(\mathbf{x})$ has a zero mean. We have approximately

$$\overline{\ln Z(\theta, g(\mathbf{x}))} = \theta^{2-\alpha} \varphi_0 + \frac{1}{2} \theta^{2-\alpha-2k} \int \int d^d \mathbf{x}_1 d^d \mathbf{x}_2 \overline{g(\mathbf{x}_1)g(\mathbf{x}_2)} \varphi_2(\mathbf{x}_1, \mathbf{x}_2) + \dots \quad (324)$$

If the integral on the r.h.s. is well-defined, the leading effect of disorder is expected to be of order $\theta^{2-\alpha-2k}$ with an exponent smaller or equal to the one pertaining to the pure system. The above argument breaks down when this exponent becomes negative

$$2 - \alpha - 2k < 0 \quad (325a)$$

or, equivalently, by using hyperscaling, i.e. $2 - \alpha = \nu d$, when

$$\Delta < \frac{1}{2}d \quad (325b)$$

as a naive perturbative expansion would suggest. This is the criterion due to Harris (1974) on the relevance of weak disorder.

It is therefore of interest to study a marginal situation corresponding to $\Delta = \frac{1}{2}d$. An example is offered by weak fluctuations of the bond couplings in a two-dimensional Ising model. The operator K is identified with the energy operator of dimension $\Delta = 1$. The marginality of disorder means that the associated field theoretic model will be just renormalizable. It turns out, as shown by Dotsenko and Dotsenko, that this model – first introduced in another context by Gross and Neveu – is infrared free in the regime of interest. As we increase the scale, the effect of disorder is weaker and weaker, to the extent that one can ascertain the leading large-distance behaviour in thermodynamic quantities. Even though some aspects of the derivation are still controversial,

in particular the delicate use of the replica trick, the underlying field theoretic model is of interest in itself.

We start directly from the continuous version as a free real Majorana field theory, with a space dependent mass term $m(\mathbf{x})$ fluctuating around a zero mean value. We are interested in finding the average of the logarithm of the partition function Z_M (the index M referring to Majorana)

$$Z_M = \int \mathcal{D}(\psi, \bar{\psi}) \exp -S_M \quad (326a)$$

$$S_M = \int \frac{d^2 x}{2\pi} \{ \psi \bar{\partial} \psi - \bar{\psi} \partial \bar{\psi} + m(\mathbf{x}) \bar{\psi} \psi \} \quad (326b)$$

There are several possible treatments. We choose to follow a presentation due to Shankar, which has the merit of succinctly exhibiting the main phenomenon. Since we are only interested in finding the effect on the free energy, we might as well double it, or equivalently square Z_M , to obtain a complex Dirac version, therefore introducing at first an $O(2)$ symmetry. Denote by $(\psi_1, \bar{\psi}_1), (\psi_2, \bar{\psi}_2)$ the two copies of Majorana fields and define the two component Dirac fields through

$$u = \begin{pmatrix} \psi_1 + i\psi_2 \\ -\psi_1 - i\psi_2 \end{pmatrix} \quad \bar{u} = (\bar{\psi}_1 - i\bar{\psi}_2, \bar{\psi}_1 + i\bar{\psi}_2) \quad (327)$$

An extra factor σ_3 has been inserted in taking the adjoint in such a way that the Dirac operator reads

$$D_m = \begin{pmatrix} \frac{1}{2}m & \partial \\ \bar{\partial} & \frac{1}{2}m \end{pmatrix} \quad (328)$$

We shall freely interchange the notations $\mathbf{x}, (z, \bar{z})$ or even z alone, to denote the coordinates of a point. Thus

$$Z_D = Z_M^2 = \int \mathcal{D}(u, \bar{u}) \exp -S_D \quad (329a)$$

$$S_D = \frac{1}{2\pi} \int d^2 x \bar{u} D_m u \quad (329b)$$

The appended index m on the Dirac operator is to remind us that the mass term varies in space. The evaluation of Z_D is in principle straightforward being the determinant of the operator D_m . The delicate question is as usual the role of modes close (or equal) to zero. We shall assume that they do not compromise the treatment

discussed below. This is where the present approach is open to criticism, but we proceed nevertheless.

Before averaging over $m(x)$ let us recall the bosonization formulas of the Dirac field (chapter 2). They arise from the existence of an $O(2)$ symmetry between the two copies of Majorana fields, or in terms of the complex Dirac field of the invariance under

$$u \rightarrow e^{i\tau} u \quad \bar{u} \rightarrow \bar{u} e^{-i\tau} \quad (330)$$

Noether's theorem leads to a conserved current

$$j^z = \frac{1}{2i} \bar{u} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} u \quad j^z = \frac{1}{2i} \bar{u} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} u \quad (331a)$$

$$\partial_x j^z + \partial_z j^x = 0 \quad (331b)$$

Accordingly, in the massless theory, we represent the currents in terms of a free massless field φ , as

$$j^z = \bar{\psi}_2 \bar{\psi}_1 \rightarrow \sqrt{2} \bar{\partial} \varphi \quad j^x = \psi_1 \psi_2 \rightarrow -\sqrt{2} \partial \varphi \quad (332)$$

These formulas have to be understood *cum grano salis*. They mean that the correlation functions computed in the corresponding theories do agree.

Recall that for $m = 0$, with a, b referring to points x_a, x_b

$$\langle \psi_a \psi_b \rangle = \frac{1}{z_{ab}} \quad \langle \bar{\psi}_a \bar{\psi}_b \rangle = -\frac{1}{\bar{z}_{ab}} \\ \langle \varphi_a \varphi_b \rangle = -\frac{1}{2} \ln z_{ab} \bar{z}_{ab}$$

We verify that the two-point functions do agree

$$\langle j_a^z j_b^z \rangle = -\frac{1}{\bar{z}_{ab}^2} = 2 \frac{\partial}{\partial \bar{z}_a} \frac{\partial}{\partial \bar{z}_b} \langle \varphi_a \varphi_b \rangle \\ \langle j_a^x j_b^x \rangle = -\frac{1}{z_{ab}^2} = 2 \frac{\partial}{\partial z_a} \frac{\partial}{\partial z_b} \langle \varphi_a \varphi_b \rangle$$

Check the agreement of higher correlations.

The fluctuating field $m(x)$ is coupled to the energy operator which can also be represented in bosonic form as

$$\frac{1}{2} \bar{u} u = \bar{\psi}_1 \psi_1 + \bar{\psi}_2 \psi_2 \rightarrow 2\Lambda \cos \sqrt{2} \varphi = 2\varepsilon \quad (333)$$

The ultraviolet cutoff $\Lambda = a^{-1}$, which takes care of the relative dimension of the operators, compensates for the absence of Wick

ordering since

$$: e^{i\alpha\varphi} := \Lambda^{\frac{1}{2}\alpha^2} e^{i\alpha\varphi} \quad (334)$$

One of the virtues of the bosonic formulation is that, while in fermionic terms the spin field is a highly nonlocal operator, it takes a local form when expressed in terms of φ in the following sense. Taking the product of spins $\sigma_1 \sigma_2$ in the two Majorana copies, we have the correspondence

$$\sigma_1 \sigma_2 \rightarrow \sqrt{2} \Lambda^{\frac{1}{2}} \sin \varphi / \sqrt{2} \quad (335)$$

allowing us to compute squares of correlations in bosonic form. The prefactor $\Lambda^{\frac{1}{2}}$ is inserted to remedy for the absence of normal ordering and emphasises the dimension of the operator.

We are now prepared to take averages over $m(x)$. Suppose the latter fluctuates around a mean value m . Instead of averaging $\ln Z_M$ (or $\frac{1}{2} \ln Z_D^2$) let us rather compute the specific heat at $m = 0$, i.e. differentiate twice the average over m then set $m = 0$. This involves the integral over the two-point function of the energy operator ε i.e. the average of

$$\langle \varepsilon(x) \varepsilon(x') \rangle = Z_B^{-1} \int \mathcal{D}(\varphi) \Lambda^2 \cos \sqrt{2} \varphi(x) \cos \sqrt{2} \varphi(x') \exp -S_B \quad (336a)$$

$$S_B = S_B^{(0)} + S_B^{(1)} \quad (336b)$$

$$S_B^{(0)} = \frac{1}{\pi} \int d^2 x \partial \varphi \bar{\partial} \varphi$$

$$S_B^{(1)} = \frac{1}{\pi} \int d^2 x m(x) \Lambda \cos \sqrt{2} \varphi \quad (336c)$$

For short we write $:\cos \alpha \varphi:$ for $\Lambda^{\alpha^2/2} \cos \alpha \varphi$. We have still to characterize the probability law obeyed by $m(x)$. The only case amenable to a complete analysis is a Gaussian distribution with zero-range correlations

$$\overline{O(m)} = \frac{1}{2} \int \frac{\mathcal{D}m O(m(x)) \exp -(1/2g^2) \int d^2 x m^2(x)}{\int \mathcal{D}m \exp -\frac{1}{2} \int d^2 x m^2(x)} \quad (337a)$$

i.e. such that

$$\overline{m(x_1) m(x_2)} = g^2 \delta^{(2)}(x_1 - x_2) \quad (337b)$$

Hence g^2 measures the intensity of disorder. To take the required average we have two options. Either to introduce n replicas of the bosonic field and let n go to zero, or to associate a Grassmannian partner. Results are formally the same and we choose the first option. Therefore

$$\langle \varepsilon(x) \varepsilon(x') \rangle = \lim_{n \rightarrow 0} \int \mathcal{D}(\varphi) : \cos \sqrt{2} \varphi^1 : : \cos \sqrt{2} \varphi^1 : \exp -S(\varphi) \quad (338a)$$

$$S(\varphi) = \int d^2x \frac{1}{\pi} \left\{ \sum_c \partial \varphi^c \bar{\partial} \varphi^c - \frac{g^2}{2\pi} \left(\sum_c : \cos \sqrt{2} \varphi^c : \right)^2 \right\} \quad (338b)$$

The replica index c runs from 1 to n . If instead of using bosonic variables we had performed a similar treatment on the original Dirac action, we would have obtained the Gross-Neveu model with a symmetry group $O(2n)$. This symmetry is not easily expressed in the present context. The complete action with $m(x)$, u^c , $\bar{u}^c$, is the easiest one on which to carry completely the renormalization program. At any rate, equation (338) indicates that, apart from the limit $n \rightarrow 0$, we deal with a renormalizable theory. The obvious question is to investigate its infrared behaviour and to figure out how the effect of disorder, namely the effective g^2 , behaves in this limit. An interesting aspect of the Gross-Neveu model is that it admits an exact treatment when $n = 1$, in which case its critical behaviour is one of those investigated in chapter 9 when dealing with theories with unit central charge and continuously varying exponents. To see this, observe that according to equations (332) and (333) we have, for a single component Dirac field, at a formal level, for the square of the current

$$\frac{1}{2} j^z j^z = \frac{1}{2} \bar{\psi}_2 \bar{\psi}_1 \psi_1 \psi_2 = \frac{1}{4} (\bar{\psi}_1 \psi_1 + \bar{\psi}_2 \psi_2)^2 \quad (339)$$

Interpreted in bosonic language, in terms of the corresponding correlations, this means that we have the equivalence

$$-\partial \varphi \bar{\partial} \varphi \leftrightarrow \left(: \cos^2 \sqrt{2} \varphi : \right)^2 \quad (340)$$

This can also be confirmed by a more careful renormalization analysis which implies that the diagonal terms in the interaction

Lagrangian (338b) can be traded for a modified kinetic term as

$$S(\varphi) = \frac{1}{\pi} \int d^2x \left(1 + \frac{g^2}{2\pi} \right) \sum_c \partial \varphi^c \bar{\partial} \varphi^c - \frac{g^2}{2\pi} \sum_{c \neq d} : \cos \sqrt{2} \varphi^c : : \cos \sqrt{2} \varphi^d : \quad (341)$$

When the number of components is equal to unity, the model is obviously equivalent to a massless free field theory where the field φ is rescaled by a factor $(1 + g^2/2\pi)^{1/2}$. This innocent looking modification alters of course the dimension of any operator of the form $K_\alpha = : \cos \alpha \varphi :$, from $\Delta_\alpha = \frac{1}{2} \alpha^2$ to $\Delta_\alpha(g) = \frac{1}{2} \alpha^2 (1 + g^2/2\pi)^{-1}$, thus leading to exponents depending continuously on g . For general n , the associated β -function has to vanish when $n = 1$, meaning that it has at least a factor $(n - 1)$. As we shall confirm below, this factor appears to the first power, hence $\beta(g)$ changes sign when n crosses unity. As a consequence, close to $g = 0$ the infrared and ultraviolet properties are interchanged. While for $n > 1$ the model is ultraviolet asymptotically free, with a spontaneously broken discrete symmetry through a dynamical generation of a fermion mass, in the opposite situation $n < 1$, we have infrared asymptotic freedom, the favourable circumstance for our present purposes.

Let us now proceed to the required computation. On the one hand, using the above observations and the form (341) of the action, we can rescale the fields φ^c according to

$$(1 + g^2/2\pi)^{1/2} \varphi^c = \tilde{\varphi}^c \quad (342)$$

Set

$$\alpha = \sqrt{\frac{2}{1 + g^2/2\pi}} \quad (343)$$

and recall that the dot prescription on exponentials implies an appropriate power of Λ . This leads to the action

$$S = \frac{1}{\pi} \int d^2x \left\{ \sum_c \partial \tilde{\varphi}^c \bar{\partial} \tilde{\varphi}^c - \frac{1}{2\pi} \left[g \Lambda^{1-1/(1+g^2/2\pi)} \right]^2 \sum_{c \neq d} : \cos \alpha \tilde{\varphi}^c : : \cos \alpha \tilde{\varphi}^d : \right\} \quad (344)$$

More accurately, we should have introduced an arbitrary finite scale μ to normalize the added power of Λ , which to leading order in g^2 should be interpreted as $(\Lambda/\mu)^{1-1/(1+g^2/2\pi)} \sim 1 + (g^2/2\pi) \ln(\Lambda/\mu)$. This first renormalization is however insufficient to insure a finite behaviour of all correlations, in particular of the energy-energy correlation (338). An additional renormalization is necessary for the operator $\sum_{a \neq d} : \cos \alpha \tilde{\varphi}^c :: \cos \alpha \tilde{\varphi}^d :$ as opposed to its individual factors. According to equation (344), the dimension of the interaction has been shifted from 2 to $2/(1+g^2/2\pi)$. But again, this it to be interpreted in a perturbative sense though an expansion in g^2 , and is justified *a posteriori* in that only the limit $g^2 \rightarrow 0$ will be of interest.

To see the additional divergences it is sufficient to investigate the perturbative expansion of the $\epsilon\epsilon$ correlation. We can carry this computation in two ways which can be usefully compared. Either we use the form (344), the $\tilde{\varphi}$ field and the corresponding Wick prescription. This requires an inequivalent treatment of the diagonal and off-diagonal terms in $\langle \epsilon^a \epsilon^b \rangle$. Or we retreat to the original more symmetrical expression (338b), with the field φ and only after understanding the first nontrivial counterterm do we apply the tricks eliminating the diagonal interaction. Indeed for $n \neq 1$, the value of g^2 appearing in the correspondence between φ and $\tilde{\varphi}$ is affected by renormalization. We choose to present the unsymmetrical version and urge the reader to follow also the alternative route, or else to return to the fermionic version. Although we only look at the off-diagonal terms in the $\langle \epsilon\epsilon \rangle$ expansion, it is also instructive to carry out the same discussion for the diagonal ones. Let us write

$$\begin{aligned} \epsilon^a &= \Lambda^{1-1/(1+g^2/2\pi)} E^a \\ E^a &= : \cos \alpha \tilde{\varphi}^a : \end{aligned} \quad (345)$$

Before taking the limit $n \rightarrow 0$, the averages can be written $\langle E^a(x_1) E^b(x_2) \rangle$. It is simpler to use the Feynman rules directly in configuration space, with

$$\langle E^a(x_1) E^b(x_2) \rangle = \delta^{ab} f(x_1 - x_2) \quad f(x) = \frac{1}{2|x|^{\alpha^2}} \quad (346)$$

Choose $a \neq b$ and compute the successive perturbative contributions

$$\langle E^a(x_1) E^b(x_2) \rangle = 2 \left[\frac{g^2}{2\pi^2} \left(\frac{\Lambda}{\mu} \right)^{2-\alpha^2} \right] f(x_1 - x_2)$$

$$\begin{aligned} &+ 4 \left[\frac{g^2}{2\pi^2} \left(\frac{\Lambda}{\mu} \right)^{2-\alpha^2} \right]^2 (n-2) \\ &\int \int d^2x d^2y f(x_1 - x) f(x_2 - y) f(x - y) + \dots \end{aligned} \quad (347)$$

The factor $(n-2)$ is specific to the off-diagonal correlation as can be seen from the interaction in equation (344). A subtraction takes care of the additional divergence if we set

$$\frac{g^2(\Lambda)}{2\pi^2} \left(\frac{\Lambda}{\mu} \right)^{2-\alpha^2} = \frac{g_R^2}{2\pi^2} - 2 \left(\frac{g_R^2}{2\pi^2} \right)^2 (n-2) \pi \frac{\Lambda^{\alpha^2-2} - \mu^{\alpha^2-2}}{\alpha^2 - 2} + \dots \quad (348)$$

which arises from the logarithmic divergent integral

$$\int_{\Lambda^{-1} < |x| < \mu^{-1}} d^2x f(x) = \pi \int_{\Lambda^{-1}}^{\mu^{-1}} \frac{r dr}{r^{\alpha^2}} = \pi \frac{\Lambda^{\alpha^2-2} - \mu^{\alpha^2-2}}{\alpha^2 - 2} \simeq \ln \frac{\Lambda}{\mu}$$

To leading order in g_R^2 ,

$$g^2(\Lambda) = g_R^2 \left[1 - 2(n-1) \frac{g_R^2}{2\pi} \ln \frac{\Lambda}{\mu} + \dots \right] \quad (349)$$

The two ingredients have conspired to change $n-2$ into the expected $n-1$. It is interesting to compare with what happens to the diagonal correlation, where a similar term (proportional to $(n-1)$ this time) has to be compensated by additional diagonal counterterms in the action, the latter being needed to introduce the correct coupling in the initial step rescaling φ to $\tilde{\varphi}$.

To leading order, the flow in coupling space close to the origin reads

$$\Lambda \frac{d}{d\Lambda} g(\Lambda) \equiv \beta(g(\Lambda)) = -(n-1) \frac{g^3(\Lambda)}{2\pi} \quad (350)$$

with the consequence that the theory is asymptotically free for $n > 1$ in the ultraviolet regime. For $n < 1$, the conclusion is reversed, at large distance the effective g decreases and perturbative computations become relevant.

In $\beta(g)$, not only is the coefficient of the g^3 term universal, but so is also the coefficient of g^5 . According to Wetzel, one has

$$\beta(g) = -(n-1) \frac{g^3}{2\pi} + (n-1) \frac{g^5}{(2\pi)^2} + \dots \quad (351)$$

The r.h.s. vanishes again when $n = 1$.

The situation is similar to that with the φ^4 theory in dimension four presented in chapter 5, section 4.2 (volume 1). We can adapt the discussion given there, in particular equation (5.248), where one computes the specific heat, i.e. the integral of (338a). The most singular term arises again from the inhomogeneous piece called $b(g)$ in chapter 5, to lowest order a g -independent constant, proportional to the $\ln \Lambda$ term in the integral $\int d^2x \langle \varepsilon(x) \varepsilon(0) \rangle_0$ pertaining to the pure system. The specific heat singular term denoted $\Gamma_{0,2}$ in chapter 5, will be written $c(\theta, g)$ here

$$c(\theta, g) \sim \int_1^\ell \frac{d\ell'}{\ell'} b(g(\ell')) \exp \left\{ 2 \int_1^{\ell'} \frac{d\ell''}{\ell''} \gamma_2(g(\ell'')) \right\} \quad (352)$$

where ℓ is an infrared length scale, measured in lattice units a for instance, and chosen in such a way that the homogeneous term in equation (5.248) be negligible. This means that, at the scale ℓ , the running temperature variable $\theta(\ell) \sim \theta_0$, where θ_0 is finite. The running temperature variable $\theta(\ell)$ is of the form $\theta(\ell) = \theta \ell$ to lowest order, since $\nu = 1$ for the pure system, hence $\ell \sim \theta_0/\theta$. As regards the other ingredients in equation (352), we need the solution of equation (350) corresponding to $n = 0$, which translated into the large length scale ℓ reads

$$g^2(\ell) = \frac{g^2}{1 + (g^2/\pi) \ln \ell} \quad (353)$$

and the anomalous dimension $\gamma_2(g)$ obtained from the rescaling of the ε operator as

$$\gamma_2(g) = \frac{d}{d \ln \Lambda} \ln \left(\Lambda^{-1+1/(1+g^2/2\pi)} \right) \sim -\frac{g^2}{2\pi} \quad (354)$$

Putting everything together, replacing $b(g)$ by $b(0)$

$$c(\theta, g) \sim \text{cst} \int_1^{\theta_0/\theta} \frac{d\ell'}{\ell'} \exp - \int_1^{\ell'} \frac{d\ell''}{\ell''} \frac{g^2}{\pi + g^2 \ln \ell''} \quad (355)$$

The integrals are straightforward and yield

$$c(\theta, g) \approx \frac{\text{cst}}{g^2} \ln \left[1 + \frac{g^2}{\pi} \ln \frac{\theta_0}{\theta} \right] \quad (356)$$

This is the surprising result found by Dotsenko and Dotsenko, showing a remaining weak singularity in $\ln \ln 1/\theta$ for exponentially small values of $\theta \leq \exp -\pi/g^2$. This subtle effect cannot be

obtained by a perturbative weak disorder expansion. Assuming that the underlying hypothesis are correct, a direct observation confirming (356) would be an other evidence of the power of the renormalization group.

The bosonic version does not give access to the average of the spin-spin correlation. Show however, following Shankar, that using similar renormalization group arguments, one can obtain the average of its square at criticality in the form

$$\overline{\langle \sigma(x_1) \sigma(x_2) \rangle^2} \sim \frac{[1 + g^2/\pi \ln |x_{12}|]^{\frac{1}{2}}}{|x_{12}|^{\frac{1}{2}}} \quad (357)$$

with logarithmic corrections typical of asymptotic freedom, correcting the $|x|^{-\frac{1}{2}}$ behaviour in the pure case. Since

$$[\overline{\langle \sigma(x_1) \sigma(x_2) \rangle}]^2 \leq \overline{\langle \sigma(x_1) \sigma(x_2) \rangle^2}$$

this implies the inequality

$$\overline{\langle \sigma(x_1) \sigma(x_2) \rangle} \leq \text{cst} \frac{[1 + g^2/\pi \ln |x_{12}|]^{\frac{1}{2}}}{|x_{12}|^{\frac{1}{2}}} \quad (358)$$

Equation (357) is reminiscent of the spin-spin correlation of the XY-model at the end point of the critical line (see equation (4.184)) with vortices playing the role of disorder. However the logarithmic enhancement occurring on the right-hand side of equation (358) might disappear in a more sophisticated analysis of the average correlation. Investigate in a similar vein the averages of squares of higher spin correlations.

Appendix 10.A The Hall conductance as a topological invariant

This appendix describes an elaboration of an argument due to Laughlin on the quantized Hall conductance for a minimally coupled interacting many body system, relating it to a topological invariant. We follow the presentation by Avron and Seiler. The idea is to combine gauge invariance and periodicity of wavefunctions when magnetic fluxes are increased by a quantum $\varphi_0 = h/e$, with a nontrivial topology including the wires and apparatus designed to measure (or to produce) the Hall voltage. The point of making such a connection with a topological invariant

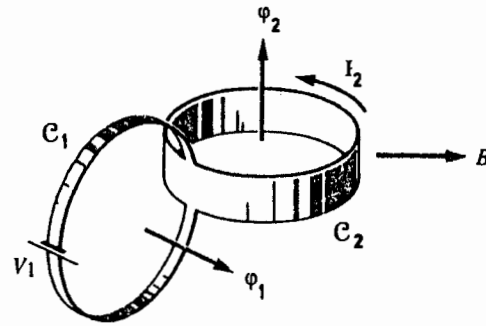


Fig. 6 The Hall setting including connecting wires and battery. The latter is replaced by a time dependent magnetic flux.

is that the latter will not be affected by smooth variations, impurities, etc.

One considers a many body Hamiltonian H_0 including the external strong (but fixed) magnetic field, as well as interparticle forces and interactions with the (possibly random) background. The discussion is carried at sufficiently low temperature that thermal effects are ignored and the crucial hypothesis is that in the absence of current and voltage, the system is in a pure nondegenerate state. One can then switch on a very weak Hall voltage (through a battery say) and observe a current in the main circuit as shown on figure 6.

The Hamiltonian H_0 includes a vector potential A_0 with a cut equal to the strong field B (and vanishing circulation around the loops C_1 and C_2 as indicated on the figure). To include the effect of the battery, as well as to have a conjugate variable for the current I_2 , one adds to A_0 an extra piece $\varphi_1 A_1 + \varphi_2 A_2$ such that

$$\oint_{C_a} dx \cdot A_b = \delta_{ab} \quad (A.1)$$

in such a way that the total magnetic flux is

$$\oint_{C_a} dx \cdot A = \varphi_a \quad (A.2a)$$

From Maxwell's equations, if φ_1 varies with time, the electromotive force induced in the loop is

$$\dot{\varphi}_1 = - \oint_{C_1} dx \cdot E = -V_1 \quad (A.2b)$$

The flux φ_1 is switched on adiabatically up to a linear regime corresponding to a very small V_1 . Crudely $\varphi_1 = -tV_1$, or in a more sophisticated manner for negative times $\varphi_1 = -V_1 e^{\eta t}/\eta$ in the limit where η tends to zero.

If A denotes the sum $A_0 + \varphi_1 A_1 + \varphi_2 A_2$, a typical Hamiltonian will have the structure

$$H = \sum_p \frac{1}{2m} \left(\frac{\hbar}{i} \nabla_p - eA(x_p) \right)^2 + V(\{x\}) \quad (A.3)$$

If we delete two cuts (to make the domain simply connected), A_1 and A_2 can be considered in a convenient gauge as gradients of two functions Λ_1 and Λ_2 . As one winds around C_a , Λ_a increases by unity. One can therefore in this cut region perform a gauge transformation to reduce the vector potential to A_0 and the Hamiltonian to H_0 , at the price of writing wavefunctions as $\psi = \exp[ie/\hbar \sum (\varphi_1 \Lambda_1 + \varphi_2 \Lambda_2)] \psi_0$. We now have to account for a discontinuity in ψ_0 , its value on one side of each cut being related to its value on the other side by a factor $\exp[ie\varphi_a/\hbar]$ for each coordinate variable x_p . This shows the periodicity of the problem as φ_1 or φ_2 is increased by a quantum $\varphi_0 = h/e$, with H_0 (including the boundary conditions) defining a doubly periodic quantum mechanical system in the parameter plane φ_1, φ_2 . Before applying this gauge transformation, the current density operator is a sum

$$J = e \sum_p v_p = \sum_p j_p \quad v_p = \frac{1}{m} \left(\frac{\hbar}{i} \nabla_p - eA(x_p) \right)$$

and

$$\frac{\partial H}{\partial \varphi_a} = -\frac{1}{2} \sum_p \{j_p, A_a(x_p)\}_+$$

The hypothesis, made in the framework of the one body theory, that the Fermi level falls in a gap of the spectrum, is replaced here by the assumption that the system is in a normalized

nondegenerate state $|\psi\rangle$. Therefore

$$\langle\psi|\frac{\partial H}{\partial\varphi_a}|\psi\rangle = -\int d^2x_1\cdots d^2x_N \sum_p \mathbf{A}_a(\mathbf{x}_p)\psi^*\mathbf{j}_p\psi$$

Expressing $\mathbf{A}_a$ as the gradient of Λ_a in the cut region, and using current conservation, we get only a boundary contribution arising from the unit discontinuity of Λ_a , of the form

$$I_a = -\langle\psi|\frac{\partial H}{\partial\varphi_a}|\psi\rangle \quad (\text{A.4})$$

This is a first justification for introducing the flux φ_2 in order to compute the current I_2 as a response function. To simplify the matter at the expense of rigour, we use $\varphi_1 = -tV_1$ to translate the time dependent Schrödinger equation satisfied by ψ into

$$-i\hbar V_1 \frac{\partial}{\partial\varphi_1} |\psi\rangle = H(\varphi_1, \varphi_2) |\psi\rangle \quad (\text{A.5})$$

Thus differentiating this equation with respect to φ_2 and taking the scalar product with the adjoint bra $\langle\psi|$ and using finally the expression (A.4) for the current I_2 , we obtain

$$-i\hbar V_1 \langle\psi|\frac{\partial^2}{\partial\varphi_1\partial\varphi_2}|\psi\rangle = -I_2 + i\hbar V_1 \left\langle\frac{\partial\psi}{\partial\varphi_1}\left|\frac{\partial\psi}{\partial\varphi_2}\right.\right\rangle$$

The Hall conductance defined as

$$I_2 = \sigma_H V_1 \quad (\text{A.6})$$

is then given by a substitute for Kubo's formula, which reads

$$\sigma_H = i\hbar \left\{ \frac{\partial}{\partial\varphi_2} \left\langle\psi\left|\frac{\partial\psi}{\partial\varphi_1}\right.\right\rangle + \left\langle\frac{\partial\psi}{\partial\varphi_1}\left|\frac{\partial\psi}{\partial\varphi_2}\right.\right\rangle - \left\langle\frac{\partial\psi}{\partial\varphi_2}\left|\frac{\partial\psi}{\partial\varphi_1}\right.\right\rangle \right\} \quad (\text{A.7})$$

Instead of computing the derivatives and then letting V_1 (hence also φ_1) go to zero as well as φ_2 , the trick is to replace the above expression by an average over each flux in an interval $[0, \varphi_0]$, the rationale being that φ_0 is very small as compared to macroscopic fluxes. This enables one to use the adiabatic approximation to leading order (for vanishingly small V_1). The latter asserts that, if for some initial value the state ψ is along an eigenstate of the φ -dependent Hamiltonian H , it remains so up to a phase. It is natural to take this state as being the normalized ground state

$$H(\varphi_1, \varphi_2) |\Omega(\varphi_1, \varphi_2)\rangle = E(\varphi_1, \varphi_2) |\Omega(\varphi_1, \varphi_2)\rangle \quad (\text{A.8})$$

where the energy eigenvalue is doubly periodic with period φ_0 . Thus, in equation (A.7), we approximate the first term on the right-hand side by $\partial/\partial\varphi_2 \langle\psi|H|\psi\rangle \approx \partial/\partial\varphi_2 E(\varphi_1, \varphi_2)$ with zero average, and identify the Hall conductance with

$$\langle\sigma_H\rangle = \frac{i\hbar}{\varphi_0^2} \int_0^{\varphi_0} \int_0^{\varphi_0} d\varphi_1 d\varphi_2 \left\{ \left\langle\frac{\partial\psi}{\partial\varphi_1}\left|\frac{\partial\psi}{\partial\varphi_2}\right.\right\rangle - \left\langle\frac{\partial\psi}{\partial\varphi_2}\left|\frac{\partial\psi}{\partial\varphi_1}\right.\right\rangle \right\} \quad (\text{A.9})$$

Equivalently with a compact notation for the two-form integrated over the torus $\mathcal{T}$ in flux space

$$\langle\sigma_H\rangle = \frac{e^2}{h} \frac{i}{2\pi} \int \int_{\mathcal{T}} \langle d\psi | d\psi \rangle = \frac{e^2}{h} \frac{1}{2i\pi} \int \int_{\mathcal{T}} \text{Tr}[dP P dP] \quad (\text{A.10})$$

Here the projector $P = |\psi\rangle\langle\psi|$ is approximated to leading order by the ground state projector $|\Omega\rangle\langle\Omega|$, both being insensitive to a flux dependent (but coordinate independent) phase of the state.

The physics is doubly periodic in flux space, but this is neither the case for the Hamiltonian, nor for the state Ω . Under a shift through φ_0 , H is changed into a unitary equivalent operator. To remedy this, we perform a unitary gauge transformation on the states corresponding to use $\mathbf{A}_a = \nabla \Lambda_a$ provided cuts are introduced in physical space. Thus with

$$\begin{aligned} F(\{\mathbf{x}\}) &= \sum_p \frac{e}{h} [\varphi_1 \Lambda_1(\mathbf{x}_p) + \varphi_2 \Lambda_2(\mathbf{x}_p)] \\ U &= e^{iF} \\ |\Omega\rangle &= U |\Omega_0\rangle \end{aligned} \quad (\text{A.11})$$

the Hamiltonian H turns into H_0 and boundary conditions on $|\Omega_0\rangle$ are periodic in φ -space. The situation is analogous to the one in solid state physics for particles in a periodic potential and Bloch waves, with φ playing the role of quasimomentum. We now have $P = U P_0 U^\dagger$. This transformation affects the integrand in equation (A.10) by a total derivative with a vanishing contribution upon integration. As P_0 is now periodic in φ , this shows both that $\langle\sigma_H\rangle$ is sensible (being now invariant under shifts in φ -space), but also, under the assumption of no degeneracy, that it is e^2/h times an integer. Henceforth the subscript zero will be omitted, it being understood that $P(\varphi)$ is a doubly periodic smooth projector on a

one-dimensional subspace. The theorem is then that the integral

$$K = \frac{1}{2i\pi} \int \int_T \text{Tr } dP P dP \quad (A.12)$$

is both topologically invariant and an integer. It is a classical result due to von Neumann and Wigner that generically level crossing is of codimension three, meaning that it requires at least a three-dimensional parameter space. When the strong external field is kept fixed, the two-dimensional φ_1, φ_2 space is insufficient to allow degeneracies in general. To explain the fractional Hall effect, one would however require degenerate ground states for a generalization of the above argument. We limit ourselves here to the integral effect and to the hypothesis of no degeneracy. The invariant K is reminiscent of similar expressions in gauge theories and can be understood as an integral of curvature of an appropriate bundle.

We conclude this appendix with a proof that K is indeed an integer. For that purpose, consider an arbitrary parametrized loop in flux space and introduce along this loop the evolution equation

$$\frac{d}{ds} |X(s)\rangle = \left(\frac{dP}{ds} P - P \frac{dP}{ds} \right) |X(s)\rangle \quad (A.13)$$

with $P(s) \equiv P(\varphi_1(s), \varphi_2(s))$, $\varphi_a(s+1) = \varphi_a(s)$. The r.h.s. operator is antihermitian so that the evolution along the loop is unitary. It is readily found that if $P(0)|X(0)\rangle = |X(0)\rangle$, then $|X(s)\rangle$ remains an eigenvector of P with eigenvalue one. The same holds true in the complementary orthogonal space. Under this assumption, after a complete turn, $|X(1)\rangle = e^{i\kappa(\gamma)} |X(0)\rangle$. The phase $e^{i\kappa(\gamma)}$ is of a type considered by Berry in his study of adiabatic transformations. To compute its value, recall that on each point of the loop $P = |\Omega\rangle\langle\Omega|$, with Ω the ground state of H , with discontinuities across the cuts. Thus

$$dP P - P dP = |d\Omega\rangle\langle\Omega| - |\Omega\rangle\langle d\Omega| + \{ \langle d\Omega|\Omega\rangle - \langle\Omega|d\Omega\rangle \} |\Omega\rangle\langle\Omega|$$

Inserting this expression in equation (A.13), we see that the projection $\langle\Omega|X\rangle$ satisfies

$$\frac{d}{ds} \langle\Omega|X\rangle = \left\langle \frac{d\Omega}{ds} | \Omega \right\rangle \langle\Omega|X\rangle \quad (A.14)$$

with $|X(0)\rangle$ and $|X(1)\rangle$ both proportional to $|\Omega(\varphi_1(0), \varphi_2(0))\rangle$. Thus the phase is given by

$$e^{i\kappa(\gamma)} = e^{\int_\gamma \langle d\Omega|\Omega\rangle} \quad (A.15)$$

On the φ -torus T , choose γ homologous to zero. It bounds two distinct domains S_1 and S_2 , and the loop integral can be written in two ways as a surface integral which can be expressed in terms of the projector P due to the nondegeneracy hypothesis

$$\exp \int_\gamma \langle d\Omega|\Omega\rangle = \exp \iint_{S_1} \text{Tr } dP P dP = \exp \iint_{S_2} \text{Tr } dP P dP \quad (A.16)$$

If γ shrinks to a point as well as S_1 , the first two exponentials become unity, forcing the remaining integral over the whole torus to become a multiple of $2i\pi$, i.e. K defined by equation (A.12) to be an integer, as was claimed.

One might argue that the definition taken for the Hall conductance as an average is not perhaps necessary. This is compensated for by the great generality of the argument based solely on the geometry of the setting and the consequences of gauge invariance in quantum mechanics. We finally note that Berry's phases such as the one introduced above occur in many significant circumstances when investigating the topology of the parameter space of quantum systems (or systems which admit a superposition principle).

Notes

The localization effect of disorder on wavefunctions is the subject of the classic paper by P.W. Anderson *Phys. Rev.* **109**,1492 (1958). One finds a number of references to the many works which have followed in the reviews by D. Thouless in *Ill Condensed Matter*, Les Houches, R. Balian, R. Maynard and G. Toulouse eds., North Holland (1979) and the articles by D. Thouless, E. Abrahams and F. Wegner in *Common Trends in Particle and Condensed Matter Physics*. E. Brézin, J.-L. Gervais, and G. Toulouse eds. *Phys. Reports* **67** (1980).

The treatment of a one-dimensional Hamiltonian with a Gaussian distributed potential is due to B. Halperin, *Phys. Rev.* **A139**, 104 (1965). This is closely connected with one-dimensional diffusion problems reviewed in S. Alexander, J. Bernasconi,

W.R. Schneider and R. Orbach, *Rev. Mod. Phys.* **53**, 175 (1981). For another interesting solvable case see J.P. Bouchaud, A. Comtet, A. Georges, P. Le Doussal, *Europhys. Lett.* **3**, 653 (1987).

The replica trick was introduced in the study of random systems by S.F. Edwards and P.W. Anderson, *J. Phys.* **F5**, 965 (1975). The supersymmetry arising when representing Jacobians by Grassmann integrals was first discussed in the context of gauge theories by C. Becchi, A. Rouet, and R. Stora in *Renormalization Theory*, G. Velo and A.S. Wightman eds, Reidel, Dordrecht (1976).

The spectrum of random chains of oscillators was investigated by F.J. Dyson, *Phys. Rev.* **92**, 1331 (1953) and D.J. Thouless, *J. Phys.* **C5**, 77 (1972). The small disorder expansion is due to B. Derrida and R. Orbach, *Phys. Rev.* **B27**, 4694 (1983). The example discussed in the text is from a joint paper with E.J. Gardner and B. Derrida, *J. Phys.* **A17**, 1093 (1984).

For a review on the quantum Hall effect, see K. von Klitzing, *Rev. Mod. Phys.* **58**, 519 (1986) and *The Quantum Hall Effect*, R.E. Prange and S.M. Girvin eds, Springer Verlag, New York (1987). The computation of the one-particle spectrum in a strong magnetic field and a random potential is due to F. Wegner, *Z. Phys.* **B51**, 279 (1983). The generalization discussed in the text was worked out with E. Brézin and D. Gross, *Nucl. Phys.* **B235** [FS 11], 24 (1984). The insensitivity of the integral quantization of the Hall conductance to the presence of impurities, as a consequence of gauge invariance, is discussed in R.B. Laughlin, *Phys. Rev.* **B23**, 5632 (1981). We present in appendix A an elaboration of this argument due to J.E. Avron and R. Seiler, *Phys. Rev. Lett.* **54**, 259 (1985) following earlier work by several authors quoted in this paper. Phases occurring in adiabatic transitions are studied in M.V. Berry, *Proc. Roy. Soc. London* **A392**, 45 (1984).

Ensembles of random matrices introduced and studied by E.P. Wigner, F.J. Dyson, M.L. Mehta, M. Gaudin and others are presented by M.L. Mehta in *Random Matrices and the Statistical Theory of Energy Levels*, Academic Press, New York (1967). For the interpretation in terms of anticommuting variables, see the contribution by E. Brézin in the proceedings of the VIIIth Sitges conference, L. Garrido ed., *Lecture Notes in Physics* **216**, 115, Springer, Berlin (1985).

The planar approximation is due to G. 't Hooft, *Nucl. Phys.* **B72**, 461 (1974) *ibid*, **B75**, 461 (1974). Algebraic aspects were investigated by J. Koplik, A. Neveu and S. Nussinov, *Nucl. Phys.* **B123**, 109 (1977). Our presentation follows work done in collaboration with E. Brézin, G. Parisi and J.-B. Zuber, *Comm. Math. Phys.* **59**, 35 (1978) and D. Bessis and J.-B. Zuber, *Adv. in Appl. Math.* **1**, 109 (1980).

The behaviour of a spin system in a random external field was investigated by Y. Imry and S.K. Ma, *Phys. Rev. Lett.* **35**, 1399 (1975). Dimensional transmutation using supersymmetry was demonstrated by G. Parisi and N. Sourlas, *Phys. Rev. Lett.* **43**, 744 (1979). A general reference on the applications of anticommuting variables in the study of random systems is K.B. Efetov, *Adv. in Phys.* **32**, 53 (1983).

A criterion for assessing the effect of weak impurities on continuous phase transitions was put forward by A.B. Harris, *J. Phys.* **C7**, 1671 (1974). The application to the two-dimensional Ising model with random bonds was investigated by Vi.S. Dotsenko and Vi.S. Dotsenko in *Adv. in Phys.* **32**, 129 (1983), and R. Shankar, *Phys. Rev. Lett.* **58**, 2466 (1987). The underlying field theoretic model is due to D.J. Gross and A. Neveu, *Phys. Rev.* **D10**, 3235 (1974). For the two-loop renormalization of this model, see W. Wetzel, *Phys. Lett.* **153B**, 297 (1985).

RANDOM GEOMETRY

Statistical models with a geometrical basis arise in many circumstances, such as the theory of liquids, membranes, polymer networks, defects, microemulsions, interfaces, etc. Gauge theories also lead to random surfaces, as does the theory of extended objects such as strings, and quantum gravity requires a generalization to four-manifolds. From a general point of view, local quantum field theory is rooted in the study of random paths. One may wish to find such a universal model, generalizing Brownian curves, to Brownian manifolds and in the first instance to surfaces (Polyakov, 1981). Despite many efforts, no such universal archetype has been found, although the endeavour towards such a model has uncovered a rich mathematical structure. By necessity, our presentation will be limited to the most elementary aspects.

In the first section we discuss random lattices in Euclidean space. The use of such lattices was advocated as a mean to restore translational invariance while keeping the advantage of a short distance cutoff (Christ, Friedberg and Lee, 1982). The formalism could be generalized to other manifolds, but we refrain from doing so, nor will we pursue the analysis of standard models on random lattices, a difficult subject. Even free field theory on such a lattice opens the Pandora's box of disordered systems. Concepts from random lattices may be useful in the study of liquids or gases.

In the second section we look at random surfaces, both in a discrete and continuous setting. Thus emerges the role of reparametrization invariance, and its associated quantum anomaly related to conformal field theory. We present in some detail the Polyakov model for smooth surfaces, with its magic dimension 26, as well as a discrete analog, pointing out the connection of the latter with planar quantum field theory.

11.1 Random lattices

Random geometry belongs to an old mathematical tradition. Recall for instance the problem of Buffon's needle. It is described in a classical book by Santalo under the name of integral geometry. In the early sixties, Bernal investigated random packing of spheres and applied them to the study of liquid structure functions. Here we describe a related topic, namely Poissonian random lattices and their local properties. They provide a means of introducing a short-distance cutoff in field theory, while hopefully maintaining continuous geometrical symmetries, assuming self-averaging of bulk quantities. This offers the opportunity to describe in elementary terms some concepts both of topological and geometrical interest. The study of models on random lattices paves the way to an understanding of field theory in curved space, since the use of a random lattice is a discrete analog of the choice of an arbitrary coordinate system in continuous geometry.

11.1.1 Poissonian lattices and cell statistics

In a large regular region of volume Ω , in Euclidean d -dimensional space, N points are chosen independently with uniform probability. We consider a limit with N and Ω tending to infinity keeping the density ρ fixed

$$N \rightarrow \infty, \quad \Omega \rightarrow \infty, \quad N/\Omega \rightarrow \rho = a^{-d} \quad (1)$$

The quantity $a = \rho^{-1/d}$ plays the role of an elementary length scale. We choose units in such a way that $a = 1$, hence $\rho = 1$. As we saw in chapter 10, in one dimension the points can be ordered and the intervals ℓ between successive points become independent variables in the thermodynamic limit, with a distribution

$$p(\ell)d\ell = e^{-\ell}d\ell \quad (2)$$

This justifies calling these lattices Poissonian.

Consider the case of two dimensions. The generalization to higher dimension will be straightforward. The following construction is due to Dirichlet and Voronoi. To each point M_i belonging to the lattice, we associate the closed cell C_i of those points of the plane which are nearer to M_i than to any M_j , $j \neq i$.

The cell C_i is a closed convex polygon, as an intersection of half

planes. For each pair i, j the intersection $C_i \cap C_j$ is either empty or a segment lying on the bisecting line between M_i and M_j and belonging to the boundary of C_i and C_j . In the latter case we say that (M_i, M_j) form a pair of neighbours. We call q_i be the local coordination number, the number of neighbours of M_i , or equivalently the number of sides of the cell C_i . For a triple $i \neq j \neq k$, if $C_i \cap C_j \cap C_k$ is not empty, it is a vertex of each of the corresponding cells, the center of the circle through M_i, M_j, M_k and we call (M_i, M_j, M_k) an elementary 2-simplex, as (M_i, M_j) was an elementary 1-simplex. The planar region is now paved by 2-simplices as well as by cells C_i . The two sets are dual to each other.

Let $N_0 = N$ be the number of points M_i , equal to the number of cells, N_1 the number of links between neighbours and N_2 the number of 2-simplices, equal to the number of cell vertices. We have the relation

$$N_0 - N_1 + N_2 = \chi \quad (3)$$

where the Euler characteristic χ (2 for the sphere, 0 for a torus, etc.) is not extensive, i.e. in the infinite volume limit $\chi/N \rightarrow 0$. Cutting the plane along the edges of the cells we get N_0 pieces (the cells) with $2N_1 = \sum_i q_i$ edges and $3N_2$ vertices. Since each cell has as many vertices as edges (Euler's relation in dimension one) we find

$$\begin{aligned} 2N_1 &= 3N_2 = \sum_i q_i \\ N_0 &= \sum_i 1 \end{aligned} \quad (4)$$

Consequently, if the average coordination number is defined as

$$n_1 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i q_i \quad (5)$$

the above relations imply

$$n_1 = 6 \quad \lim_{N \rightarrow \infty} \frac{N_1}{N_0} = 3 \quad \lim_{N \rightarrow \infty} \frac{N_2}{N_0} = 2 \quad (6)$$

Equivalently, if n_a is the average number of a -simplices incident on a 0-simplex (a point) (with $n_1 \equiv q$)

$$n_1 = n_2 = 6 \quad (7)$$

The two-dimensional case is remarkable in that on the triangular regular lattice, the relations (6) and (7) are true locally. The random lattice may therefore be loosely considered as arising from defects in a regular triangular lattice, the defects being the points where the local coordination number are different from six. These points are the analogs of charges (the deviation from six), with equation (6) implying global neutrality.

Let us now introduce metric properties. We use the notation

$$\begin{aligned} \ell_{ij} &= \text{distance between the neighbours } (M_i, M_j) & \langle \ell_{ij} \rangle &= \ell_1 \\ \ell_{ijk} &= \text{area of simplex } (M_i, M_j, M_k) & \langle \ell_{ijk} \rangle &= \ell_2 \end{aligned}$$

Similarly

$$\begin{aligned} \sigma_i &= \text{area of cell } C_i & \langle \sigma_i \rangle &= \sigma_2 \\ \sigma_{ij} &= \text{length of edge perpendicular to the link } (ij) & \langle \sigma_{ij} \rangle &= \sigma_1 \end{aligned}$$

Since by convention the density is unity, we have

$$\sigma_2 = 1 \quad \ell_2 = \frac{1}{2} \quad (8)$$

The computation of ℓ_1 and σ_1 is more tricky and the reasoning is typical of random lattice problems. Let, for short, (i, j, k) be a 2-simplex with 0 as the dual vertex common to the cells C_i, C_j and C_k . The circle Γ of radius R centered in 0 is circumscribed to the triangle (i, j, k) (figure 1), and there can be no lattice point inside Γ , since for $\ell \neq i, j, k$ the distance $|OM_\ell|$ is greater or equal to $|OM_i|, |OM_j|$, and $|OM_k|$.

Given a point M_1 belonging to the lattice, the probability that two other points M_2 and M_3 , lying within d^2x_2 and d^2x_3 respectively, form an elementary 2-simplex is therefore

$$dp = \frac{1}{n_2} \times \frac{1}{2}(N-1)(N-2) \times \left(1 - \frac{\pi R^2}{\Omega}\right)^{N-3} \frac{d^2x_2 d^2x_3}{\Omega \Omega}$$

where the first factor n_2 is a normalization of the average number of 2-simplices per 0-simplex, the second expresses the numbers of choices of M_2 and M_3 among the $N-1$ remaining points, and the third factor is the integrated probability of the $(N-3)$ remaining lattice points lying outside the circle Γ . In the limit $N \rightarrow \infty$, the required probability

$$dp = \frac{1}{12} e^{-\pi R^2} d^2x_2 d^2x_3 \quad (9)$$

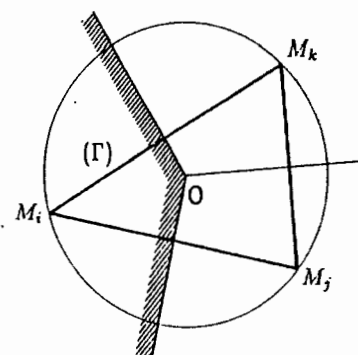


Fig. 1 The Dirichlet-Voronoi construction in two dimensions.

has a form reminiscent of the Poisson distribution (2).

Using as parameters the radius R of the circumscribed circle and the polar angles $\varphi_1, \varphi_2, \varphi_3$ of OM_1, OM_2, OM_3 , with

$$d^2x_2 d^2x_3 = R^3 |\sin(\varphi_1 - \varphi_2) + \sin(\varphi_2 - \varphi_3) + \sin(\varphi_3 - \varphi_1)| dR d\varphi_1 d\varphi_2 d\varphi_3$$

check that the probability (9) is correctly normalized

$$\begin{aligned} \int dp &= \frac{1}{3}\pi \int_0^\infty dR R^3 e^{-\pi R^2} \int_0^{2\pi} d\varphi_3 \\ &\quad \int_0^{\varphi_3} d\varphi_2 \{ \sin(\varphi_3 - \varphi_2) - \sin \varphi_3 + \sin \varphi_2 \} \\ &= 2\pi^2 \int_0^\infty dR R^3 e^{-\pi R^2} = 1 \end{aligned}$$

To obtain ℓ_1 , the average separation between neighbours, we average $|M_1 M_2| = 2R |\sin \frac{1}{2}(\varphi_2 - \varphi_1)|$ over dp , i.e.,

$$\begin{aligned} \ell_1 &= \frac{2}{3}\pi \int_0^\infty dR R^4 e^{-\pi R^2} \\ &\quad \times \int_0^{2\pi} d\varphi_3 \int_0^{\varphi_3} d\varphi_2 \sin \frac{1}{2}\varphi_2 \{ \sin(\varphi_3 - \varphi_2) - \sin \varphi_3 + \sin \varphi_2 \} \end{aligned}$$

Thus

$$\ell_1 = \frac{32}{9\pi} = 1.1317684 \quad (10)$$

- i) In two dimensions the procedure followed above, to compute the average side of triangle and identifying it with the average link side is justified by the fact that each link belongs to a fixed number of triangles, namely two. Indeed one has

$$\begin{aligned} \ell_1 &= \frac{1}{N_1} \sum_{\text{links}} \ell_{ij} = \frac{1}{2N_1} \sum_{\text{triangles } T} \sum_{\ell_{ij} \in T} \ell_{ij} \\ &= \frac{3N_2}{2N_1} \times \text{average of a triangle side} \end{aligned}$$

and $3N_2/2N \rightarrow 1$. In three dimensions, the number of tetrahedra sharing a link is not a constant throughout the lattice, so the average length of a tetrahedron side is not the average of a link of the lattice. But the procedure is correct for the average area of a triangular face. For lack of a better procedure the quantities referred to as $\ell_{(i)}$ in dimension d below are in fact computed as averages of the corresponding elements for a typical simplex.

- ii) The quantity ℓ_1 for a random lattice may be compared to the equivalent on a regular triangular lattice with the same density of points i.e. $2^{\frac{1}{2}}3^{-\frac{1}{4}} = 1.0745699$ showing that the random lattice is somehow "looser".
- iii) Using the probability density (9), check that ℓ_2 is $\frac{1}{2}$ as it should, and that its relative variance is quite large

$$\frac{\delta \ell_2}{\ell_2} = \left(\frac{\langle \ell_{ijk}^2 \rangle}{\langle \ell_{ijk} \rangle^2} - 1 \right)^{\frac{1}{2}} = \left(\frac{35}{2\pi^2} - 1 \right)^{\frac{1}{2}} = 0.8793$$

The average perimeter of cells is obtained using a clever argument due to Meijering. Let M be a lattice point and consider for any other lattice point M_k the bisecting line Δ_k of the segment MM_k . The number of such lines at distance between r and $r + dr$ is the number of points M_k at distance between $2r$ and $2r + 2dr$, namely $8\pi r dr$. The fraction of length of such a line at a distance between R and $R + dR$ is $(R^2 - r^2)^{-\frac{1}{2}} 2R dR$ (Figure 2). The total amount of would-be cell boundary length between R and $R + dR$ is obtained by integrating the product of these two factors over r between the values 0 and R ,

$$16\pi R dR \int_0^R \frac{r dr}{\sqrt{R^2 - r^2}} = 16\pi R^2 dR$$

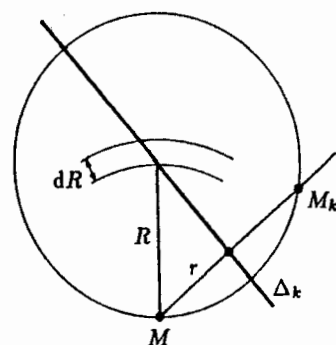


Fig. 2 Geometric construction involved in Meijering's argument.

An element of this length will actually belong to the boundary of the cell around M if the circle of radius R centered at the length element and (therefore going through M) does not contain any interior lattice point. This occurs with probability $e^{-\pi R^2}$. Thus the amount of perimeter of the cell boundary at a distance between R and $R + dR$ is

$$16\pi R^2 e^{-\pi R^2} dR$$

Integrating over R yields the mean perimeter

$$P = 16\pi \int_0^\infty R^2 e^{-\pi R^2} dR = 4 \quad (11)$$

Since on average each cell has six sides, we find

$$\sigma_1 = \frac{2}{3} \quad (12)$$

to be compared with the value $2^{1/3}3^{-1/3} = 0.62204032$ on a regular triangular lattice. Table I summarizes these two-dimensional local results.

It is possible to write an analytical (but untractable) expression for the probability p_n of finding a cell with n sides. Let M_0 at the origin of the coordinate system stand for a lattice point, and let $M_1, \dots, M_n$ with coordinates $x_1, \dots, x_n$, be its neighbours. We denote by A the area of the union of circles centered at $M_1, \dots, M_n$, and going through M_0 , and by $\chi(M_1, \dots, M_n)$ a characteristic function with value one if the bisecting lines of the segments (M_0, M_i)

Table I. Local averages in two dimensions corresponding to unit density.

Dimension of simplex	Number	Average number incident on site	Mean size	Mean size (dual)
0	$N_0 = N$	$n_0 = 1$		$\sigma_2 = 1$
1	$N_1 = 3N$	$n_1 = 6$	$\ell_1 = \frac{32}{9\pi}$	$\sigma_1 = \frac{2}{3}$
2	$N_2 = 2N$	$n_2 = 6$	$\ell_2 = \frac{1}{2}$	

Table II. Probability distribution for n -sided cells in a Poissonian two-dimensional lattice.

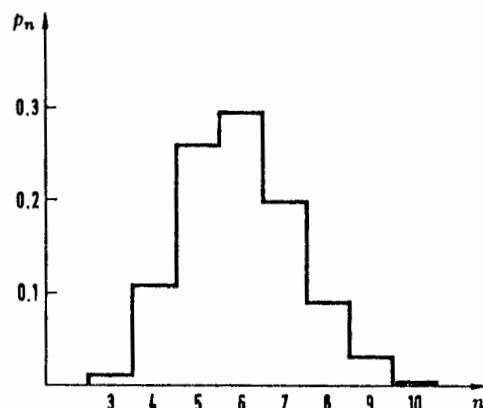
n	p_n
3	$(1.127 \pm 0.008)10^{-2}$
4	$(1.077 \pm 0.0011)10^{-1}$
5	0.258 ± 0.002
6	0.294 ± 0.003
7	0.198 ± 0.003
8	0.090 ± 0.020
9	$(2.88 \pm 0.07)10^{-2}$
10	$(6.95 \pm 0.20)10^{-3}$
20	$(1.5 \pm 0.8)10^{-13}$
35	$3.6 \cdot 10^{-40}$
50	$1.5 \cdot 10^{-73}$

generate an n -sided convex polygon, and zero otherwise. Then

$$p_n = \frac{1}{n!} \int \prod_{i=1}^n d^2 x_i \chi(x_1, \dots, x_n) e^{-A} \quad (13)$$

Integrating (numerically) this expression leads to the values quoted in table II. Although the average value $q = \sum_3^\infty n p_n = 6$ is also the most probable, occurring approximately in 30 percent of the cases, other small coordination numbers have an appreciable probability. For large n , p_n decreases very fast as $n^{-\alpha n}$, where α is between 1 and 2. The corresponding histogram is shown on figure 3.

Similarly one can study the mean distance λ_n to the perimeter of the cell or A_n the average area of an n -sided cell. One observes

Fig. 3 Probability distribution of n -sided cells.

that A_n is almost linear in n with $A_n \sim \frac{1}{4}n + \text{cst}$ for large n , as if the cell was growing in an "unfriendly" environment. The ratio A_n/λ_n^2 seems to converge to a value close to π , suggesting that a very large cell behaves in certain respects as a circle.

The Dirichlet-Voronoi construction carries over in higher dimensions. Let us present some details in three dimensions. Let $N_0 = N$ be the number of points, N_1 of links, N_2 of triangles and N_3 of tetrahedra. The Euler characteristic vanishes in odd dimensions, thus

$$N_0 - N_1 + N_2 - N_3 = 0 \quad (14)$$

Each triangle is shared by two tetrahedra, and there are four triangles on the boundary of a tetrahedron, thus

$$2N_2 = 4N_3 \quad (15)$$

If as above n_a stands for the number of a -simplices incident on a point

$$n_a = (a+1)N_a/N \quad (16)$$

topology alone yields the relations

$$\begin{aligned} n_1 &= 2 + \frac{1}{2}n_3 \\ n_2 &= \frac{3}{2}n_3 \end{aligned} \quad (17)$$

leaving one unknown, for instance n_3 . The probability of finding a tetrahedron with one vertex M_1 at the origin and the three other ones M_2, M_3, M_4 at $\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ within infinitesimal volumes d^3x_i is

$$dp = \frac{1}{3!n_3} e^{-\frac{1}{2}\pi R^3} \prod_{i=2}^4 d^3x_i \quad (18)$$

In this relation R is the radius of the sphere circumscribed to the tetrahedron. Thus

$$n_3 = \frac{1}{3!} \int e^{-\frac{1}{2}\pi R^3} \prod_{i=2}^4 d^3x_i \quad (19)$$

To evaluate this integral one can use the following coordinates. Let O be the center of the circumscribed sphere and $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ unit vectors along OM_1, OM_2, OM_3 and OM_4 . Call R^3w the volume of the tetrahedron with

$$w = \pm \frac{1}{3!} \det \begin{vmatrix} 1 & 1 & 1 & 1 \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \mathbf{u}_4 \end{vmatrix} \quad (20)$$

with a sign chosen in such a way that w be positive. Then

$$\frac{1}{3!} \prod_{i=2}^4 d^3x_i = R^3 w dR d^2\hat{u}_1 d^2\hat{u}_2 d^2\hat{u}_3 d^2\hat{u}_4 \quad (21)$$

with $d^2\hat{u}$ the measure on the sphere. Then

$$n_3 = 72\pi \langle w \rangle \quad (22)$$

where w stands for the rotational average

$$\langle w \rangle = \int w \prod_{i=1}^4 \frac{d^2\hat{u}_i}{4\pi} \quad (23)$$

We shall obtain below (equation (36)) the d -dimensional generalization of this integral. Restricted to $d=3$ it yields

$$n_3 = \frac{96}{35}\pi^2 = 27.0709 \dots$$

$$n_1 = 2 + \frac{48}{35}\pi^2 = 15.53546 \dots \quad n_2 = \frac{144}{35}\pi^2 = 40.6064 \quad (24)$$

where use has been made of equation (17). Alternatively this means that

$$\begin{aligned} N_3/N &= \frac{24}{35}\pi^2 & N_2/N &= \frac{48}{35}\pi^2 & N_1/N &= 1 + \frac{24}{35}\pi^2 \end{aligned} \quad (25)$$

Of course these numbers do not correspond to any regular lattice.

The average volume of a tetrahedron in natural units is the reciprocal of N_3/N , i.e.

$$\ell_3 = 35/24\pi^2 = 0.1478 \dots \quad (26)$$

From the probability density (18), one can also compute the average area of a triangle

$$\ell_2 = \left(\frac{3}{4\pi}\right)^{\frac{1}{2}} \frac{875}{243} \pi^{-1} \Gamma\left(\frac{2}{3}\right) = 0.5973 \dots \quad (27)$$

and the mean length of a 1-simplex, i.e. the average distance between neighbours, with the caveat explained above

$$\ell_1 = \left(\frac{3}{4\pi}\right)^{\frac{1}{2}} \frac{1715}{2304} \Gamma\left(\frac{1}{3}\right) = 1.2371 \quad (28)$$

For the dual lattice $\sigma_3 = 1$. The average area of cell boundary is obtained using Meijering's reasoning as

$$\text{average area of cell boundary} = \left(\frac{4\pi}{3}\right)^{\frac{1}{2}} \frac{8}{3} \Gamma\left(\frac{2}{3}\right) = 5.821 \quad (29)$$

Thus σ_2 follows by dividing out by the connectivity n_1

$$\sigma_2 = \frac{\left(\frac{4\pi}{3}\right)^{\frac{1}{2}} \frac{4}{3} \Gamma\left(\frac{2}{3}\right)}{1 + \frac{24}{35} \pi^2} = 0.3747 \dots \quad (30)$$

These data are collected in table III while table IV present similar data pertaining to four dimensions.

The average coordination number n_1 grows very fast with dimension being already $\frac{340}{9} = 37.778$ in four dimensions as opposed to 8 on a regular hypercubical lattice.

The above computations are rather arduous, and fewer results are known in general dimension d . Let us quote some of them.

- i) Since each $(d-1)$ -simplex belongs to two d -simplices and since $d+1$ $(d-1)$ -simplices are incident on a d -simplex

$$2N_{d-1} = (d+1)N_d \quad n_{d-1} = \frac{1}{2}dn_d \quad (31)$$

Euler's relation applied to the boundary of a cell, homeomorphic to a sphere yields

$$n_1 - n_2 + \dots + (-1)^{d-1}n_d = 1 - (-1)^d \quad (32)$$

Table III. Cell data in three dimensions

Dimension of simplex	Number	Number incident on a point	Average size	Average size (dual)
0	$N_0 = N$	$n_0 = 1$		
1	$N_1 = \frac{1+24\pi^2}{35}N$	$n_1 = 2 + \frac{48}{35}\pi^2$	$\ell_1 = \left(\frac{2}{4\pi}\right)^{\frac{1}{2}} \frac{1715}{2304} \Gamma\left(\frac{1}{3}\right)$	$\sigma_2 = \frac{\left(\frac{4\pi}{3}\right)^{\frac{1}{2}} \frac{4}{3} \Gamma\left(\frac{2}{3}\right)}{1 + \frac{24}{35}\pi^2}$
2	$N_2 = \frac{48}{35}\pi^2 N$	$n_2 = \frac{144}{35}\pi^2$	$\ell_2 = \left(\frac{3}{4\pi}\right)^{\frac{1}{2}} \frac{875}{243} \Gamma\left(\frac{2}{3}\right)$	
3	$N_3 = \frac{24}{35}\pi^2 N$	$n_3 = \frac{96}{35}\pi^2$		

Table IV. Cell data in four dimensions

Dimension of simplex	Number	Number incident on a point	Average size	Average size (dual)
0	$N_0 = N$	$n_0 = 1$		
1	$N_1 = \frac{170}{9}N$	$n_1 = \frac{340}{9}$	$\ell_1 = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{3^3}{7 \times 11} \frac{16}{15} \frac{1}{\pi} \Gamma\left(\frac{1}{4}\right)$	$\sigma_3 = \frac{9}{225} (2\pi^2)^{\frac{1}{2}} \frac{1}{\pi} \Gamma\left(\frac{3}{4}\right)$
2	$N_2 = \frac{590}{9}N$	$n_2 = \frac{1770}{9}$	$\ell_2 = 2^{\frac{1}{2}} \frac{2^7}{11^2 \times 13} \frac{16}{17} \frac{1}{\pi} \Gamma\left(\frac{1}{2}\right)$	
3	$N_3 = \frac{715}{9}N$	$n_3 = \frac{2860}{9}$	$\ell_3 = \left(\frac{2}{\pi^2}\right)^{\frac{1}{2}} \frac{2^7 7^2}{3 \times 13^2 \times 57} \frac{18}{17} \frac{1}{\pi} \Gamma\left(\frac{3}{4}\right)$	
4	$N_4 = \frac{286}{9}N$	$n_4 = \frac{1470}{9}$	$\ell_4 = \frac{9}{286}$	

More generally if $n_{k,m}$, $k > m$, ($n_{k,0} \equiv n_k$) denotes the number of k -simplices incident on an m -simplex

$$\sum_{k=m+1}^d (-1)^{k-m-1} n_{k,m} = 1 - (-1)^{d-m} \quad (33)$$

Equivalently since

$$n_{k,m} = \frac{(k+1)!}{(k-m)!(m+1)!} \frac{N_k}{N_m} \quad (34)$$

$$\sum_{k=m}^d \frac{(k+1)!}{(k-m)!(m+1)!} (-1)^{d-k} N_k = N_m \quad (35)$$

The case $m = d$ is trivial, $m = d-1$ or $d-2$ give equivalent information, and so on. Hence equations (35) leave only $\frac{1}{2}d-1$ (d even) or $\frac{1}{2}(d-1)$ (d odd) unknown quantities among the d ratios $N_1/N, \dots, N_d/N$. Thus topology alone fixes them in two dimensions, leaves one unknown in three and four dimensions etc...

- ii) It is possible to obtain the average number n_d of d -simplices incident on a point in closed form. It is given by the integral, generalizing the expression found in (19) for the three-dimensional case, as

$$n_d = \frac{1}{d!} \int d^d x_2 \cdots d^d x_{d+1} \exp -V$$

where V is the volume of the sphere circumscribed to the d -simplex $(0, x_2, \dots, x_{d+1})$ the radius of which we denote by R , thus $V = d^{-1} R^d S_d$ with S_d the area of the unit sphere

$$S_d = \frac{2\pi^{\frac{1}{2}d}}{\Gamma(\frac{1}{2}d)}$$

Taking the origin 0 at the center of the sphere with $\hat{u}_1, \dots, \hat{u}_{d+1}$ the unit vectors pointing along $OM_1, \dots, OM_{d+1}$ and w_d standing for the obvious generalization of (20), one finds

$$n_d = \int_0^\infty dR R^{d^2-1} S_d^{d+1} e^{-R^d S_d/d} \langle w_d \rangle = \frac{S_d d!}{d^2} \langle w_d \rangle$$

with $\langle w_d \rangle$ the average of a $(d+1) \times (d+1)$ determinant

$$\begin{aligned} \langle w_d \rangle &= \int \prod_0^d \frac{d^{d-1} \hat{u}_k}{S_d} \frac{1}{d!} \left| \det \begin{pmatrix} 1 & \cdots & 1 \\ \hat{u}_0 & \cdots & \hat{u}_d \end{pmatrix} \right| \\ &= \int \prod_0^d \frac{d^{d-1} \hat{u}_k}{S_d} \frac{1}{d!} (\det A_{\alpha\beta})^{\frac{1}{2}} \end{aligned}$$

The indices α, β range from 0 to d , $A_{\alpha\beta} = 1 + \hat{u}_\alpha \cdot \hat{u}_\beta$, and the integration is over $d+1$ unit vectors. For a regular simplex $\hat{u}_\alpha \cdot \hat{u}_\beta = \delta_{\alpha\beta} - (1 - \delta_{\alpha\beta})/d$ and $w_{\text{reg}} = [d! d^{\frac{1}{2}d}]^{-1} (1+d)^{\frac{1}{2}(1+d)}$. The end points of the d vectors $\hat{u}_1, \dots, \hat{u}_d$ build a $(d-1)$ -simplex of volume Δ in a hyperplane at a distance h from $\hat{u}_0$ and $w = h\Delta/d$. Let $\hat{u}$ stand for a unit vector along the normal of this hyperplane with two possible orientations. In general

$$\begin{aligned} \int f(\hat{u}_1, \dots, \hat{u}_d) \prod_1^d d^{d-1} \hat{u}_k &= \int \frac{1}{2} d^{d-1} \hat{u} (d-1)! \Delta \prod_1^d d^{d-1} \hat{u}_k \\ &\quad f(\hat{u}_1, \dots, \hat{u}_d) \delta(\hat{u} \cdot [\hat{u}_2 - \hat{u}_1]) \cdots \delta(\hat{u} \cdot [\hat{u}_d - \hat{u}_1]) \end{aligned}$$

Define $\hat{u}_k = \cos \theta_k \hat{u} + \sin \theta_k \hat{v}_k$ where $\hat{v}_k$ is a unit $(d-1)$ -dimensional vector orthogonal to $\hat{u}$, and

$$d^{d-1} \hat{u}_k = \sin \theta_k^{d-3} d \cos \theta_k d^{d-2} \hat{v}_k$$

The arguments of the δ -functions are $\cos \theta_k - \cos \theta_1$, $k \geq 2$, and force $\hat{u}$ to be orthogonal to all $\hat{v}_k$'s. We substitute $f \equiv h\Delta/d$, and observe that $\Delta = \sin \theta^{d-1} w_{d-1}$. Thus performing the integration over $\hat{u}_0$

$$\begin{aligned} \langle w_d \rangle &= \frac{1}{d} \left(\frac{S_{d-1}}{S_d} \right)^d \frac{1}{2} (d-1)! S_d \langle w_{d-1}^2 \rangle \\ &\quad \times \int_0^\pi d \cos \theta (\sin \theta)^{d(d-3)+2(d-1)} \\ &\quad \times \frac{\int_0^\pi d\varphi \sin \varphi^{d-2} |\cos \varphi - \cos \theta|}{\int_0^\pi d\varphi \sin \varphi^{d-2}} \end{aligned}$$

The last ratio comes from the average over the height h . Consequently

$$\langle w_d \rangle = \frac{(d-2)!(d+1)}{d^2} \frac{S_{d-1}^{d+1} \Gamma(\frac{1}{2}) \Gamma(\frac{1}{2}(d^2-1))}{S_d^d \Gamma(\frac{1}{2}d^2)} \langle w_{d-1}^2 \rangle$$

$$\langle w_{d-1}^2 \rangle = \frac{1}{(d-1)!^2} (\det B_{ab}) \quad B_{ab} = 1 + \hat{v}_a \cdot \hat{v}_b \quad 1 \leq a, b \leq d$$

Expanding the $d \times d$ determinant as a sum over products of cycles, observing that each cycle contributes a factorized term, using

$$\langle \hat{v}_a^k \hat{v}_b^{k'} \rangle = \delta^{kk'} \delta_{ab} \frac{1}{d-1}$$

$$\langle B_{12} B_{23} \cdots B_{\ell 1} \rangle = 1 + \langle \hat{v}_1 \cdot \hat{v}_2 \hat{v}_2 \cdot \hat{v}_3 \cdots \hat{v}_\ell \cdot \hat{v}_1 \rangle = 1 + \frac{1}{(d-1)^{\ell-1}}$$

we find

$$\begin{aligned} \eta_d &= \langle \det B_{ab} \rangle \\ &= \sum_{\text{permutations } \mathcal{P}} (-1)^{\mathcal{P}} \prod_{\text{cyclic decomposition of } \mathcal{P}} \left(1 + \frac{1}{(d-1)^{\ell-1}} \right) \end{aligned}$$

We group together all terms with the same cyclic decomposition. Let $\alpha_\ell \geq 0$ be the number of cycles of length ℓ with $\sum_\ell \ell \alpha_\ell = d$. The number of permutations having this cyclic decomposition is

$$\frac{d!}{\prod_{\ell=1}^d \ell^{\alpha_\ell} \alpha_\ell!}$$

and the signature is $(-1)^{\mathcal{P}} = (-1)^{\sum_\ell \ell(\ell-1)\alpha_\ell}$. Therefore

$$\begin{aligned} \frac{1}{d!} \eta_d &= \sum_{\{\alpha_\ell\}} \frac{1}{\prod_{\ell=1}^d \ell^{\alpha_\ell} \alpha_\ell!} \prod_{\ell} (-1)^{\alpha_\ell \ell(\ell-1)} \left[1 + \frac{1}{(d-1)^{\ell-1}} \right]^{\alpha_\ell} \\ &= \text{coefficient of } t^d \text{ in } \exp \sum_{\ell=1}^{\infty} (-1)^{\ell-1} \left[1 + \frac{1}{(d-1)^{\ell-1}} \right] \frac{t^\ell}{\ell} \\ &= \text{coefficient of } t^d \text{ in } (1+t) \left(1 + \frac{t}{d-1} \right)^{d-1} = \frac{1}{(d-1)^{d-1}} \end{aligned}$$

Combining all this information, we get

$$\langle w_d \rangle = \frac{d+1}{d} \frac{S_{d-1}^{d+1}}{[(d-1)S_d]^d} \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{1}{2}(d^2-1))}{\Gamma(\frac{1}{2}d^2)} \quad (36a)$$

and as a result,

$$n_d = \frac{2}{d} \frac{\Gamma(d)}{\Gamma(\frac{1}{2}(d+1))^2} \left[\frac{\Gamma(\frac{1}{2}) \Gamma(\frac{1}{2}d+1)}{\Gamma(\frac{1}{2}(d+1))} \right]^{d-1} \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{1}{2}(d^2+1))}{\Gamma(\frac{1}{2}d^2)} \quad (36b)$$

Asymptotically for large d

$$n_d \sim \frac{2}{d} e^{\frac{1}{2}} (2\pi d)^{d-\frac{1}{2}} \quad (37)$$

The average volume of a d -simplex is

$$\ell_d = \frac{d+1}{n_d} \quad (38)$$

Similarly generalizing the reasoning in low dimensions one can obtain the average area of a cell boundary, call it P , as

$$P = 2^{d+1} \frac{1}{d} \Gamma\left(2 - \frac{1}{d}\right) \frac{1}{\Gamma(d - \frac{1}{2})} [\Gamma(\frac{1}{2}d + 1)]^{2-1/d} \quad (39)$$

Dividing out by n_{d-1} given by equations (31) and (36) we obtain what was called above ℓ_{d-1}

$$\begin{aligned} \ell_{d-1} &= \frac{d^2}{\pi^{\frac{1}{2}}} \frac{\Gamma(d + (d-1)/d)}{[\Gamma(d)]^2} \left[\frac{\Gamma(\frac{1}{2}d + 1)}{\pi^{d/2}} \right]^{(d-1)/d} \\ &\times \frac{[\Gamma(\frac{1}{2}(d+1))]^d}{[\Gamma(\frac{1}{2}d + 1)]^{d-1}} \left[\frac{\Gamma(\frac{1}{2}d^2)}{\Gamma(\frac{1}{2}(d^2+1))} \right]^2 \frac{\Gamma(\frac{1}{2}(d^2+2))}{\Gamma(\frac{1}{2}(d^2+d-1))} \end{aligned} \quad (40)$$

Up to now we have only studied local properties. Even though the points are chosen at random and independently, the Dirichlet construction correlates all of them. It would be of interest to understand the nature of correlations among distant cells. This is an intricate problem, to which no conclusive answer seems to be known. Other interesting constructions involve random packings, in particular of spheres.

As an application of the power of elementary topological considerations such as those developed in this section, let us derive the five regular polyhedra (or Platonic solids) in three dimensions. A finite subgroup of rotations acts transitively as substitutions on the N_0 vertices, N_1 links and N_2 faces, which means that all vertices, all links and all faces respectively, are equivalent. Let q denote the coordination number, i.e. the number of neighbours of a vertex, and $\tilde{q}$ the number of edges of each face, i.e. the number of neighbouring faces. Each link is common to two faces and joins two vertices, and the Euler characteristic χ is equal to 2, thus

$$2N_1 = \tilde{q}N_2 = qN_0 \quad N_0 - N_1 + N_2 = 2$$

Equivalently

$$2(N_0 + N_2) = 4 + \tilde{q}N_2 = 4 + qN_0$$

Consequently $(4 - q)N_0 + (4 - \bar{q})N_2 = 8$, showing that q or $\bar{q}$ has to be smaller than 4. Since both are larger or equal to three (otherwise we have no solid), one concludes that the faces are either triangles ($\bar{q} = 3$) or the vertices have a coordination number equal to 3 ($q = 3$). From the above relations one also deduces that $(6 - q)N_0 = 12 + 2(\bar{q} - 3)N_2$ showing that q can only take the values 3, 4 and 5 since the r.h.s. is larger or equal to 12. Similarly $\bar{q}$ can only take the same values. Thus we have (with a complete duality) either $q = 3$ and $\bar{q} = 3, 4$ or 5 or vice versa. From the original equalities we compute N_1 as

$$N_1 = \frac{2}{(2/q) + (2/\bar{q}) - 1}$$

which has to be an integer. For $q = 3$ and $\bar{q}$ respectively equal to 3, 4 or 5 we find $N_1 = 6, 12$ and 30. Similarly for $\bar{q} = 3$ and $q = 3, 4, 5$. This leads to the table of regular solids, where in the first column we recall the nature of the faces (given by the number $\bar{q}$ of their edges) and in the last column the order of the corresponding rotation group, equal to $2N_1$, since it acts transitively on the oriented links. The solids are divided into three groups, with the octahedron dual to the cube, and the icosahedron dual to the dodecahedron.

Table V. Regular solids in three dimensions.

Faces	$\bar{q}$	q	N_0	N_1	N_2	Order of rotation group
4 triangles (Tetrahedron)	3	3	4	6	4	12
8 triangles (Octahedron)	3	4	6	12	8	24
6 squares (Cube, hexahedron)	4	3	8	12	6	24
20 triangles (Icosahedron)	3	5	12	30	20	60
12 pentagons (Dodecahedron)	5	3	20	30	12	60

Can the reader extend these arguments in higher dimension?

If, to the above list of finite rotation groups, we add the two infinite series of cyclic and dihedral groups, we recover the A-D-E classification encountered in section 3.4 of chapter 9.

11.1.2 Field equations

Given an arbitrary lattice we want to set up a statistical, or field theoretical model, using some of the metrical properties attached to points, links, etc. As a matter of fact we have constructed two lattices. The first one is a simplicial lattice L made of 0, 1, 2, ... simplices. A 0-simplex is a point i (we set $\ell_i \equiv 1$), a 1-simplex is a link joining two neighbours (ij) at a distance $\ell_{ij} = \ell_{ji}$, crossing the hyperplane of face common to the two cells around i and j , and so on. The dual lattice $\bar{L}$ is made of d -, $(d-1)$ -, ... cells. We call i the d -cell pertaining to point i and σ_i its (hyper-) volume, (i, j) is then the $(d-1)$ -cell dual to the link (ij) and common to cells i and j , with σ_{ij} its "area" etc.

Define $\mathcal{F}_0$ as the set of 0-forms, i.e. functions defined on sites, $i \rightarrow \varphi_i$. Similarly $\mathcal{F}_1$ is the set of antisymmetric 1-forms, defined on oriented 1-simplices, i.e. links, $\varphi_{ij} = -\varphi_{ji}$; $\mathcal{F}_2$ is made of antisymmetric 2-forms assigned to 2-simplices $\varphi_{ijk} = -\varphi_{jik} = \dots$, and so on. In parallel, let us call $\bar{\mathcal{F}}_d$ the set of d -densities, namely functions ψ_i defined on positively oriented d -cells. The $(d-1)$ -densities ψ_{ij} , defined on oriented $(d-1)$ -cells build $\bar{\mathcal{F}}_{d-1}$, and so on. The orientation on the cell $(ij\dots)$ and on the dual simplex $(ij\dots)$ are chosen to be compatible, i.e. their product is unity.

The forms $\varphi_i, \varphi_{ij}, \varphi_{ijk}, \dots$ may be considered as restrictions on the lattice of scalar, vector, antisymmetric tensors fields and so on, defined in the continuum. The correspondence is as follows

$$\begin{aligned} \varphi_i &\leftrightarrow \varphi(x_i) \\ \varphi_{ij} &\leftrightarrow \frac{1}{\ell_{ij}} \int_i^j dx^\mu \varphi_\mu(x) \\ \varphi_{ijk} &\leftrightarrow \frac{1}{\ell_{ijk}} \int \int_{(ijk)} dx^\mu \wedge dx^\nu \varphi_{\mu\nu}(x) \\ &\dots \end{aligned} \quad (41)$$

Between the dual vector spaces $\mathcal{F}_p$ and $\bar{\mathcal{F}}_{d-p}$ there exists a natural scalar product. With φ in $\mathcal{F}_p$, ψ in $\bar{\mathcal{F}}_{d-p}$, we have

$$\begin{aligned} p=0 \quad \langle \varphi | \psi \rangle &= \frac{1}{1!} \sum_i \varphi_i \psi_i \\ p=1 \quad \langle \varphi | \psi \rangle &= \frac{1}{2!} \sum_{ij} \varphi_{ij} \psi_{ij} \end{aligned}$$

$$p=2 \quad \langle \varphi | \psi \rangle = \frac{1}{3!} \sum_{ijk} \varphi_{ijk} \psi_{ijk} \quad (42)$$

...

The factorial compensates for overcounting and the sum extends over sites, links, triangles etc. Should one wish it, the scalar product among real quantities could be extended to a Hermitian one among complex ones. The obvious duality between $\mathcal{F}_p$ and $\tilde{\mathcal{F}}_{d-p}$ is set up as

$$\begin{aligned} \varphi &\leftrightarrow \psi \\ \mathcal{F}_p &\leftrightarrow \tilde{\mathcal{F}}_{d-p} \\ \ell_{ijk\dots} \varphi_{ijk\dots} &= \frac{\psi_{ijk\dots}}{\sigma_{ijk\dots}} \end{aligned} \quad (43)$$

where no summation is understood. For any p , this entails the following relation among dimensions

$$[\psi] = [\varphi] [\text{length}]^d \quad (44)$$

The dimensional factor is the volume of a cell when $p=0$. When $p=1$, it is d -times the sum of the volumes of two pyramids of apex i and j respectively and base σ_{ij} , and so on. We denote by $\psi = \tilde{\varphi}$ or $\varphi = \tilde{\psi}$ the above correspondences.

On $\mathcal{F}_p$ we can therefore set up a quadratic potential (or mass) term as

$$p=0 \quad V(\varphi) = \frac{1}{2} \langle \varphi | \tilde{\varphi} \rangle \quad (45)$$

$$p=1 \quad V(\varphi) = \frac{1}{2} \sum_i \sigma_i \varphi_i^2$$

$$p=1 \quad V(\varphi) = \frac{1}{2} \sum_{(ij)} \ell_{ij} \sigma_{ij} \varphi_{ij}^2 \quad (46)$$

$$p=2 \quad V(\varphi) = \frac{1}{2} \sum_{(ijk)} \ell_{ijk} \sigma_{ijk} \varphi_{ijk}^2$$

...

To construct kinetic terms, we need the analogs of gradient and divergence. Define the operator d (the analog of the exterior derivative) as a mapping from $\mathcal{F}_p$ to $\mathcal{F}_{p+1}$ (giving zero on $\mathcal{F}_d$) as

$$d: \mathcal{F}_p \rightarrow \mathcal{F}_{p+1}$$

$$\begin{aligned} (d\varphi)_{ij} &= \frac{\varphi_i - \varphi_j}{\ell_{ij}} \\ (d\varphi)_{ijk} &= \frac{\ell_{ij}\varphi_{ij} + \ell_{jk}\varphi_{jk} + \ell_{ki}\varphi_{ki}}{\ell_{ijk}} \\ (d\varphi)_{i_0 i_1 \dots i_q} &= \sum_{s=0}^q \frac{(-1)^s \ell_{i_0 \dots i_{s-1} i_{s+1} \dots i_q} \varphi_{i_0 \dots i_{s-1} i_{s+1} \dots i_q}}{\ell_{i_0 i_1 \dots i_q}} \end{aligned} \quad (47)$$

Clearly the operator d has a vanishing square

$$d^2 = 0 \quad (48)$$

while from equation (42), its transpose $\tilde{d}$ maps $\tilde{\mathcal{F}}_p$ into $\tilde{\mathcal{F}}_{p+1}$ according to

$$\langle \varphi | \tilde{d}\psi \rangle = \langle d\varphi | \psi \rangle \quad (49)$$

Explicitly

$$\tilde{\mathcal{F}}_{d-1} \rightarrow \tilde{\mathcal{F}}_d \quad (\tilde{d}\psi)_i = \sum_{j(i)} \frac{\psi_{ij}}{\ell_{ij}} \quad (50)$$

$$\tilde{\mathcal{F}}_{d-2} \rightarrow \tilde{\mathcal{F}}_{d-1} \quad (\tilde{d}\psi)_{ij} = \sum_{k(ij)} \frac{\ell_{ij} \psi_{ijk}}{\ell_{ijk}} \quad (51)$$

...

Obviously

$$\tilde{d}^2 = 0 \quad (52)$$

The divergence operator d^* is then obtained by pulling back $\tilde{d}$ using the duality map (43) according to the scheme

$$\begin{array}{ccc} \mathcal{F}_p & \xrightarrow{\text{duality}} & \tilde{\mathcal{F}}_{d-p} \\ \downarrow & & \downarrow \\ \mathcal{F}_{p-1} & \xleftarrow{\text{duality}} & \tilde{\mathcal{F}}_{d-p+1} \end{array} \quad \tilde{d} \quad (53)$$

For instance d^* maps $\mathcal{F}_1$ on $\mathcal{F}_0$. We start from φ_{ij} , apply duality, and get $\psi_{ij} = \tilde{\varphi}_{ij} = \ell_{ij} \sigma_{ij} \varphi_{ij}$. Acting with $\tilde{d}$ we obtain

$$(\tilde{d}\psi)_i = \sum_{j(i)} \frac{\psi_{ij}}{\ell_{ij}} = \sum_{j(i)} \sigma_{ij} \varphi_{ij}$$

Returning by duality to $\mathcal{F}_0$, the final expression is

$$(d^*\varphi)_i = \frac{1}{\sigma_i} \sum_{j(i)} \sigma_{ij} \varphi_{ij} \quad (54)$$

which is seen to be the discrete analog of the divergence. As an example, equation (54) clearly exhibits Gauss's relation between the flux of the electric field and the integral of the charge density. Again

$$d^*2 = 0 \quad (55)$$

and d^* annihilates $\mathcal{F}_0$. One can also construct a similar nilpotent operator $\tilde{d}^*$.

A kinetic part in an action can then be defined as

$$\begin{aligned} K(\varphi) &= V(d\varphi) = \frac{1}{2} \langle d\varphi | \tilde{d}\varphi \rangle \\ &= \frac{1}{2} \sum_{i_0, \dots, i_p} \frac{\sigma_{i_0 \dots i_p}}{\ell_{i_0 \dots i_p}} \left(\sum_{s=0}^p (-1)^s \ell_{i_0 \dots i_{s-1} i_{s+1} \dots i_p} \varphi_{i_0 \dots i_{s-1} i_{s+1} \dots i_p} \right)^2 \end{aligned} \quad (56)$$

Specifically, for scalars and vectors respectively

$$\varphi \in \mathcal{F}_0 \quad K(\varphi) = \frac{1}{2} \sum_{(ij)} \frac{\sigma_{ij}}{\ell_{ij}} (\varphi_i - \varphi_j)^2 \quad (57a)$$

$$\varphi \in \mathcal{F}_1 \quad K(\varphi) = \frac{1}{2} \sum_{(ijk)} \frac{\sigma_{ijk}}{\ell_{ijk}} (\ell_{ij}\varphi_{ij} + \ell_{jk}\varphi_{jk} + \ell_{ki}\varphi_{ki})^2 \quad (57b)$$

As in the continuum theory, the dimension of K is

$$[K] = [\varphi]^2 [\text{length}]^{d-2} \quad (58)$$

To obtain free massless fields, we use the kinetic term as an action. The classical field equations are obtained by looking for extrema of K in the form

$$\delta K = \langle \cdot | \tilde{\delta}\varphi \rangle = 0$$

therefore

$$d^*d\varphi = 0 \quad (59)$$

The operator d^*d has a dimension of $[\text{length}]^{-2}$ and maps $\mathcal{F}_p$ on $\mathcal{F}_p$.

Let us look first at the scalar case. Since $d^*\mathcal{F}_0 = 0$, we can replace d^*d by $(d^*d + dd^*) = (d + d^*)^2 = -\Delta$, where Δ is the Laplacian on scalars, such that

$$(-\Delta\varphi)_i = (d^*d\varphi)_i = \frac{1}{\sigma_i} \sum_{j(i)} \frac{\sigma_{ij}}{\ell_{ij}} (\varphi_i - \varphi_j) \quad (60)$$

Define

$$p_{j,i} = \frac{1}{2d} \frac{\sigma_{ij}\ell_{ij}}{\sigma_i} \quad (61)$$

a dimensionless positive hopping probability from a site i to a neighbouring site j satisfying

$$\sum_j p_{j,i} = 1 \quad (62)$$

Hence

$$\frac{1}{2d} (-\Delta\varphi)_i = \sum_{j(i)} p_{j,i} \frac{\varphi_i - \varphi_j}{\ell_{ij}^2} \quad (63)$$

Show that the functions

$$\varphi_i = a + \mathbf{k} \cdot \mathbf{x}_i \quad (64)$$

where $\mathbf{k}$ is a constant vector, are harmonic, i.e. satisfy

$$(-\Delta\varphi)_i = 0 \quad (65)$$

For $p \geq 1$, the field equations (59) are not equivalent to the Laplace equation, very much as Maxwell's equations for the (four-) vector potential, which correspond to the case $p = 1$, do not entail without a special choice of gauge that each component be harmonic. There also exists here a gauge invariance of the form $\varphi \rightarrow \varphi' = \varphi + d\varphi''$ since $d^*d\varphi = d^*d\varphi'$. For instance, when $p = 1$, we can rewrite equation (59) explicitly as

$$(d^*d\varphi)_{ij} \equiv \frac{1}{\sigma_{ij}} \sum_{k(ij)} \frac{\sigma_{ijk}}{\ell_{ijk}} (\ell_{ij}\varphi_{ij} + \ell_{jk}\varphi_{jk} + \ell_{ki}\varphi_{ki}) = 0 \quad (66)$$

Show that if one assumes that φ_{ij} results from an underlying vector field $A_\mu(\mathbf{x})$ defined in the continuum

$$\varphi_{ij} \leftrightarrow \frac{1}{\ell_{ij}} \int_i^j d\mathbf{x}^\mu A_\mu(\mathbf{x})$$

then equation (66) is a discretized version of Maxwell's equation

$$\partial^\mu F_{\mu\nu} \equiv \Delta A_\nu - \partial_\nu(\partial \cdot A) = 0$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In the continuum we have

$$\Delta A_\nu = \partial^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\nu(\partial \cdot A)$$

and more generally a similar definition of the Laplacian extends to any antisymmetric tensor field. In discretized form, the Laplacian is defined as

$$-\Delta = d^*d + dd^* = (d + d^*)^2 \quad (67)$$

Thus $-\Delta$ appears as the square of the so-called Dirac-Kähler operator defined on a set $\Phi \equiv \{\varphi_i, \varphi_{ij}, \varphi_{ijk}, \dots\}$ of forms of any degree as

$$(d + d^*)\Phi = 0 \quad (68)$$

This has been suggested as a possible extension of Dirac's equation on a lattice, although the tensorial transformation properties do not correspond to the standard ones in the continuum.

One can study on random lattices other types of field equations than the free massless linear ones. Attempts have been made at numerical investigations of four-dimensional gauge theories on such lattices. The task is formidable. Even two-dimensional models present serious difficulties. One has to face the question of how the geometric and topological disorder interact with the intrinsic dynamics, meaning a competition between several length scales. In all cases, we are looking for the way in which the large distance properties of the random lattice allow the restoration of continuous symmetries.

11.1.3 The spectrum of the Laplacian

The simplest example of a comparison between the behaviour of a dynamical system on a random lattice and a regular one, is the study of the spectrum of the scalar Laplacian given by equations (60) and (63). We have already presented in chapter 10, section 1.4, the easiest one-dimensional case. Let us look now at some general properties in arbitrary dimension. The difficulty is that we can not use Fourier analysis anymore. Let us consider a large, but finite box, with N sites, and let $\varphi^{(\alpha)}$ and ω_α stand for the eigenfunctions and eigenvalues of the Laplacian. The eigenfunctions are normalized according to

$$\langle \varphi^{(\alpha)} | \tilde{\varphi}^{(\beta)} \rangle \equiv \sum_i \sigma_i \varphi_i^{(\alpha)} \varphi_i^{(\beta)} = \delta^{\alpha\beta} \quad (69)$$

An arbitrary function φ such that

$$\gamma_\alpha = \langle \varphi | \tilde{\varphi}^{(\alpha)} \rangle \equiv \sum_i \sigma_i \varphi_i \varphi_i^{(\alpha)} \quad (70)$$

admits an expansion of the form

$$\varphi_i = \sum_\alpha \gamma_\alpha \varphi_i^{(\alpha)} \quad (71)$$

with

$$(-\Delta\varphi)_i = \sum_\alpha \gamma_\alpha \omega_\alpha \varphi_i^{(\alpha)} \quad (72)$$

Here the functions $\varphi^{(\alpha)}$ which replace the plane waves, and the corresponding eigenvalues ω_α are still stochastic variables. We look for the most likely, or rather average distribution of eigenvalues.

As the number of sites N tends to infinity, with the index α running through N values, let $N\rho(\omega)$ denote the number of eigenvalues in an interval $d\omega$ around ω . We have

$$\int_0^\infty d\omega \rho(\omega) = 1 \quad (73)$$

We expect that the thermodynamic limit ensures that $\rho(\omega)$ is already the most likely distribution. Our aim will be to ascertain the behaviour of $\rho(\omega)$ for small as well as large values of ω . For $\omega \rightarrow 0$, if eigenfunctions can be approximated by plane waves over large distances, it is likely that $d\omega \rho(\omega) \sim d^d \mathbf{k} / (2\pi)^d$ with $\omega = \mathbf{k}^2$, i.e.

$$\rho(\omega) \underset{\omega \rightarrow 0}{\sim} \frac{\omega^{\frac{1}{2}d-1}}{(4\pi)^{\frac{1}{2}d} \Gamma(\frac{1}{2}d)} \quad (74)$$

To lend credence to this estimate, let us study how the Laplacian acts on the functions

$$\varphi_i(\mathbf{k}) = \exp i\mathbf{k} \cdot \mathbf{x}_i \quad (75)$$

Consider the local quantities

$$\omega_i(\mathbf{k}) = \varphi_i^*(\mathbf{k}) (-\Delta\varphi(\mathbf{k}))_i = \frac{1}{\sigma_i} \sum_{j(i)} \frac{\sigma_{ij}}{\ell_{ij}} [1 - e^{i\mathbf{k} \cdot (\mathbf{x}_j - \mathbf{x}_i)}] \quad (76)$$

For small k , $\omega_i(k)$ is of order k^2 , according to equations (64) and (65). The average

$$\overline{\omega(k)} = \frac{\sum_i \sigma_i \omega_i(k)}{\sum \sigma_i} \quad (77)$$

can be interpreted as the expectation value of $(-\Delta)$ in the state $\varphi(k)$. For $N \rightarrow \infty$, there exists no preferred direction in the lattice, so the expression (77) can be averaged over the directions of the wavevector k

$$\overline{\omega(k)} = \sum_{s=1}^{\infty} (-1)^{s-1} \frac{k^{2s}}{2^{2s}} \frac{1}{s!} \frac{\Gamma(\frac{1}{2}d)}{\Gamma(\frac{1}{2}d+s)} \ell^{2s-1} \quad (78)$$

The last average stands for

$$\ell^{2s-1} = \frac{1}{N} \sum_i \sum_{j(i)} \sigma_{ij} \ell_{ij}^{2s-1} \quad (79)$$

It can be computed following Meijering's argument of section 1.1, with the result that

$$\begin{aligned} \overline{\omega(k)} &= \frac{2^d}{d\pi^{\frac{1}{2}}} \Gamma(1 + \frac{1}{2}d)^2 \sum_{s=1}^{\infty} \left[\frac{-\Gamma(1 + \frac{1}{2}d)^{2/d}}{\pi} \right]^{s-1} \\ &\quad \times \frac{\Gamma(s + \frac{d-1}{2}) \Gamma(2 + \frac{2(s-1)}{d}) k^{2s}}{\Gamma(s + \frac{d}{2}) \Gamma(s + d - 1) s!} \end{aligned} \quad (80)$$

This rather cumbersome formula reduces to simple expressions in low dimensions, $d = 1$ or 2 . We have

$$d = 1 \quad \overline{\omega(k)} = \ln(1 + k^2) \quad (81a)$$

$$d = 2 \quad \overline{\omega(k)} = 2\pi \left[1 - e^{-k^2/2\pi} I_0(k^2/2\pi) \right] \quad (81b)$$

where $I_0(z)$ is the modified Bessel function

$$I_0(z) = \sum_{n=0}^{\infty} \left(\frac{z^2}{4} \right)^n \frac{1}{(n!)^2} \underset{z \rightarrow \infty}{\sim} \frac{e^z}{\sqrt{2\pi z}} \left(1 + \frac{1}{8z} + \dots \right) \quad (82)$$

In all cases $\overline{\omega(k)}$ behaves as k^2 for small k . This comforts the estimate (74). The computation of $\overline{\omega(k)}$ for any k^2 allows us also to draw some interesting conclusions. In dimension one, we observe that it is unbounded, giving a proof that the support of $\rho(\omega)$ is also unbounded. However, in dimension larger than one, $\overline{\omega(k)}$ is bounded as k^2 tends to infinity. For instance for $d = 2$,

$\overline{\omega(k)} \rightarrow 2\pi(1 - 1/|k| + \dots)$. In general, this limit can be obtained by dropping the oscillatory term in averaging (76)

$$\lim_{k \rightarrow \infty} \overline{\omega(k)} = \pi d \frac{\Gamma(2 - 2/d)}{[\Gamma(1 + \frac{1}{2}d)]^{2/d}} \quad (83)$$

showing that the support of $-\Delta$ extends at least up to this value.

It is most likely that for small ω , in large enough dimension, the eigenstates are extended ones, with unit probability. At the other extreme, we expect localized states for large ω . A very crude argument, based on this expectation, yields an estimate of $\rho(\omega)$ for ω tending to infinity. Looking at the form of the Laplace operator (60) or (63), we see that "weak links", where σ_{ij}/ℓ_{ij} is exceptionally large, play a significant role in localizing the wavefunctions. Assume that σ_{ij}/ℓ_{ij} is large due to the unfrequent occurrence of two points i, j very close to each other. A bold and extreme approximation is that an eigenstate has a wavefunction φ essentially concentrated on the two neighbouring sites i, j , with

$$\begin{aligned} \omega \varphi_i &= \frac{1}{\sigma_i} \frac{\ell_{ij}}{\sigma_{ij}} (\varphi_i - \varphi_j) \\ \omega \varphi_j &= \frac{1}{\sigma_j} \frac{\ell_{ij}}{\sigma_{ij}} (\varphi_j - \varphi_i) \end{aligned} \quad (84)$$

Therefore

$$\omega = \frac{\sigma_{ij}}{\ell_{ij}} \left(\frac{1}{\sigma_i} + \frac{1}{\sigma_j} \right) \quad (85)$$

It is of course assumed that the right-hand side of this expression is very large (with a small probability). Since i and j are very close, we can think, in the limit, of the two neighbouring cells as a unique cell, call it $\mathcal{C}$, partitioned by a plane bisecting the infinitesimal segment (ij) , or equivalently an arbitrary plane through the center of $\mathcal{C}$, cutting it into two parts of volumes σ_+ and σ_- , and having an area σ_0 inside $\mathcal{C}$. Then

$$\rho(\omega) d\omega \underset{\ell \rightarrow 0}{\sim} p(\ell) d\ell \quad (86a)$$

$$\omega \sim \frac{1}{\ell} \sigma_0 \left(\frac{1}{\sigma_+} + \frac{1}{\sigma_-} \right) \quad (86b)$$

where $p(\ell)d\ell$ behaves as $S_d \ell^{d-1}d\ell$ and the average in (86b) is both on the cell C and the direction of the plane. This results in

$$\rho(\omega) \geq \frac{2\pi^{\frac{1}{2}d}}{\Gamma(\frac{1}{2}d)} \frac{1}{\omega^{d+1}} \quad (87)$$

This power law decrease is in agreement with the unit integral of $\rho(\omega)$, equation (73), and is verified for $d = 1$. A tail in the spectrum of the type (87) is of course absent in the case of a regular lattice. The above argument shows that it is related to localization effects. To determine the localization threshold ω_c is a delicate matter. While for $d = 1, \omega_c = 0$, it is not known what is even the lower critical dimension (possibly $d = 2$) where ω_c departs from zero.

- i) Show that from equation (83) one can derive the exact sum rule

$$\int_0^\infty d\omega \rho(\omega) \omega = \pi d \frac{\Gamma(2 - 2/d)}{[\Gamma(1 + \frac{1}{2}d)]^{2/d}} \quad (88)$$

This is infinite for $d = 1$ due to a divergence at large ω , in agreement with the estimate (87), but eventually grows like d , due to the suppression of small and large ω 's in large dimension.

- ii) Derive the inequality

$$\frac{\langle \omega^2 \rangle}{\langle \omega \rangle^2} - 1 \geq \frac{1}{q} \quad (89)$$

where q is the average coordination number of the lattice, and prove that it becomes an equality on any regular lattice. As q grows very rapidly with d on a random lattice, and $\rho(\omega)$ tends to be strongly peaked, it is likely that (89) becomes asymptotically an equality.

11.2 Random surfaces

We will not repeat the motivations for attempting to find models of random surfaces. A compact and basic formulation is still missing. As a consequence the following presentation will necessarily be fragmentary. We have to give a suitable definition of a surface since it involves various degrees of internal or external structure. We first have topological properties such as connectedness, orientability, Euler characteristics, number of boundaries.

The generalization in higher dimension implies homology, and if a differentiable structure is assumed, dual cohomology groups. We have then the metric properties, parallel transport and associated curvature, with generalizations to other fibre bundles and their connections on surfaces. One can also study complex structures. These are intrinsic properties, but one can also wish to investigate the embeddings in other manifolds, in practical terms mostly Euclidean (or Minkovskian) space. An important circumstance is the one of interfaces, mostly in $\mathcal{R}^3$ of course.

An aspect, which distinguishes surfaces from curves, is the possible existence of nontrivial boundaries. Typical of these questions is the Plateau problem of finding minimal surfaces bounding a given curve, or the study of Wilson loops in gauge theories.

Yet another question, which was not on the forefront for random curves, but could also have been discussed in this case, is the one of reparametrization invariance, or general covariance as it is called in the context of gravitation theory. A correlative point of view is the one of dynamical generation of surfaces. A curve could be viewed as the path traced by a structureless point as it evolves in time, but a surface is generated by a string with infinitely many degrees of freedom (hence the name *string field theory* now in common use).

To sum up, the much richer set of possibilities afforded by the concept of surface opens a vast new domain. We start with a short description of piecewise linear manifolds as exemplified by triangulated surfaces.

11.2.1 Piecewise linear surfaces

We consider an assembly of triangles endowed with a flat Euclidean metric and compatible incidence relations among the edges (and therefore the vertices). This implies of course the equality of lengths for identified sides. We assume for simplicity the resultant manifold to be connected, without boundary, and orientable. Its topology is then dictated by its Euler characteristic $\chi = 2 - 2g$ where g (denoted H in chapter 10 section 4) is the genus, or number of handles. The incidence relations define the notion of nearest neighbour pairs (or links) and 2-simplices (the original triangles) and the construction allows us to compute distances. The mani-

fold has an obvious differentiable structure except possibly at singular points (the vertices). As usual N_0, N_1, N_2 are the numbers of vertices, links and triangles related by

$$2N_1 = 3N_2 \quad \chi = N_0 - N_1 + N_2 \quad (90a)$$

Thus in terms of the number N of vertices

$$\begin{aligned} N_0 &= N \\ N_1 &= 3(N - \chi) \\ N_2 &= 2(N - \chi) \end{aligned} \quad (90b)$$

We shall discuss later the question of inequivalent triangulations and their statistical weights.

A practical way of implementing the above construction is to pick a model, or parameter space, as a given piecewise flat triangulated connected and orientable manifold, to map its N points in Euclidean space $\mathcal{R}^d$ and to interpolate linearly between the images of neighbours. This produces in the target space another triangulated surface (with possible self-intersections), or equivalently we may consider that it endows the original model with a different metric structure. As an example, we may start from a finite region in a Euclidean plane carrying a regular triangular lattice with the shape of a torus ($g = 1$) and allow for various maps in $\mathcal{R}^d$. There are of course numerous other ways of defining surfaces such as those encountered in chapter 6 when discussing lattice gauge theories, as collections of plaquettes from regular lattices with specific rules for assembling them.

The triangles are Euclidean, hence their internal angles $\theta_1, \theta_2, \theta_3$ add to π . For each triangle we can therefore split unity into $1 = (\theta_1 + \theta_2 + \theta_3)/\pi$, with each θ between 0 and π . Summing over triangles yields N_2 , but we can also collect first all θ 's pertaining to the same vertex, then sum over vertices. Call θ_i the sum at each vertex. Then $N_2 = \sum_i \theta_i/\pi$. From equation (90) this is also $2N - 2\chi = -2\chi + 2\sum_i 1$. Identification leads to

$$\chi = \sum_i \left(1 - \frac{\theta_i}{2\pi}\right) \quad (91)$$

which is the discrete form of the Gauss-Bonnet formula, in terms of deficit angles $2\pi - \theta$. When the deficit angle vanishes at a vertex the corresponding triangles fit in flat space and do not contribute to χ . We have the following interpretation: deficit

angle $\Leftrightarrow$ frustration from a planar situation $\Leftrightarrow$ curvature. Positive curvature is related to a positive deficit angle ($\theta_i < 2\pi$), negative curvature means overabundance of "matter" and negative deficit angle ($2\pi < \theta_i$).

This can be understood in the example of a sphere of radius r , with curvature $R = 1/r^2$, the product of the two inverse principal radii of curvature, here equal. For a spherical triangle with inner angles $\alpha_1, \alpha_2, \alpha_3$, the area A is given by $RA = (\alpha_1 + \alpha_2 + \alpha_3 - \pi)$. The total amount that a tangent vector has rotated in one circumnavigation (rounding vertices) is

$$\begin{aligned} \theta &= (\pi - \alpha_1) + (\pi - \alpha_2) + (\pi - \alpha_3) \\ &= 2\pi - RA \quad \text{i.e.} \quad RA = 2\pi - \theta \end{aligned}$$

the total angular deficit. On the triangulated surface, the curvature is concentrated at the vertices, with RA taking a finite limit as A shrinks to zero and R tends to have a δ -function singularity. With the present normalization (which may differ from the one found in other sources), it follows that the Gauss-Bonnet formula has the following form in the continuum

$$\chi = \frac{1}{2\pi} \int dA R \quad (92)$$

i.e. $\chi = 2$ for a sphere where $A = 4\pi r^2, R = r^{-2}$.

For surfaces, curvature is a scalar quantity. In a discrete setting, equation (91) implies a dissymmetry between positive and negative curvature, as we cannot have more than a 2π deficit ($\theta = 0$), while one can have as much as one wants of "extra matter" for large θ .

One can envision higher dimensional analogs of piecewise flat manifolds, with d -dimensional simplices replacing triangles. The Euler characteristics $\chi = N_0 - N_1 + N_2 - N_3 + \dots$ remains a topological invariant. Indeed, adding a vertex will produce the substitutions $N_0 \rightarrow N_0 + 1, N_1 \rightarrow N_1 + C_{d+1}^1, N_2 \rightarrow N_2 + C_{d+1}^2, \dots$, with the last change being $N_d \rightarrow N_d + C_{d+1}^d - 1$. Therefore $\chi \rightarrow \chi + (1 - 1)^{d+1} = \chi$. One would like to get a relation as simple as (91) relating χ to metric properties, with a smooth limit in the continuum case and such that contributions from regions where the simplices fit in flat space vanish (as was true for the expression

(91)). This goal can apparently not be fulfilled. The following is the best substitute.

Consider a p -simplex in the triangulation incident on a given vertex and round off the corresponding corner, to a spherical shape, which by rescaling corresponds to a fraction $\varphi^{(p)}$ of the area of the unit $(p-1)$ -dimensional sphere S_{p-1} . The case of a one-simplex - a link - corresponds to the sphere S_0 , a set of two points, in which case we assign $\varphi^{(1)} = \frac{1}{2}$ to each end point. For a vertex itself, we define $\varphi^{(0)} = 1$. If we sum the $\varphi^{(p)}$'s over the $p+1$ vertices of a given p -simplex, we obtain the full area of S_{p-1} , thus the $\varphi^{(p)}$ add to unity. The sum of unity over the p -simplices being N_p , we have, by collecting first the contributions pertaining to a vertex i denoted $\varphi_i^{(p)}$

$$\varphi_i^{(p)} = \sum_{\substack{p \text{ simplices} \\ \text{incident on } i}} \varphi^{(p)}$$

the obvious relation

$$N_p = \sum_i \varphi_i^{(p)}$$

Thus for the Euler characteristic

$$\chi = \sum_i \sum_{p=0}^d (-1)^p \varphi_i^{(p)} \quad (93)$$

When $d=2$ we recover formula (91), but in higher dimension this expression does not exhibit explicitly that contributions from flat regions vanish. It is possible to re-express (93) in terms of interior angles, but in any case the relation with the curvature tensor is hard to see (Cheeger, Müller and Schrader).

We also note here that the four-dimensional Hilbert-Einstein action for gravity has been discretized by Regge for a simplicial complex in the form

$$\sum a_2 \left(1 - \frac{\theta}{2\pi}\right)$$

generalizing equation (92) where the sum runs over 2-simplices of area a_2 with θ the corresponding sum of dihedral angles. The conditions under which this approximates the continuous version

$$\int d^4x \sqrt{g} R$$

require some careful analysis (Feinberg, Friedberg, Lee and Ren).

Returning to triangulated surfaces with their metric properties, we can easily generalize the field equations for random lattices. In the following, we shall be interested mostly by free massless scalar fields. If the surface were smooth with local parameters α^1, α^2 , and a metric positive definite tensor $g_{ab}(\alpha)$ such that $g = \det g_{ab}$, with g^{ab} the inverse of g_{ab} , we would write a reparametrization invariant action as

$$S = \int d^2\alpha \sqrt{g} \frac{1}{2} g^{ab}(\alpha) \partial_a \varphi \partial_b \varphi \quad (94)$$

We assume for simplicity that the surface has no boundary and observe that in general one would require several overlapping coordinate patches to be able to define the above integral. If instead of taking real values, φ is the image of a map from the surface to a Riemannian d -dimensional manifold M , with local coordinates φ^μ , and a metric $G_{\mu\nu}(\varphi)$, the action would be replaced by

$$S^{(M)} = \int d^2\alpha \sqrt{g} \frac{1}{2} g^{ab}(\alpha) \partial_a \varphi^\mu \partial_b \varphi^\nu G_{\mu\nu}(\varphi) \quad (95)$$

This defines a general σ -model. As a particular case, if the manifold M coincides with the Euclidean space $\mathcal{R}^d$, with a flat metric $G_{\mu\nu}(\varphi) = \delta_{\mu\nu}$, the action (95) amounts to the sum of d -copies of the scalar action (94).

Let us look first at the structureless one-component action and extend its definition to a piecewise linear triangulated surface embedded in $\mathcal{R}^d$. Take on each triangle with vertices x_1, x_2, x_3 a set of barycentric coordinates

$$x(\alpha) = \alpha^1 x_1 + \alpha^2 x_2 + \alpha^3 x_3, \quad \alpha_i \geq 0, \quad \alpha^1 + \alpha^2 + \alpha^3 = 1 \quad (96)$$

If φ takes values φ_i at the vertices we extend it linearly inside the triangle through

$$\varphi(\alpha) = \alpha^1 \varphi_1 + \alpha^2 \varphi_2 + \alpha^3 \varphi_3 \quad (97)$$

this being a natural harmonic function. Computing the action (94) as a sum over triangles, we find

$$S_{\text{discrete}} = \sum_{(ijk)} \frac{1}{8} \left[\varphi_i(x_j - x_k) + \varphi_j(x_k - x_i) + \varphi_k(x_i - x_j) \right]^2 / \ell_{ijk} \quad (98)$$

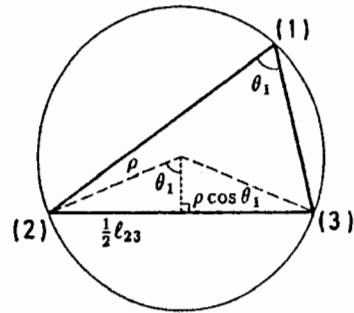


Fig. 4 Construction of the dual quantities.

The sum is over triangles of area ℓ_{ijk} , related to the lengths of the sides through

$$16\ell_{ijk}^2 = (\ell_{ij} + \ell_{jk} + \ell_{ki})(\ell_{ij} + \ell_{jk} - \ell_{ki})(\ell_{ij} - \ell_{jk} + \ell_{ki})(-\ell_{ij} + \ell_{jk} + \ell_{ki}) \quad (99)$$

Translations, rotations or constant shifts of the field leave the action invariant. Henceforth we will omit the suffix "discrete", since the context will clearly indicate when we deal with triangulated surfaces.

In spite of the fact that an *a priori* notion of duality seems to be missing, it is possible to give to the r.h.s. in (98) a form similar to the one corresponding to flat random lattices (equation (57a)).

The quantities $\ell_i = 1$, ℓ_{ij} , and ℓ_{ijk} have obvious definitions the second being the length of the side (ij) , the third the area of the triangle ℓ_{ijk} as above. Define now $\sigma_{ijk} = 1/\sigma_{ij}$, corresponding to a (virtual) dual lattice as follows. Pick one triangle and call it $(1,2,3)$ for short with interior angles $\theta_1, \theta_2, \theta_3$ between 0 and π . The corresponding contribution to the action (98) is

$$S_{123} = \frac{1}{4} [\cot \theta_1 (\varphi_2 - \varphi_3)^2 + \cot \theta_2 (\varphi_3 - \varphi_1)^2 + \cot \theta_3 (\varphi_1 - \varphi_2)^2] \quad (100)$$

The coefficients can take either sign, but the whole expression is positive. Let ρ denote the radius of the circle of center O circumscribed to the triangle. We have

$$\frac{1}{2} \cot \theta_i = (\rho \cos \theta_i) / \ell_{jk}$$

with $\rho \cos \theta_i$ the (algebraic) distance of O to the edge jk (positive if i and O are on the same side of the chord jk). Two and only two triangles $(23i_1)$ and $(23i_2)$ share a given edge (23) . We define the indefinite quantity σ_{23} , the (algebraic) length of the virtual link dual to (23) , as the sum of the two quantities $\rho_1 \cos \theta_{i_1}$ and $\rho_2 \cos \theta_{i_2}$ pertaining to these two triangles. Thus

$$\sigma_{ij} = \rho_1 \cos \theta_{i_1} + \rho_2 \cos \theta_{i_2} \quad (101)$$

in such a way that $\frac{1}{2} \sigma_{ij} / \ell_{ij}$ will be the coefficient of $(\varphi_i - \varphi_j)^2$ in the action. This is consistent with the flat lattice definition and yields

$$S = \frac{1}{2} \sum_{(ij)} \frac{\sigma_{ij}}{\ell_{ij}} (\varphi_i - \varphi_j)^2 \quad (102)$$

The positivity of the action entails inequalities such as

$$\sum_{j(i)} \frac{\sigma_{ij}}{\ell_{ij}} > 0$$

The remaining quantity to be defined is σ_i , the area of the virtual cell i , which should satisfy $\sum_i \sigma_i = \sum_{(ijk)} \ell_{ijk} = A$, where A stands for the total area. A natural choice fulfilling this condition is

$$\sigma_i = \frac{1}{4} \sum_{j(i)} \sigma_{ij} \ell_{ij} \quad (103)$$

which reduces, in the flat case, to area of the Voronoi cell. In the present context, some σ_i 's can however be negative. It should be observed that the construction is well-defined once the lengths ℓ_{ij} of the edges are given. Thus an abstract collection of metric triangles with compatible incidence relations is all that is required, independently of any embedding.

With the ℓ 's and σ 's as above, the definition of the operators d and d^* goes through as in equations (47), (53) and (54), with the obvious restrictions to a two-dimensional manifold. Thus

$$\begin{aligned} \varphi_i &\rightarrow (d\varphi)_{ij} = (\varphi_i - \varphi_j) / \ell_{ij} \\ \varphi_{ij} &\rightarrow (d\varphi)_{ijk} = (\ell_{ij}\varphi_{ij} + \ell_{jk}\varphi_{jk} + \ell_{ki}\varphi_{ki}) / \ell_{ijk} \end{aligned} \quad (104a)$$

and

$$\begin{aligned} \varphi_{ij} &\rightarrow (d^*\varphi)_i = \frac{1}{\sigma_i} \sum_{j(i)} \sigma_{ij} \varphi_{ij} \\ \varphi_{ijk} &\rightarrow (d^*\varphi)_{ij} = \frac{1}{\sigma_{ij}} \sum_{k(ij)} \varphi_{ijk} \end{aligned} \quad (104b)$$

The scalar Laplacian is

$$(-\Delta\varphi)_i = (d^*d\varphi)_i = \frac{1}{\sigma_i} \sum_{j(i)} \frac{\sigma_{ij}}{\ell_{ij}} (\varphi_i - \varphi_j) \quad (105)$$

It enables one, in the absence of boundaries, to rewrite equation (98) and (102) as

$$S = \frac{1}{2} \sum_i \varphi_i (-\Delta\varphi)_i \quad (106)$$

On a compact surface, the action is positive as soon as φ is not a constant, as readily follows from (98), thus the only harmonic functions are constants and $-\Delta$ is a non-negative operator. For pseudoscalar fields, the action, analogous, but distinct, from (102), is

$$\tilde{S} = \frac{1}{2} \sum_{(ijk)} \ell_{ijk} \varphi_{ijk} (dd^*\varphi)_{ijk} = \frac{1}{2} \sum_{(ij)} \frac{\ell_{ij}}{\sigma_{ij}} (\varphi_{ijk_1} - \varphi_{ijk_2})^2 \quad (107)$$

where (ijk_1) and (ijk_2) are the two triangles which share the link (ij) . At least if all σ_{ij} 's are positive and since the surface is orientable, the unique harmonic pseudoscalar is a multiple of η_{ijk} , the orientation tensor with values ± 1 on each (oriented) triangle.

For vector fields φ_{ij} , let us show that harmonicity is equivalent to having both a vanishing curl and divergence

$$[(dd^* + d^*d)\varphi]_{ij} = 0 \Leftrightarrow \{(d\varphi)_{ijk} = 0, (d^*\varphi)_i = 0\} \quad (108)$$

Indeed if φ_{ij} is harmonic, $\psi_i = (d^*\varphi)_i$ satisfies

$$d^*d\psi = d^*dd^*\varphi = -d^{*2}d\varphi = 0$$

so that ψ_i is a constant ψ given by

$$\sum_i \sigma_i \psi_i = \psi A = \sum_i \sigma_i (d^*\varphi)_i = \sum_{(ij)} \sigma_{ij} \varphi_{ij} = 0$$

from $\varphi_{ij} = -\varphi_{ji}$. Hence $\psi = 0$, and we have the second part in equation (108). Therefore $d^*d\varphi = 0$, meaning that $d\varphi$ is harmonic, hence a multiple $\tilde{\psi}$ of η_{ijk} . Again

$$\tilde{\psi} = \sum_{(ijk)} \eta_{ijk} \ell_{ijk} (d\varphi)_{ijk} = \sum_{(ijk)} \eta_{ijk} (\ell_{ij} \varphi_{ij} + \ell_{jk} \varphi_{jk} + \ell_{ki} \varphi_{ki}) = 0 \quad (109)$$

where the last equality follows by adding the two contributions for each factor φ_{ij} . Thus $\tilde{\psi} = 0$.

The counting of linearly independent harmonic vector fields yields a classical relation with topology. There are N_1 independent components φ_{ij} . Equation (108) implies $(N_0 - 1) + (N_2 - 1)$

conditions since $\sum_{(ijk)} \ell_{ijk} \eta_{ijk} (d\varphi)_{ijk} = 0$ and $\sum_i \sigma_i (d^*\varphi)_i = 0$. Therefore there are $N_1 - N_0 - N_2 + 2 = 2 - \chi = 2g$ linearly independent harmonic vector fields. We recall that the same counting holds in the continuum and, when an analytic structure is prescribed, this amounts to the existence of g holomorphic vector fields or equivalently g abelian differentials of the first kind (i.e. holomorphic differentials).

11.2.2 The conformal anomaly and the Liouville action

Given a compact, connected, oriented surface, without boundaries, with its metric structure, our aim is to promote classical massless field theory to its quantum version, namely to use in the discrete case the exponential of the scalar action (102) as a Boltzmann weight in a path integral over the scalar field φ_i . For the time being, the geometry is frozen and we assume that the surface is regular enough for the various σ 's to be positive.

We proceed by setting up the Gaussian path integral in such a way that it yields a result proportional to $[\det'(-\Delta)]^{-1/2}$ where the prime indicates the omission of the zero mode, or equivalently the determinant of $-\Delta$ in the space orthogonal to a constant φ with respect to the square norm $\sum_i \sigma_i \varphi_i^2$, it being understood that φ_i 's are real. For a finite triangulated surface, the operator Δ is a finite $N \times N$ matrix, with $N \equiv N_0$ the number of vertices.

Let $E_1, \dots, E_{N-1}$ be the positive eigenvalues of $-\Delta$ and $E_0 = 0$ corresponds to the unique zero mode. The associated orthonormal set of eigenfunctions will be denoted $\varphi^{(0)}, \varphi^{(1)}, \dots, \varphi^{(N-1)}$. Expanding the field as

$$\varphi_i = \sum_{0 \leq n \leq N-1} C_n \varphi_i^{(n)} \quad (110)$$

it is seen that

$$S = \frac{1}{2} \sum_{1 \leq n \leq N-1} E_n C_n^2 \quad (111)$$

We define the partition function Z as

$$\begin{aligned} Z &= \frac{1}{[\det'(-\Delta)]^{\frac{1}{2}}} = \frac{1}{\left(\prod_{1 \leq n \leq N-1} E_n\right)^{\frac{1}{2}}} \\ &= \int \prod_n \frac{dC_n}{\sqrt{2\pi}} \delta(C_0) \exp -S \end{aligned} \quad (112)$$

The real eigenfunctions are normalized according to

$$\sum_i \sigma_i \varphi_i^{(n_1)} \varphi_i^{(n_2)} = \delta_{n_1, n_2} \quad (113)$$

As a result, the σ_i 's being positive,

$$\prod_0^{N-1} dC_n = \left(\prod_0^{N-1} \sigma_i \right)^{\frac{1}{2}} \prod_i d\varphi_i \quad (114)$$

and

$$\varphi^{(0)} = \frac{1}{\sqrt{A}} \quad (115)$$

Thus we rewrite the path integral in (112) as

$$Z = \sqrt{2\pi A} \int \prod_i \frac{d\varphi_i}{\sqrt{2\pi}} \left(\prod_i \sigma_i \right)^{\frac{1}{2}} \delta \left(\sum_i \sigma_i \varphi_i \right) e^{-S(\varphi)} \quad (116)$$

This could readily be generalized to include a source term either in the form $\sum_i J_i \varphi_i$ or $\sum_i \sigma_i j_i \varphi_i$.

This is as far as one can go in general terms in the piecewise linear case. We can however proceed much further if we make the bold assumption that, at the price of some renormalization, the discrete and well-defined path integral is a valid approximation to a continuous theory on a smooth surface with the action written as in equation (94). We then observe that the absence of a mass term implies that it is conformally invariant, in the sense that the combination $\sqrt{g}g^{ab}$ is unaffected by a point dependent rescaling of the metric $g_{ab}(\alpha) \rightarrow \rho^2(\alpha)g_{ab}(\alpha)$, which amounts to rescaling lengths locally by the factor ρ . While this holds at the classical level, it fails at the quantum level, leading to the conformal anomaly amply discussed in chapter 9.

In two dimensions, it is always possible in a coordinate patch homeomorphic to the interior of a circle to find appropriate coordinates in which an arbitrary metric takes the form

$$g_{ab}(\alpha) = \rho^2(\alpha) \delta_{ab} \quad (117)$$

For instance, on the upper hemisphere of a sphere of radius r , a stereographic projection yields for the square length element

$$ds^2 = \frac{(d\alpha^1)^2 + (d\alpha^2)^2}{(1 + (\alpha_1^2 + \alpha_2^2)/4r^2)^2} \quad (118)$$

In such a neighbourhood, the coordinates themselves are harmonic functions (these are the so-called thermal coordinates), as they are in flat space. The curvature R (r^{-2} for the sphere) is given by

$$R = -\Delta \ln \rho = -\frac{1}{\rho^2} \Delta_0 \ln \rho \quad (119)$$

where Δ stands for the Laplacian in curved space and Δ_0 in flat space relative to the metric $g_{ab}^{(0)} \equiv \delta_{ab}$. If in each coordinate patch one uses thermal coordinates, the conformal factor ρ drops out in the expression for the action, but not of course in the Laplace operator.

In the continuous case and for a compact surface, the set of eigenvalues is discrete, and an expansion similar to (110), but extending over an infinite number of modes is still valid. The partition function, $(\det' - \Delta)^{-\frac{1}{2}}$, requires subtractions. If we first ignore them, a variation of the free energy will take the form

$$\delta \ln Z = \delta \ln \left\{ \int \prod' \left(\frac{dC_n}{\sqrt{2\pi}} \right) \exp \left(-\frac{1}{2} \sum E_n C_n^2 \right) \right\} \quad (120)$$

The zero mode is given by equation (115).

We concentrate on a variation with respect to an infinitesimal conformal change of the form $g_{ab} \rightarrow (1 + 2\delta\epsilon)g_{ab}$, where $\delta\epsilon$ varies from point to point. This induces a change in the eigenvalues and eigenfunctions, and, for a fixed field $\varphi(\alpha)$, in the components C_n , but the combination $\sum E_n C_n^2$ is invariant, this being precisely the consequence of classical conformal invariance. Thus in equation (120) the variation arises solely from the Jacobian of the transformation from C_n to $C_n + \delta C_n$. Since the zero mode had to be subtracted, we find

$$\delta \ln Z = \left\langle \left(\sum_0^\infty \frac{\partial \delta C_n}{\partial C_n} - \frac{\partial \delta C_0}{\partial C_0} \right) \right\rangle \quad (121)$$

From the invariance of the field φ

$$\delta \varphi = \sum_0^\infty (\delta C_n \varphi^{(n)} + C_n \delta \varphi^{(n)}) = 0$$

Since

$$\int d^2\alpha \sqrt{g} \varphi^{(n_1)} \varphi^{(n_2)} = \delta_{n_1, n_2}$$

we have

$$\frac{\partial \delta C_n}{\partial C_n} = - \int d^2 \alpha \sqrt{g} \varphi^{(n)} \delta \varphi^{(n)} = \int d^2 \alpha \sqrt{g} \left(\frac{1}{4} \frac{\delta g}{g} \right) [\varphi^{(n)}]^2$$

and

$$0 = \int d^2 \alpha \sqrt{g} \varphi^{(n)} \delta \varphi^{(n)} + \int d^2 \alpha \sqrt{g} \left(\frac{1}{4} \frac{\delta g}{g} \right) [\varphi^{(n)}]^2$$

Hence formally

$$\delta \ln Z = \int d^2 \alpha \sqrt{g} \frac{1}{4} \frac{\delta g}{g} \sum_0^\infty [\varphi^{(n)}]^2 - \frac{1}{2} \frac{\delta A}{A} \quad (122)$$

The Gaussian average implied in equation (121) has dropped, but the quantity $\sum_0^\infty [\varphi^{(n)}]^2$ is infinite, as a consequence of the infinite number of modes, itself due to the continuum theory. This is an ultraviolet divergence and, instead of curing it by a ζ -function regularization, we shall employ an equivalent device, namely replace $\sum_0^\infty [\varphi^{(n)}(\alpha)]^2$ by the heat kernel at coinciding points $U(s; \alpha, \alpha)$

$$\sum_0^\infty [\varphi^{(n)}(\alpha)]^2 \rightarrow U(s; \alpha, \alpha) = \sum_0^\infty [\varphi^{(n)}(\alpha)]^2 \exp -s E_n \quad (123)$$

This is the trick used by Fujikawa in various similar circumstances. In general the symmetric heat kernel is given by

$$U(s; \alpha, \beta) = U(s; \beta, \alpha) = \sum_0^\infty \varphi^{(n)}(\alpha) \varphi^{(n)}(\beta) e^{-s E_n}$$

$$\left(\frac{\partial}{\partial s} - \Delta_\alpha \right) U(s; \alpha, \beta) = 0 \quad s > 0 \quad (124)$$

$$\lim_{s \rightarrow +0} U(s; \alpha, \beta) = \delta(\alpha, \beta)$$

The invariant δ -function satisfies

$$\varphi(\alpha) = \int d^2 \beta \sqrt{g(\beta)} \delta(\alpha, \beta) \varphi(\beta) \quad (125)$$

Thus we have

$$\delta \ln Z_{\text{reg}}(s) = \int d^2 \alpha \sqrt{g} \frac{1}{4} \frac{\delta g}{g} U(s; \alpha, \alpha) - \frac{1}{2} \frac{\delta A}{A} \quad (126)$$

and we ask for the small s behaviour. We expect to find terms of order s^{-1} plus finite ones in this two-dimensional case, in the

absence of boundaries (which would generate terms of order $s^{-\frac{1}{2}}$). In flat space

$$U_{\text{flat}}(s; \alpha, \beta) = \frac{\exp(-(\alpha - \beta)^2/4s)}{4\pi s} \quad (127)$$

Therefore $U_{\text{flat}}(s; \alpha, \alpha) = 1/(4\pi s)$ independently of α , which is recognized as a consequence of translational invariance. In curved space, if $R(\alpha)$ stands for the curvature, we expect on purely dimensional grounds

$$U_{\text{curved}}(s; \alpha, \alpha) = \frac{1}{4\pi s} [1 + \lambda s R(\alpha) + \dots]$$

since s has dimension $[\text{length}]^2$ and R $[\text{length}]^{-2}$, while λ must be a universal constant. It is therefore sufficient to know the result for any specific curved manifold, a sphere of radius r for instance, to deduce the value of λ .

For the metric (118) on the sphere in the vicinity of the point with coordinates $\alpha^1 = \alpha^2 = 0$, one readily derives that the heat kernel is of the form

$$U_{\text{sphere}}(s; \alpha, 0) = \frac{e^{-\alpha^2/4s}}{4\pi s} \left[1 + \frac{1}{3} R s \left(1 + \frac{\alpha^2}{4s} + \frac{(\alpha^2)^2}{8s^2} \right) + \dots \right] \quad (128)$$

Thus

$$U_{\text{curved}}(s; \alpha, \alpha) = \frac{1}{4\pi s} [1 + \frac{1}{3} R(\alpha) s + \dots] \quad (129)$$

With a metric locally of the form (117), it follows from equation (119) that

$$U_{\text{curved}}(s; \alpha, \alpha) = \frac{1}{4\pi s} - \frac{1}{12\pi} \Delta \ln \rho + \dots$$

$$= \frac{1}{4\pi s} - \frac{1}{12\pi} \frac{1}{\rho^2} \Delta_0 \ln \rho \quad (130)$$

With this choice of metric, we have $g = \rho^4$, hence $\frac{1}{4} \delta g/g = \delta \ln \rho$. Upon inserting these expressions into the regularized form (126), we obtain, neglecting terms which vanish with s ,

$$\delta \ln Z_{\text{reg}}(s) = \delta \left[\frac{A}{8\pi s} - \frac{1}{2} \ln \frac{A}{A_0} + \frac{1}{24\pi} \int d^2 \alpha \rho^2 \ln \rho (-\Delta \ln \rho) \right] \quad (131)$$

The short time expansion (129) for the heat kernel at coinciding points can be slightly generalized to near coincidence as follows

$$U_{\text{curved}}(s; \alpha, \beta) = \frac{\exp(-d_{\alpha\beta}^2/4s)}{4\pi s} \left[1 + \frac{1}{3} R_{\alpha} \left(s + \frac{1}{4} d_{\alpha\beta}^2 \right) + \dots \right] \quad (132)$$

with $d_{\alpha\beta}$ is the geodesic distance from α to β . The factor $(4\pi s)^{-1}$ in the flat case can be interpreted as the inverse of the area of a circle of radius $r_{\text{eff}} = 2\sqrt{s}$. This is in agreement with U being thought as a probability density for a Brownian motion. On a curved surface (curvature R), a circle with the same small radius has an area $\pi R^{-1} 2(1 - \cos \theta)$ where $\theta = R^{\frac{1}{2}} r_{\text{eff}} = 2R^{\frac{1}{2}} \sqrt{s}$, i.e. an area $4\pi s (1 - \frac{1}{3} R s + \dots)$. This provides an interpretation for the leading correction in equation (129). Can the reader generalize the argument (i) to include the second correction in equation (132) (ii) to a curved d -dimensional manifold.

The infinite term proportional to s^{-1} in (131) has to be renormalized to a finite one with an arbitrary coefficient of dimension $[\text{length}]^{-2}$. The safest way to make the derivation sensible is to assume that the scale variation was confined to a compact support within a coordinate patch in thermal coordinates.

The reader should be warned of the pitfalls in using the expression found above or rather, that it should be handled with due respect to boundary conditions. For instance, on a sphere with a point deleted, consider the system of coordinates with elementary arc length square given in equation (118). While it is true that the integral

$$\int_{\mathcal{R}^2} d^2 \alpha \rho^2 \ln \rho (-\Delta \ln \rho) \equiv \int_{\mathcal{R}^2} d^2 \alpha \ln \rho (-\Delta_0 \ln \rho)$$

is finite, the one obtained through an integration by parts

$$\int_{\mathcal{R}^2} d^2 \alpha \left[\left(\frac{\partial}{\partial \alpha^1} \ln \rho \right)^2 + \left(\frac{\partial}{\partial \alpha^2} \ln \rho \right)^2 \right]$$

is not. In general, since $-\Delta \ln \rho = R$, it is clear that if the integral of R does not vanish, i.e. if the Euler characteristics is nonzero, from

$$\chi = \frac{1}{2\pi} \int d^2 \alpha \rho^2 R \quad (133)$$

it follows that the quantity $\ln \rho$ cannot be defined everywhere as a nonsingular function. To avoid this difficulty, one could assume that the compact surface has the topology of a torus ($\chi = 0$) in which case it admits a flat metric.

The conclusion is that, while the classical massless field theory exhibits a local scale invariance, this is not true of the quantum theory where the renormalized free energy has a dependence on the local scale which assumes for a compact surface the following form

$$\ln \frac{Z}{Z_0} = -\frac{1}{2} \ln \frac{A}{A_0} + \frac{1}{24\pi} \int d^2 \alpha \rho^2 [\ln \rho (-\Delta \ln \rho) + \mu^2] \quad (134)$$

In this expression, A_0 is an arbitrary area to make $\ln Z$ dimensionless, and μ is an arbitrary mass scale since $\int d^2 \alpha \rho^2 = A$. As was emphasized, equation (134) makes better sense when comparing two choices of local scales the ratio being of compact support within a thermal coordinate patch. The second term in (134) is called the Liouville action, and the above result is due to Polyakov (1981). The scale independent term might also require a (logarithmic) renormalization, as a more careful analysis will show in the sequel.

Equation (134) implies a lack of conformal invariance of the renormalized free energy of the form

$$\delta \ln \left[\left(\frac{A}{A_0} \right)^{\frac{1}{2}} Z \right] = \frac{1}{12\pi} \int dA R \delta \psi \quad (135)$$

where $dA = d^2 \alpha \sqrt{g}$ is the invariant infinitesimal element of area, and the metric is rescaled as $g \rightarrow e^{2\delta \psi} g$.

An alternative form of the Liouville action is as follows. Let $G(\alpha, \beta)$ be the subtracted Green function on the surface

$$G(\alpha, \beta) = \sum_n \frac{1}{E_n} \varphi^{(n)}(\alpha) \varphi^{(n)}(\beta) \quad (136)$$

$$-\Delta G(\alpha, \beta) = \delta(\alpha, \beta) - \frac{1}{A}$$

Then

$$-\Delta \int dA_{\beta} G(\alpha, \beta) R_{\beta} = R_{\alpha} - \frac{2\pi}{A} \chi \quad (137)$$

Thus, if we assume for simplicity that the Euler characteristic χ vanishes,

$$\ln \left[\left(\frac{A}{A_0} \right)^{\frac{1}{2}} \frac{Z}{Z_0} \right] = \frac{1}{24\pi} \iint dA_{\alpha} R_{\alpha} G(\alpha, \beta) dA_{\beta} R_{\beta} + \frac{1}{24\pi} \int dA \mu^2 \quad (138)$$

This admits of a simple discretized form (again for $\chi = 0$). The second term is proportional to the area. Set

$$dAR \rightarrow q_i = \left(1 - \frac{\theta_i}{2\pi}\right) \quad (139)$$

such that $\sum_i q_i = 0$. Define a (massive) discrete symmetric Green's function such that

$$\sum_{k(i)} \frac{\sigma_{ik}}{\ell_{ik}} [G(i, j; m^2) - G(k, j; m^2)] + m^2 \sigma_i G(i, j; m^2) = \delta_{ij} \quad (140)$$

Then the first term on the r.h.s. in (138) assumes the form

$$S_{\text{Liouville}} = \frac{1}{24\pi} \lim_{m^2 \rightarrow 0} \sum_{i,j} q_i G(i, j; m^2) q_j \quad (141)$$

The zero mode appearing in the limit $m^2 \rightarrow 0$ is harmless since $\sum_i q_i = 0$. This nonlocal action is reminiscent of a Coulomb gas with the charges being replaced by the local curvature.

11.2.3 Sums over smooth surfaces

We present now Polyakov's continuous model of random surfaces which serves as a foundation for string field theory. The applications to a tentative unified theory of forces (including gravity) in particle physics will not be covered here, but motivates the formalism.

The surprising outcome of the analysis will be that dimension 26 is an upper bound in which the expressions make sense. In this dimension the model enjoys an additional quantum reparametrization invariance. This holds for the bosonic version discussed below, but turns into $d = 10$ for a generalized version including fermionic degrees of freedom. This extension might also be of interest in statistical physics, since there is a hope to interpret some three-dimensional systems, such as the Ising model, in terms of random fermionic surfaces using their dual gauge version.

The amount of analytical sophistication required to cope with the continuous theory is unfortunately rather heavy. The unfamiliar but interested reader is urged to refer to some of the reviews quoted in the notes, for instance those of Friedan and Alvarez which expanded on Polyakov's original work and provide the mathematical background.

In the continuum limit, rather than using the arc length, it was found that a natural action for a random path is proportional to

$$S_{\text{path}} = \int d\tau \left(\frac{dx(\tau)}{d\tau} \right)^2 \quad (142)$$

with τ a parameter along the curve, the proper time in relativistic quantum mechanics, or length according to the metric induced from ambient space. Although this is not invariant under reparametrization, one could generalize this expression by introducing a monotonic function $\tau = f(s)$ and formulate the path integral in such a way that, after gauge fixing, the action recovers the form (142).

For smooth surfaces embedded in Euclidean space $\mathcal{R}^d$, one wishes similarly to replace the Nambu-Goto action (proportional to the area computed using the induced metric), by a suitable form analogous to (142). Rather than choosing a specific metric, Polyakov suggested, by analogy with quantum gravity, to average over all possible metrics defining the (square) infinitesimal arc length

$$ds^2 = g_{ab}(\alpha) d\alpha^a d\alpha^b \quad (143)$$

Thus the path integral takes the form

$$Z = \sum_{\text{topologies}} \lambda_0^x \int_{\text{metrics}} \mathcal{D}g \int_{\text{embeddings}} \mathcal{D}(X) \exp - \left[S(g, X) + \mu_0^2 \int d^2\alpha \sqrt{g} \right] \quad (144)$$

Here λ_0 is a bare activity conjugate to the characteristic χ given by equation (133). The term proportional to the area $A = \int d^2\alpha \sqrt{g}$, computed with the metric g , leaves us the option of absorbing an infinite similar counterterm in a suitable renormalization.

The vacuum functional in (144) involves only compact orientable surfaces (the latter for simplicity) and we shall also assume a restriction to connected surfaces. Otherwise one should presumably include combinatorial factors. The d fields $X^\mu(\alpha)$ map the surface in $\mathcal{R}^d$, and the main part of the action reads

$$S(g, X) = \frac{1}{2} \int d^2\alpha \sqrt{g} g^{ab}(\alpha) \sum_\mu \frac{\partial}{\partial \alpha^a} X^\mu \frac{\partial}{\partial \alpha^b} X^\mu \quad (145)$$

The target space could be replaced by a d -dimensional manifold with metric $G_{\mu\nu}(X)$ as indicated in equation (95). We refrain from doing so.

The quadratic action (145) is similar to the one-dimensional one with the appropriate geometric factor, $g = \det g_{ab}$, and g^{ab} the inverse of g_{ab} . It enjoys important invariance properties. These are of three types.

- i) Reparametrization invariance, accompanied of course by the corresponding changes in the metric.
- ii) Rescaling of the metric $g_{ab} \rightarrow \rho^2 g_{ab}$, under which the combination $\sqrt{g} g^{ab}$ is invariant. This is also called Weyl invariance and affects lengths and areas but not angles. Metrics which differ by a scale factor are therefore said to belong to the same conformal class. It should be stressed that if we only look at the action on metrics, the effect of reparametrization can partly be identified with a rescaling and this will play a role in the sequel.
- iii) The action is also invariant when X is transformed by an element of the Euclidean group. Only the noncompact translation part is a dangerous one and will be avoided by separating the center of mass motion or constant mode.

Constrained by the requirement of involving only X through its gradient, the action (145) supplemented by the terms conjugate to $\ln \lambda_0$ and μ_0^2 contains all possible renormalizable combinations compatible with these symmetries, given that a physical length scale has been absorbed in X .

Our task is to make sense of the formal expression (144), more precisely of each piece individually. Therefore in the following, we deal with a fixed topology and wish to extract a factor corresponding to the infinite reparametrization group (more pedantically the diffeomorphism group). In other words, we want to fix the gauge, keeping the benefits of gauge invariance. It will be seen that at the quantum level with regard to scale invariance (assuming the renormalized μ^2 to vanish), this can only be achieved in dimension 26 as announced.

When μ_0^2 vanishes, the classical equations of motion obtained by varying in turn g_{ab} and X^μ are

$$T_{ab} = - \sum_{\mu} \partial_a X^\mu \partial_b X^\mu + \frac{1}{2} g_{ab} \sum_{\mu} \partial^c X^\mu \partial_c X^\mu = 0 \quad (146)$$

$$\Delta X^\mu \equiv \frac{1}{\sqrt{g}} \partial_a \sqrt{g} \partial^a X^\mu = 0 \quad (147)$$

We recognize in (146) the vanishing of the symmetric traceless energy momentum tensor which requires g_{ab} to be proportional to the metric $\sum_{\mu} \partial_a X^\mu \partial_b X^\mu$ induced from the ambient Euclidean space on the surface. The field coordinates X^μ have then to be harmonic for this metric. This is the condition for minimal surfaces (i.e. surfaces which minimize the area locally). Reinserted in the action, these expressions yield the induced total area. Clearly this would make more sense for surfaces with boundaries.

To begin with, we recall a property stated in equation (117), namely that any metric can be taken locally in a diagonal form by an appropriate choice of coordinates. One can even further insure that at some point the conformal factor is unity and that its gradient vanishes. An additional stronger property is that each conformal class of metrics contains a metric g_0 of constant curvature. This says that given g , there exists a choice of ψ such that $g = e^{2\psi} g_0$ with R_0 constant, i.e. from equation (119) this requires solving for ψ the classical Liouville equation

$$-\Delta \psi = R - e^{-2\psi} R_0 \quad (148)$$

The reader will recall that the uniformization of Riemann surfaces implies that for $\chi = 0$ a torus may be obtained by taking the quotient of the complex plane by a lattice of translations, while for $\chi < 0$ it may be obtained as the quotient of the upper half complex plane by an appropriate discrete crystallographic subgroup of $SL(2R)$. A reference metric is $dz d\bar{z}$, or $dz d\bar{z}/(\text{Im} z)^2$ respectively, in appropriate coordinates. The case of the sphere has been exemplified above. To see what is involved, use complex coordinates $z = \alpha^1 + i\alpha^2$, $\bar{z} = \alpha^1 - i\alpha^2$. A general metric takes the form $ds^2 = g_{zz}^2 dz^2 + g_{\bar{z}\bar{z}}^2 d\bar{z}^2 + 2g_{z\bar{z}} dz d\bar{z}$, ($g_{\bar{z}\bar{z}} = \overline{g_{zz}}$, $g_{z\bar{z}} = \overline{g_{\bar{z}z}}$) with the positivity condition being $g_{z\bar{z}}^2 > g_{zz} g_{\bar{z}\bar{z}}$, $g_{z\bar{z}} > 0$. Set

$$\begin{aligned} \frac{\mu}{\bar{\mu}} &= \frac{g_{\bar{z}\bar{z}}}{g_{zz}} & \frac{1}{2}(1 + \mu\bar{\mu}) &= \left| \frac{g_{z\bar{z}}}{g_{zz}} \right|^2 \left(1 - \sqrt{1 - \left| \frac{g_{z\bar{z}}}{g_{zz}} \right|^2} \right) \\ \lambda &= \frac{2g_{z\bar{z}}}{1 + \mu\bar{\mu}} \end{aligned} \quad (149a)$$

Positivity means $\lambda > 0$ and $|\mu| < 1$, while the metric reads

$$ds^2 = \lambda |dz + \mu d\bar{z}|^2 \quad (149b)$$

The above property amounts to saying (with due attention paid to the crystallographic invariance for $\chi < 0$) that it is possible to find a smooth change of variable $u = u(z, \bar{z})$, $\bar{u} = \bar{u}(z, \bar{z})$, such that in the variables $u, \bar{u}$, the metric assumes the (anti-) diagonal form $ds^2 \propto du d\bar{u}$, i.e. to solve the Beltrami equation $\partial u / \partial \bar{z} = \mu \partial u / \partial z$.

We return to a specific term in the sum (144) with the aim to "fix the gauge" by taking g conformally equivalent to a given metric g_0 of constant curvature (i.e. $g = e^{2\psi} g_0$ with ψ varying). To do so, we study variations δg around g . One defines a square norm on deformations (with an obvious generalization to tensor fields)

$$\|\delta g\|^2 = \int d^2\alpha \sqrt{g} g^{ac} g^{bd} \delta g_{ab} \delta g_{cd} \quad (149)$$

This enables one to obtain a volume element $\mathcal{D}g$ by analogy with Riemann manifolds in a standard fashion. This definition implies that the metric g on the surface. The variation δg can be split into two orthogonal parts according to (150), a symmetric traceless part δh and a piece proportional to g itself (the trace part) as

$$\delta g = 2\delta\varphi g + \delta h \quad (151)$$

Since these pieces are orthogonal, $\mathcal{D}g$ can be factored into

$$\mathcal{D}g = \mathcal{D}\varphi \mathcal{D}h \quad (152)$$

both factors being reparametrization invariant. One has in mind that this corresponds to a factorization into an integration over the conformal factor and reparametrization both acting on the reference metric g_0 . To see this, we examine how an infinitesimal reparametrization $\alpha \rightarrow \alpha + \delta\alpha$ affects infinitesimally the metric

$$(g_{ab} + \delta g_{ab}) d\alpha^a d\alpha^b = g(\alpha + \delta\alpha)_{ab} d(\alpha^a + \delta\alpha^a) d(\alpha^b + \delta\alpha^b) \quad (153)$$

i.e.

$$\delta_v g_{ab} = \delta\alpha^c \frac{\partial}{\partial\alpha^c} g_{ab} + g_{ac} \frac{\partial\delta\alpha^c}{\partial\alpha^b} + g_{cb} \frac{\partial\delta\alpha^c}{\partial\alpha^a} = D_b v_a + D_a v_b \quad (154)$$

where we have set

$$v_a = g_{ab} \delta\alpha^b \quad (154)$$

for the infinitesimal (covariant) vector field associated to the reparametrization, and D is the covariant derivative involving

the Riemann-Christoffel symbol

$$D_a v_b = \partial_a v_b - \Gamma_{ab}^c v_c \quad (155)$$

$$\Gamma_{ab}^c = \frac{1}{2} g^{cd} \left\{ \frac{\partial g_{ad}}{\partial\alpha^b} + \frac{\partial g_{bd}}{\partial\alpha^a} - \frac{\partial g_{ab}}{\partial\alpha^d} \right\}$$

This abuse of tensor indices will readily be remedied by the judicious choice of reference metric. We also have under rescaling

$$\delta_\psi g_{ab} = 2\delta\psi g_{ab} \quad (156)$$

The sum $(\delta_v + \delta_\psi)g$ can now be compared to the general orthogonal decomposition (151), observing that we have to subtract from $\delta_v g$ its trace. Thus

$$(\delta_v + \delta_\psi)g = 2\delta\varphi g + \delta h \quad (157)$$

$$\delta\varphi = \delta\psi + \frac{1}{2} g^{ab} D_a v_b \quad \delta h = D_a v_b + D_b v_a - g_{ab} (g^{cd} D_c v_d)$$

The standard notation is

$$\delta h = P_1 v \quad (158)$$

where P_1 is a first order differential operator which maps vectors (hence the index 1) onto symmetric traceless 2-tensors. We now trade the variables φ and h for ψ (associated to a conformal factor) and v which represents reparametrization, taking care of a Jacobian functional determinant (the Faddeev-Popov determinant)

$$\mathcal{D}\varphi \mathcal{D}h = \mathcal{D}\psi \mathcal{D}v \left| \frac{D(\varphi, h)}{D(\psi, v)} \right| \quad (159a)$$

with a symbolic presentation

$$\frac{D(\varphi, h)}{D(\psi, v)} = \det \begin{vmatrix} I & X \\ 0 & P_1 \end{vmatrix} = \det P_1$$

As we are only interested in the absolute value, we insert in (159) instead of $\det P_1$ the positive quantity $(\det P_1^\dagger P_1)^{\frac{1}{2}}$ where the operator $P_1^\dagger P_1$ has the advantage that it maps vectors fields on vector fields. Thus

$$\mathcal{D}\varphi \mathcal{D}h = \mathcal{D}\psi \mathcal{D}v (\det P_1^\dagger P_1)^{\frac{1}{2}} \quad (159b)$$

One could represent the determinant as a Grassmannian integral over "ghost" fields. This familiar device is useful for further developments in string field theory, especially to give a compact

form to the expansion (144) interpreted as a perturbative series of a would-be second quantized interacting theory. The analogy is with the Feynman expansion expressed in terms of one-dimensional graphs. The Grassmannian integral, similar to the one over the X -fields, assumes a simple form using complex coordinates introduced below, emphasizing its conformal invariance. We leave as an exercise for the reader to establish its expression.

To define the measure Dv , we use a norm on vector fields, which has the merit of being independent from the conformal factor ψ , setting by analogy with equation (150)

$$\|v\|^2 = \int d^2\alpha \sqrt{g} g^{ab} v_a v_b = \int d^2\alpha \sqrt{g_0} g_0^{ab} v_a v_b$$

The (infinite and formal) integral $\int Dv$ may be interpreted as the volume of the reparametrization group.

We have to pay attention also to the fact that the operator P_1 may have zero modes i.e. may annihilate certain vector fields (called conformal Killing fields) corresponding to reparametrizations which just rescale the metric. Those are already included in the integral over ψ . To avoid double counting, and to define a positive Jacobian $(\det' P_1^\dagger P_1)^{\frac{1}{2}}$, we must integrate not over the full space of v 's but on the orthogonal complement of the (finite-dimensional) space of zero modes. This has the drawback that the corresponding integral yields only part of the "volume" of the reparametrization group. Finally, it was implicitly assumed that the most general δg can be obtained in the form $(\delta_\psi + \delta_v)g$. This need not always be the case and must be investigated as a separate issue. The trace part does not present a problem. The question is whether every traceless δh belongs to the image of P_1 . This is in general not the case and leaves (for $\chi \leq 0$) a finite-dimensional space, with a measure denoted Dh_M , which is called the Teichmüller or moduli space, with an intricate global structure. To be precise, the Teichmüller space is a covering of the space of moduli. So we have more precisely

$$Dg = Dh_M Dv^\perp D\psi (\det' P_1^\dagger P_1)^{\frac{1}{2}} \quad (159c)$$

where the prime refers to the determinant over nonzero modes.

Let the series in (144) be written as a sum of contributions with different topologies

$$Z = \sum_x Z_x \quad (160)$$

An individual term can be written

$$Z_x = \lambda_0^x \int Dv^\perp \int Dh_M D\psi \int \prod_\mu dC_0^\mu e^{-\mu_0^2 \int d^2\alpha \sqrt{g}} (\det' P_1^\dagger P_1)^{\frac{1}{2}} (\det' [-\Delta])^{-\frac{1}{2}d} \quad (161)$$

where we have performed the integral over the X -fields, separating the integral over constant normalized zero modes C_0^μ (hence the prime over the determinant of the Laplacian, computed of course with the metric g). Since all the quantities involved are reparametrization invariant, the integral $\int Dv^\perp$ factors out and will simply be omitted in a renormalized version of Z_x which is still a formal one, since the determinants are of course ill-defined and also require renormalization. We follow the heuristic physicist's tradition of proceeding from nonsensical infinite quantities towards renormalized ones, rather than an axiomatic presentation of the final product which would hide the motivation. The reader might also worry (and we worry with him) that different topologies have different integrals over the reparametrization group, so that at this stage the way to combine the various Z_x 's into a unique sum remains problematic.

Our main concern is with the definition of the determinants and with the integral over moduli. We expect from section 2.2 to see the Liouville action emerging again from the ψ dependence. This leads us into a technical detour of intrinsic interest, which elaborates on the schematic discussion of the massless field given above.

To investigate the determinants occurring in equation (161), we use complex coordinates $z = \alpha^1 + i\alpha^2$, $\bar{z} = \alpha^1 - i\alpha^2$. The metric reads $g = e^{2\psi} g_0$, with a reference metric g_0 such that locally $ds^2 = e^{2\psi_0} dz d\bar{z}$, corresponding to a constant curvature as agreed above. Thus only $g_{z\bar{z}}$ is nonzero and positive while $g^{z\bar{z}} = g_{z\bar{z}}^{-1}$.

We study analytic transformations acting on tensors with contravariant and covariant indices (the keyword is analytic). These transformations leave the metric structure invariant. The metric g allows us to raise or lower indices (its transformation law is of course

$g_{zz} dzd\bar{z} = g'_{zz} dz'd\bar{z}'$). Thus we can restrict ourselves to covariant indices say, in which case (as in chapter 9) the transformation law is specified by requiring the combination $t_{h,\bar{h}} dz^h d\bar{z}^{\bar{h}}$ to be invariant. The suffixes h and $\bar{h}$ (both integers here) are the numbers of z and $\bar{z}$ lower indices. The simplification brought about by analytic transformations, to which we limit ourselves henceforth, is that a component $t_{h,\bar{h}}$ is simply multiplied by a factor. It is slightly more convenient in the present context to trade the $\bar{h}$ lower $\bar{z}$ indices for $\bar{h}$ upper z indices. We denote the tensor $t^{\bar{h}}_h = (g^{z\bar{z}})^{\bar{h}} t_{h,\bar{h}}$, so that the transformation law reads

$$(t^{\bar{h}}_h)' = \left(\frac{dz'}{dz}\right)^{\bar{h}-h} t^{\bar{h}}_h \quad (162)$$

To conform to usage, we shall adopt the opposite convention from chapter 9 by using n (and not s) for the difference $\bar{h} - h$ ($s = -n$). For fixed n (positive or negative) the set of (one-component) tensorial quantities which transform as indicated will be called T^n .

The operators we consider such as $-\Delta$ or P_1 map symmetric traceless tensors onto symmetric traceless tensors. The statement is empty for the scalar Laplacian and for P_1 , it maps vectors on symmetric traceless 2-tensors. Such real tensors have only two independent components in two dimensions.

To see this in Euclidean geometry, associate to such a tensor the homogeneous polynomial $P_n(x, y) = \sum_{k+k'=n} T_{11\dots 22\dots} x^k y^{k'}$ with *a priori* $n+1$ coefficients. The tensor is traceless if the ordinary Laplacian $\partial^2/\partial x^2 + \partial^2/\partial y^2$ acting on P_n gives zero. This provides us with a homogeneous polynomial of degree $n-2$, hence with $n-1$ conditions (or zero if $n=1$), thus leaving two independent coefficients such that $P_n(x, y) = t_+(x-iy)^n + t_-(x+iy)^n$. If the tensor is real the components are complex conjugate. The argument goes through in curved space. In complex coordinates, one component belongs to T^n , call it t , while the second component $\bar{t}$ would have n $\bar{z}$ upper indices. To conform to the above convention, it is multiplied by $(g_{z\bar{z}})^n$ to obtain an element of T^{-n} . By complexification, the set of (real) symmetric traceless tensors of rank n can therefore be considered as belonging to the direct sum $T^n \oplus T^{-n}$, with an association

$$T \equiv \{t, (g_{z\bar{z}})^n \bar{t}\} \quad (163)$$

where it is understood that t is in T^n . The corresponding space of associated real symmetric traceless tensors is denoted S^n . The simpler case $n=0$ with only one component is set aside. Using the special form of the metric, the covariant derivative has two

components D^z and $D_{\bar{z}}$ (with an appended suffix n if necessary to recall on which space it operates, as D^n_z and $D^n_{\bar{z}}$)

$$\begin{aligned} T^n \xrightarrow{D^n_z} T^{n+1} & \quad t \rightarrow g^{z\bar{z}} \bar{\partial} t \\ T^n \xrightarrow{D^n_{\bar{z}}} T^{n-1} & \quad t \rightarrow (g^{z\bar{z}})^n \partial_z (g_{z\bar{z}})^n t \end{aligned} \quad (164)$$

and an invariant scalar product is derived from the norm

$$\|t\|^2 = \int d^2\alpha \sqrt{g} (g_{z\bar{z}})^n \bar{t} t \quad (165a)$$

One can easily verify these formulas and check the transformation laws.

In essence D^z is a decoration of the $\bar{\partial}$ operator and the factors in $D_{\bar{z}}$ compensate for the passage from D^z to $D_{\bar{z}}$ when acting on the appropriate space. A similar remark applies to the norm which we extend (with a reality condition) to S_n by requiring for T written as in equation (163) that

$$\langle T_1, T_2 \rangle = \int d^2\alpha \sqrt{g} (g_{z\bar{z}})^n (\bar{t}_1 t_2 + t_1 \bar{t}_2) \quad (165b)$$

Using these notations, the adjoint of D^z is acting from T^{n+1} to T^n and

$$(D^n_z)^\dagger = -D^{n+1}_{\bar{z}} \quad (166)$$

There are now two possible definitions of the Laplacian acting on T^n . Either we can climb one rung on the ladder then retreat, or do the opposite. In formulas

$$\begin{aligned} T_n \rightarrow T_n & \quad -\delta_n^{(+)} = -2D^{n+1}_{\bar{z}} D^n_z \equiv 2(D^n_z)^\dagger D^n_z \\ & \quad -\delta_n^{(-)} = -2D^n_{\bar{z}-1} D^n_z \equiv 2D^n_{\bar{z}-1} (D^n_{\bar{z}-1})^\dagger \end{aligned} \quad (167)$$

The factor 2 is included to agree with the flat space definition (surely had Laplace lived long enough, he would have included a minus sign into the definition of his operator to make it positive!). These operators are semidefinite positive.

On the space S^n , operators are acting on the two conjugate components

$$\begin{aligned} P_n &= (D^n_z, D^n_{\bar{z}-n}) \\ S^n &\equiv S^{n+1} & P_n^\dagger &= (-D^{n+1}_{\bar{z}}, -D^{n+1}_{\bar{z}-(n+1)}) \end{aligned} \quad (168)$$

and P_1 and $P_1^\dagger$ are those occurring in formula (161). The associated Laplacians on S are

$$S_n \rightarrow S_n \quad -\Delta_n^{(+)} = 2P_n^\dagger P_n = (-\delta_n^{(+)}, -\delta_{-n}^{(-)})$$

$$S_{n+1} \rightarrow S_{n+1} \quad -\Delta_{n+1}^{(-)} = 2P_n P_n^\dagger = \left(-\delta_{n+1}^{(-)}, -\delta_{-(n+1)}^{(+)} \right) \quad (169)$$

The relation between the spectra of $P_n^\dagger P_n$ and $P_n P_n^\dagger$ (in particular for $n = 1$) is crucial for the following. Both operators are non-negative but may have zero modes which span their kernel. For an operator O , the kernel of $O^\dagger$ is orthogonal to the image of O , and the same relation holds with O and $O^\dagger$ interchanged. Thus in terms of orthogonal complements

$$\left(\ker -\Delta_n^{(+)} \right)^\perp \leftrightarrow \left[\ker -\Delta_{n+1}^{(-)} \right]^\perp \quad (170)$$

The left-hand side is also the complement of $\ker P_n$ the right-hand side of $\ker P_n^\dagger$. To make this correspondence precise, let $\{T_r\}$ be an orthonormal set of nonzero eigenmodes of $-\Delta_n^{(+)}$ with corresponding eigenvalues λ_r (the scalar product being understood in the sense of equations (165)), then

$$2P_n^\dagger P_n T_r = \lambda_r T_r \quad \lambda_r > 0 \quad (171)$$

Thus $2P_n P_n^\dagger (P_n T_r) = \lambda_r (P_n T_r)$ and $P_n T_r$ is an eigenvector of $-\Delta_{n+1}^{(-)}$ with the same eigenvalue λ_r . These modes are orthogonal. To normalize them, we note that

$$\langle P_n T_r, P_n T_{r'} \rangle = \langle T_r, P_n^\dagger P_n T_{r'} \rangle = \frac{1}{2} \lambda_r \delta_{r,r'}$$

Therefore an orthonormal basis for $\left[\ker \left(-\Delta_{n+1}^{(-)} \right) \right]^\perp$ is

$$\left\{ \tilde{T}_r = \sqrt{\frac{2}{\lambda_r}} P_n T_r \right\}$$

justifying equation (170).

In the particular case $n = 1$, the zero modes of P_1 are the conformal Killing vectors, while the harmonic traceless 2-tensors, zero modes of $P_1^\dagger$, span those deformations of the metric which can not be described by rescaling or changes of coordinates and are associated to the moduli, which left a finite-dimensional measure in equations (159c) and (161). In general the zero modes are harmonic tensors in their respective spaces. We turn to the determinants themselves and use as before the heat kernel method with

$$\ln \det' \left(-\Delta_n^{(+)} \right) = - \int_\epsilon^\infty \frac{dt}{t} \text{Tr}' \exp t \Delta_n^{(+)} \quad (172)$$

since

$$- \int_\epsilon^\infty \frac{dt}{t} e^{-t\lambda} = \ln \epsilon \lambda + \gamma + o(\epsilon) \quad (173)$$

This is a (brute force) regularization for $\sum_r' \ln \lambda_r$, the sum running over positive eigenvalues, and ϵ is an ultraviolet cutoff, the square of an elementary length scale. One should recall that $-\Delta_n^{(+)}$ operates in the direct sum of two spaces (for $n \neq 0$). The method to extract information from (172) is, as in section 2.2, to study the short time behaviour of the appropriate kernel.

The calculation parallels the one already presented. We shall nevertheless give some details pertaining to the modified Laplacians occurring here, using perturbation theory. Let α denote a point at which we want to estimate the diagonal heat kernel $G(\alpha, \alpha; t) = \langle \alpha | e^{t \Delta_n^{(+)}} | \alpha \rangle$. This is really a two-by-two diagonal matrix. We recall that G behaves as t^{-1} for small t , and that length scales as $t^{1/2}$. We look for terms up to order t^0 . Thus

$$\langle \alpha | e^{t \Delta_n^{(+)}} | \alpha \rangle = \frac{G_{-1}}{t} + G_0 + \dots \quad (174a)$$

where $G_{-1}, G_0, \dots$ are two-by-two diagonal matrices which depend of course on the point under consideration. Inserting this expansion in (173), we have

$$\begin{aligned} \ln \det' \left(-\Delta_n^{(+)} \right) &= -\frac{1}{\epsilon} \int d^2 \alpha \sqrt{g} \text{Tr} G_{-1} \\ &+ \ln \epsilon \left[\int d^2 \alpha \sqrt{g} \text{Tr} G_0 - \dim \text{Ker} \left(-\Delta_n^{(+)} \right) \right] + \text{finite terms} \end{aligned} \quad (174b)$$

The second term in $\ln \epsilon$ arises from the original prescription of summing over nonzero modes.

Suppose that we vary the scale factor ψ in $g = e^{2\psi} g_0$. Suppressing all indices, for a nonzero eigenmode in S^n

$$\lambda \langle T, T \rangle = 2 \langle PT | PT \rangle$$

reads explicitly

$$\lambda \int d^2 \alpha \sqrt{g} (g_{z\bar{z}})^n \bar{t} t = 2 \int d^2 z \sqrt{g} (g_{z\bar{z}})^{n+1} (g^{z\bar{z}} \bar{\partial} t) (g^{z\bar{z}} \partial \bar{t})$$

Hence from the normalization, we have

$$\delta \lambda + \lambda 2(n+1) \langle T, \delta \psi T \rangle = 4n \langle PT, \delta \psi PT \rangle$$

Recalling that $\sqrt{\frac{2}{\lambda}} T = \tilde{T}$ is the corresponding normalized eigenmode in S^{n+1}

$$\delta \lambda = 2\lambda \left\{ n \langle \tilde{T}, \delta \psi \tilde{T} \rangle - (n+1) \langle T, \delta \psi T \rangle \right\}$$

and

$$\begin{aligned}\delta \ln \det' (-\Delta_n^{(+)}) &= \int_{\epsilon}^{\infty} dt \sum_r' \delta \lambda_r e^{-t\lambda_r} \\ &= 2 \int_{\epsilon}^{\infty} dt \frac{d}{dt} \sum_r' \left[(n+1) \langle T_r, \delta \psi T_r \rangle - n \langle \tilde{T}_r, \delta \psi \tilde{T}_r \rangle \right] e^{-t\lambda_r} \\ &= 2n \text{Tr}' \delta \psi e^{\epsilon \Delta_{n+1}^{(-)}} - 2(n+1) \text{Tr}' \delta \psi e^{\epsilon \Delta_n^{(+)}}\end{aligned}\quad (175)$$

To evaluate this quantity we need the analog of equation (174a) for the operator $\Delta_{n+1}^{(-)}$. Explicitly, writing in a column the two diagonal elements of the operator

$$-\Delta_n^{(+)} = 2(g_{z\bar{z}})^{-1} \begin{bmatrix} [\partial\bar{\partial} + n(\partial \ln g_{z\bar{z}})\bar{\partial}] \\ [\partial\bar{\partial} - n(\partial \ln g_{z\bar{z}})\bar{\partial} - n(\partial\bar{\partial} \ln g_{z\bar{z}})] \end{bmatrix}$$

An alternative derivation of a result similar to equation (130) uses a perturbative method by setting $g_{z\bar{z}} = \frac{1}{2}e^{2\psi}$ with a coordinate system adapted to the point α at which we evaluate the diagonal element, this point being chosen at the origin with $\psi(0) = 0$, $\partial\psi(0) = \bar{\partial}\psi(0) = 0$. With reference to the traditional (Euclidean) notation of quantum mechanics, we write P^2 for (minus) the flat Laplacian $P^2 = -4\partial\bar{\partial}$.

$$-\Delta_n^{(+)} = \begin{pmatrix} P^2 + (e^{-2\psi} - 1)P^2 - 8n\partial\psi e^{-2\psi}\bar{\partial} \\ P^2 + (e^{-2\psi} - 1)P^2 + 8n\partial\psi e^{-2\psi}\bar{\partial} + 8ne^{-2\psi}(\partial\bar{\partial}\psi) \end{pmatrix}$$

$$\psi(z, \bar{z}) = \frac{1}{2}z^2\partial^2\psi(0) + z\bar{z}\partial\bar{\partial}\psi(0) + \frac{1}{2}\bar{z}^2\bar{\partial}^2\psi(0) + \dots$$

Since $|z|$ is of order $\sqrt{t}$, we need only to keep terms up to second order and perturbation theory up to first order will suffice in the form

$$e^{-P^2 t - tV} = e^{-P^2 t} - \int_0^t dt' e^{-P^2(t-t')} V e^{-t'P^2} + \dots$$

For the upper and lower components respectively, we have

$$\begin{aligned}V_{up} &= V_0 - V_1 \\ V_{down} &= V_0 + V_1 + V_2 \\ V_0 &= -[z^2\partial^2\psi(0) + 2z\bar{z}\partial\bar{\partial}\psi(0) + \bar{z}^2\bar{\partial}^2\psi(0)] \\ V_2 &= 8n\partial\bar{\partial}\psi(0)\end{aligned}$$

We are only interested in $\text{Tr} \left(0 \left| e^{t\Delta_n^{(+)}} \right| 0 \right)$, where the trace refers to the sum of contributions of upper and lower components, so the contribution of V_1 cancels. The contribution of V_0 has already been obtained in equation (130). The reader might enjoy doing the

computation in the present formalism. It gets doubled here due to the trace, as does the zeroth order term. Finally $V_2 = 2n\Delta\psi(0)$ contributes to the trace with a factor

$$- \int_0^t dt' \int \frac{d^2p}{(2\pi)^2} e^{-p^2 t'} = -\frac{1}{4\pi}$$

Putting everything together

$$\text{Tr} \left(\alpha \left| e^{t\Delta_n^{(+)}} \right| \alpha \right) = \frac{1}{2\pi t} - \frac{1}{6\pi} \Delta\psi(\alpha) - \frac{n}{2\pi} \Delta\psi(\alpha) + o(t)$$

where we have successively distinguished the various contributions. Thus with f diagonal in configuration space, and recognizing that the curvature is given by equation (119), $R = -\Delta\psi$,

$$\text{Tr} f e^{\epsilon \Delta_n^{(+)}} = \int d^2\alpha \sqrt{g} f \left\{ \frac{1}{2\pi\epsilon} + \frac{(3n+1)}{6\pi} R \right\} + o(\epsilon) \quad (176a)$$

In a similar vein

$$\text{Tr} f e^{\epsilon \Delta_{n+1}^{(-)}} = \int d\alpha \sqrt{g} f \left\{ \frac{1}{2\pi\epsilon} - \frac{(3n+2)}{6\pi} R \right\} + o(\epsilon) \quad (176b)$$

Finally for $n = 0$, if Δ stands for the scalar Laplacian, we found before that

$$\text{Tr} f e^{\epsilon \Delta} = \int d\alpha \sqrt{g} f \left\{ \frac{1}{4\pi\epsilon} + \frac{1}{12\pi} R \right\} + o(\epsilon) \quad (176c)$$

An immediate consequence of these computations is the following beautiful observation leading to a special case of the (Atiyah-Singer) index theorem. The index is defined as

$$I_n = \dim[\ker P_n] - \dim[\ker P_n^\dagger] \quad (177)$$

Since $-\Delta_n^{(+)} = 2P_n^\dagger P_n$ and $-\Delta_{n+1}^{(-)} = 2P_n P_n^\dagger$ have an identical nonzero spectrum, for any t ,

$$I_n = \text{Tr} e^{t\Delta_n^{(+)}} - \text{Tr} e^{t\Delta_{n+1}^{(-)}}$$

Being t -independent, this quantity can be computed in the limit $t \rightarrow 0$ with the help of equations (176). Setting $f \equiv 1$, the t^{-1} terms cancel, and from the finite terms emerges the relation with topology through the Euler characteristic

$$I_n = \frac{(2n+1)}{2\pi} \int d^2\alpha \sqrt{g} R = (2n+1)\chi \quad (178)$$

giving the difference between the number of linearly independent real harmonic symmetric tensors of rank n and $n+1$ and proving that I_n is a topological invariant.

For $n = 1$, the index is $I_1 = 3\chi$. The dimensions involved in (177) are on real numbers. The number of linearly independent conformal Killing vectors ($\dim[\text{Ker } P_1]$) is the (real) dimension of the group of analytic diffeomorphism called the global conformal group in chapter 9. Thus $\dim[\text{Ker } P_1] = 6, 2, 0$ if $\chi = 2, 0$ or negative, corresponding to the groups $SL(2, \mathbb{C})$ for the sphere or complex translations in the case of the torus. Compact Riemann surfaces of genus larger than one admit at most a finite group of automorphisms. Thus $\dim[\text{Ker } P_1] = \dim(\text{Ker } P_1) - 3\chi$ is zero for genus zero, two for genus one and $6(g-1)$ for genus larger than one, all these on real numbers. But this is precisely the dimension of the space of those traceless tensors h_{ab} which are not obtained by rescaling or reparametrizations, in other words the (real) dimension of the moduli space. Of course this result is familiar for $g = 0$ and 1.

We have now all the tools to extract from the determinants all divergent terms as well as the finite ones that depend on the conformal scale ψ . From equations (174) and (176a)

$$\begin{aligned} & \ln \det' (-\Delta_n^{(+)}) \\ &= -\frac{1}{2\pi\epsilon} \int d^2\alpha \sqrt{g} + \ln \epsilon \left[(n + \frac{1}{3}) \chi - \dim[\text{Ker } P_n] \right] + \text{finite} \end{aligned} \quad (179)$$

This is dimensionally correct since ϵ has dimension $[\text{length}]^2$ and could be written in terms of an ultraviolet cutoff $\epsilon = \Lambda^{-2}$. The coefficient of the $\ln \epsilon$ term is of course a pure number. This is still insufficient, since we are also interested in the finite part which contains the anomaly. Up to the logarithmic term multiplied by $\dim(\text{Ker } P_n)$ (which shall soon be disposed of) the remaining divergences can be taken care by a renormalization of the potential λ_0 and the bare "tension" μ_0^2 in equation (161).

The variational formula (175) involves primed traces. If we use instead full traces as computed in (176ab), we have to subtract the contributions from the zero mode subspaces. Let $\{T_s^{(0)}\}$ be a basis of the kernel of P_n . Consider the finite symmetric matrix

$$(H_n)_{ss'} = \langle T_s^{(0)}, T_{s'}^{(0)} \rangle = \int d^2\alpha \sqrt{g} (g_{zz})^n \left\{ t_s^{(0)} \bar{t}_{s'}^{(0)} + \text{c.c.} \right\} \quad (180)$$

Recall that $t_s^{(0)}$ are analytic in z and can be chosen as ψ -independent functions. We have therefore

$$\delta H_{ss'} = 2(n+1) \langle T_s^{(0)} \delta \psi T_{s'}^{(0)} \rangle$$

The projector on $\text{ker } P_n$ can be written in an obvious notation

$$P_n = \sum_{ss'} |T_s^{(0)}\rangle (H_n^{-1})_{ss'} \langle T_{s'}^{(0)}|$$

and

$$\begin{aligned} 2(n+1) \text{Tr } \delta \psi P_n &= 2(n+1) \sum_{ss'} \langle T_s^{(0)} \delta \psi T_{s'}^{(0)} \rangle (H_n^{-1})_{ss'} \\ &= \text{Tr } \delta H_n H_n^{-1} = \delta \text{Tr } \ln H_n = \delta \ln \det H_n \end{aligned} \quad (181)$$

For the projector $\tilde{P}_{n+1}$ on $\text{ker } P_{n+1}^\dagger$, the corresponding functions $\tilde{t}_u^{(0)}$ are such that $(g_{z\bar{z}})^{n+1} \tilde{t}_u^{(0)} = \tilde{\bar{t}}_u^{(0)}$ are analytic functions of $\bar{z}$. Choose these $\tilde{t}$'s independent of ψ . The same reasoning applied to the matrix

$$(\tilde{H}_n)_{uu'} = \langle \tilde{T}_u^{(0)}, \tilde{T}_{u'}^{(0)} \rangle \quad (182)$$

shows that

$$-2n \text{Tr } \delta \psi \tilde{P}_{n+1} = \delta \ln \det \tilde{H}_n \quad (183)$$

Combining equation (175) with (176ab) and the above results, we find that under a scale variation and using the reference metric $g = e^{2\psi} g_0$, with $R = e^{-2\psi} [R_0 - \Delta_0 \psi_0]$

$$\begin{aligned} & \delta \ln \det' (-\Delta_n^{(+)}) \\ &= -\frac{1}{\pi} \int d^2\alpha \sqrt{g_0} \delta \psi \left[\frac{e^{2\psi}}{\epsilon} + \frac{6n(n+1)+1}{3} (R_0 - \Delta_0 \psi) \right] \\ &+ \delta \ln \det H_n \det \tilde{H}_n \end{aligned} \quad (184)$$

$$\begin{aligned} & \ln \det' (-\Delta_n^{(+)}) \\ &= -\frac{1}{2\pi\epsilon} \int d^2\alpha \sqrt{g_0} e^{2\psi} \\ &- \frac{6n(n+1)+1}{6\pi} \int d^2\alpha \sqrt{g_0} \{ g_0^{ab} \partial_a \psi \partial_b \psi + 2R_0 \psi \} \\ &+ \ln \det H_n \det \tilde{H}_n + f_n \end{aligned} \quad (185)$$

with f_n a term independent of ψ (but still singular). We obtain as expected Liouville's action as part of the finite contribution and observe that the ψ -dependent singular terms agree with those in equation (179). Comparing both expressions, we conclude that

$$\ln \det' (-\Delta_n^{(+)})$$

$$\begin{aligned}
&= -\frac{1}{2\pi\epsilon} \int d^2\alpha \sqrt{g_0} e^{2\psi} + \ln \epsilon \left(\frac{1}{3}(3n+1)\chi - \dim \text{Ker} P_n \right) \\
&\quad - \frac{6n(n+1)+1}{6\pi} \int d^2\alpha \sqrt{g_0} [g_0^{ab} \partial_a \psi \partial_b \psi + 2R_0 \psi] \\
&\quad + \ln (\det H_n \det \tilde{H}_n) + F_n
\end{aligned} \tag{186}$$

With F_n a finite, ψ -independent contribution (therefore function only of the moduli, or equivalently of the reference metric). In the final analysis this is the hardest term to obtain, but also the most interesting one. In chapter 9 when discussing free fields on a torus we had an example of such an evaluation. The expression (186) is needed for $n=1$. That $-\Delta_1^{(+)}$ differs of $P_1^\dagger P_1$ by a factor 2 is of course irrelevant since any constant factor $2, \sqrt{2\pi} \dots$ can be absorbed in the definition of the integration measure. The analogous result for the scalar Laplacian, with the conventions on regularization of this section and the extra care paid to topology is left as an exercise. It reads

$$\begin{aligned}
\ln \det'(-\Delta) &= -\frac{1}{4\pi\epsilon} \int d^2\alpha \sqrt{g_0} e^{2\psi} + \left[\frac{1}{8}\chi - 1 \right] \ln \epsilon \\
&\quad - \frac{1}{12\pi} \int d^2\alpha \sqrt{g_0} [g_0^{ab} \partial_a \psi \partial_b \psi \\
&\quad + 2R_0 \psi] + \ln \int d^2\alpha \sqrt{g_0} e^{2\psi} + F_0
\end{aligned} \tag{187}$$

where F_0 is finite and only depends on the reference metric. This expression agrees with equation (134) (since $\ln Z = -\frac{1}{2} \ln \det'(-\Delta)$) in its dependence on the scale factor.

The final ingredient before putting everything together is the integration measure. The finite-dimensional measure on moduli was written $\mathcal{D}h_M$. We can now use an expansion of h in $\text{ker} P_1^\dagger$ over the zero modes of $P_1^\dagger$ to write

$$dh = \sum_u d\tau_u \tilde{T}_u \tag{188}$$

Therefore

$$\mathcal{D}h_M = \prod_u d\tau_u (\det \tilde{H}_1)^{\frac{1}{2}} \tag{189}$$

This term is absent for $g=0$, is a double integral for $g=1$, and otherwise the index u runs from 1 to $6(g-1)$.

We now collect the results from equations (186), for $n=1$ and (187), where we emphasize that F_1 and F_0 are functions of τ 's, to

write

$$\begin{aligned}
Z_\chi &= [\lambda_0 \epsilon^{\frac{1}{12}(8-d)}]^X \left(\int \mathcal{D}v^\perp \right) \int \mathcal{D}\psi \prod_{\mu=1}^d \left[\left\{ \frac{\epsilon}{\ln \int d^2\alpha \sqrt{g_0} e^{2\psi}} \right\}^{\frac{1}{2}} dC_0^\mu \right] \\
&\quad \times \prod_u d\tau_u \det \tilde{H}_1 \det H_1^{\frac{1}{2}} \epsilon^{-\frac{1}{2} \dim(\text{ker} P_1)} \\
&\quad \times \exp - \left\{ \frac{(26-d)}{24\pi} \int d^2\alpha \sqrt{g_0} [g_0^{ab} \partial_a \psi \partial_b \psi + 2R_0 \psi] + \right. \\
&\quad \left. \left(\mu_0^2 + \frac{1}{4\pi\epsilon} (1 - \frac{1}{2}d) \right) \int d^2\alpha \sqrt{g_0} e^{2\psi} + \frac{1}{2} [dF_0(\tau) - F_1(\tau)] \right\}
\end{aligned} \tag{190}$$

We have still been slightly careless in writing the measure. Indeed since we defined dimensionless determinants (cf. equation (173) where the dimensionless combination $\epsilon\lambda$ occurs) each original Gaussian integral ought to have been integrated with a measure $dx\epsilon^{-\frac{1}{2}}$. Thus $dC_0^\mu \rightarrow dC_0^\mu \epsilon^{-\frac{1}{2}}$, $(\prod_u d\tau_u) \rightarrow (\prod_u d\tau_u) \epsilon^{-\frac{1}{2} \dim(\text{ker} P_1)}$. Furthermore recall that $\int \mathcal{D}v^\perp$ is not the "volume" of the infinite-dimensional group but for $g=0,1$ (where the dimension of $\text{ker} P_1$ is nonvanishing) there is an additional piece missing, the infinite "volume" of the global conformal group. Thus one would like to factor out $\int \mathcal{D}v^\perp \times$ "volume of global transformations" and divide out by a renormalized factor (we take it to be one) times $\epsilon^{-\dim(\text{ker} P_1)}$ (where $\dim(\text{ker} P_1)$ is of course the dimension of this group). That a factor $\frac{1}{2}$ is missing in the exponent is a consequence

of the fact that the vanishing eigenvalues of $(P_1^\dagger P_1)^{\frac{1}{2}}$ would have dimension 1. In any case, these subtleties just imply that the factor $(\lambda_0 \epsilon^{\frac{1}{12}(8-d)})^X \epsilon^{-\frac{1}{2} \dim(\text{ker} P_1)}$ goes over into $(\lambda_0 \epsilon^{\frac{1}{12}(8-d)})^X \epsilon^{I_1/2}$ with $I_1 = 3\chi$. Altogether this means

$$(\lambda_0 \epsilon^{\frac{1}{12}(8-d)})^X \epsilon^{-\frac{1}{2} \dim(\text{ker} P_1)} \rightarrow (\lambda_0 \epsilon^{\frac{1}{12}(26-d)})^X$$

We are now in position to define the renormalized quantities

$$\begin{aligned}
\lambda &= \lambda_0 \epsilon^{\frac{1}{12}(26-d)} \\
\mu^2 &= \mu_0^2 + \frac{1}{4\pi\epsilon} (1 - \frac{1}{2}d)
\end{aligned} \tag{191}$$

The measure on the moduli space is

$$dM_\tau = \prod_u d\tau_u \det \tilde{H}_1 (\det H_1)^{\frac{1}{2}} \tag{192}$$

It would seem that the different pieces in this quantity depend on the scale ψ according to the variational equations (181) and (183),

and everything (including the metric g_0) depends on the moduli. For $g > 1$, $\det H_1 = 1$. After the cancellation of ϵ factors, the measure over translational zero modes can then be written

$$\prod_{\mu=1}^d \left\{ \frac{dC_0^\mu}{(\int d^2\alpha \sqrt{g_0} e^{2\psi})^{\frac{1}{2}}} \right\} = \prod_{\mu=1}^d dX_0^\mu \quad (193)$$

where X_0^μ is the center of mass (or translational) coordinate since the expansion of X^μ involved the zero mode contribution $C_0^\mu \times (\int d^2\alpha \sqrt{g_0} e^{2\psi})^{-\frac{1}{2}} + \dots$. Finally we rescale Z_x to Z_x^{Ren} by the "volume" of the full reparametrizations group and write

$$S(\tau) = \frac{1}{2} [dF_0(\tau) - F_1(\tau)] \quad (194)$$

the unknown, but probably most significant, part of the action.

The end result of this analysis is the expression (Polyakov, Friedan)

$$\begin{aligned} Z_x^{Ren} &= \lambda^x \int \mathcal{D}\psi \int \prod_1^d dX_0^\mu \int dM_\tau \exp -S_{tot} \\ S_{tot} &= S(\tau) + \frac{26-d}{24\pi} \int d^2\alpha \sqrt{g_0} [g_0^{ab} \partial_a \psi \partial_b \psi + 2R_0 \psi] \\ &\quad + \mu^2 \int d^2\alpha \sqrt{g_0} e^{2\psi} \end{aligned} \quad (195)$$

There are admittedly many points which ought to be clarified in the derivation. In particular finding the precise domain of integration in moduli space is a difficult question. Moreover the superposition of terms with different topologies might suffer from serious drawbacks. One point is obvious however. The proposed bosonic continuous theory for random surfaces does not make sense in dimension higher than 26, since the kinetic contribution of the ψ -field becomes negative. In dimension smaller than 26, there appears an essential degree of freedom, namely the ψ -field, due to the quantum scale anomaly. This has been called the Liouville theory. Dimension 26 is marginal. At the price of setting $\mu^2 = 0$, the ψ contribution factors out (possibly by rescaling correctly the measure dM_τ) and can be absorbed into a further renormalization of Z_x .

There are two types of observables that can be introduced to study the behaviour of these random surfaces. The nonlocal ones

such as occur in gauge theories are boundary loops. The whole derivation in the continuum can be carried over in the presence of prescribed boundaries with due care paid to reparametrization of the boundary itself. Depending on the type of boundary conditions, extra terms may occur, proportional to the area as well as the length of the boundary. In the context of particle physics, this yields the open string model. Other types of observables are local in the sense that the surfaces are required to go through fixed points $\{x^\mu\}$ in target space. This introduces in the path integral factors of the type $\delta^{(d)}(X - x) = \int (d^d p / (2\pi)^d) \exp(-ip \cdot x) \exp(ip \cdot X)$, the so-called vertex-operators, which allow to estimate correlations on the surface.

The continuous theory of surfaces is obviously involved and considerable sophisticated work (still in progress) is needed in order to extract information. Discretized versions partly amenable to numerical simulations are therefore of great interest in providing an alternative probe.

11.2.4 Discretized models

It would seem that a discretized approach avoids the question of reparametrization invariance. It will turn out that this is a little too optimistic.

It is worth remarking that a triangulation does not *a priori* approximate as uniformly a smooth surface as does a piecewise linear path with respect to a curve. A classical example is to approximate a vertical cylinder by triangles obtained by joining the vertices of horizontal regular polygons with n sides, centered on the axis at points of ordinate kb and inscribed in circles of radius a , each one rotated by an angle π/n with respect to the previous one (figure 5). For a layer of K such families of $2n$ triangles, the area of the cylinder is $2\pi abK$. On the other hand, each triangle has an area $\frac{1}{2} \cdot 2a \sin(\pi/n) \times \sqrt{b^2 + 4a^2 \sin^2(\pi/2n)}$, hence the ratio of triangular area to the continuous one is

$$\frac{n}{\pi} \sin \frac{\pi}{n} \left\{ 1 + \frac{4a^2}{b^2} \sin^2 \frac{\pi}{2n} \right\}^{\frac{1}{2}}$$

Let n go to infinity. If a^2/b^2 is kept fixed, the ratio of areas tends to unity. If however a^2/b^2 is allowed to vary with n , we can obtain

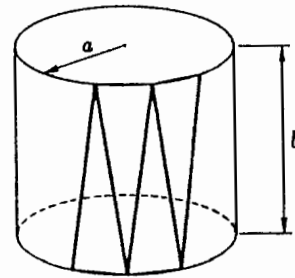


Fig. 5 Approximating a cylinder by a regular triangulation

any result (in this case larger to one) and even infinity, even though the area of each elementary triangle tends to 0.

Let us describe first the simplest approach. Fix the topology, a torus ($\chi = 0$) being the easiest one to treat, with a map in $\mathcal{R}^d$ for each point of a given triangulation. Take an action proportional to the embedded area, each triangle contributing a term ℓ_{ijk} , and integrate over all sites but one

$$Z = \int \prod' d^d X_i e^{-\beta S} \quad (196)$$

$$S = \sum_{(ijk)} \ell_{ijk}$$

We have set a short distance cutoff (analogous to $\varepsilon^{\frac{1}{2}}$ in the continuum) equal to unity, or absorbed it in the definition of β . We are interested in the case where the mesh of the triangulation becomes small. The most regular one is a regular triangular lattice in base space on a torus with N sites, $3N$ links and $2N$ triangles. On the base space, we have then a discrete translational symmetry. If we scale each $\mathbf{X}$ by λ , the partition function Z transforms according to

$$Z(\beta) = \lambda^{d(N-1)} Z(\lambda^2 \beta)$$

i.e.

$$\langle \beta S \rangle = 2\beta N \langle \ell_{ijk} \rangle = \frac{1}{2} d(N-1)$$

For N large,

$$\beta \langle \ell_{ijk} \rangle = \frac{1}{4} d \quad (197)$$

Hence a saddle point method is valid for large d . One introduces conjugate parameters for each link, enforcing the condition $(\mathbf{X}_i - \mathbf{X}_j)^2 = \ell_{ij}^2$. Recall that the area ℓ_{ijk} is given in terms of the ℓ_{ij} 's by equation (99). The partition function reads

$$Z = \int \prod' d^d X_i \int \prod_{(jk)} \left(\frac{d\lambda_{jk}}{2i\pi} d\ell_{jk} \right) \exp - \left\{ \sum_{(jk)} \lambda_{jk} \left[(\mathbf{X}_j - \mathbf{X}_k)^2 - \ell_{jk}^2 \right] + \beta S \right\} \quad (198)$$

Assuming a homogeneous mean field ($\lambda_{jk} = \lambda$ and $\ell_{jk} = \ell$ independent of the link), the saddle point condition leads to

$$Z \sim \exp - S_{eff}(\lambda, \ell) \quad (199)$$

$$S_{eff} = N \left[\frac{1}{2} d \ln \lambda - 3\lambda \ell^2 + \frac{1}{2} \sqrt{3} \beta \ell^2 \right]$$

with stationary conditions in ℓ and λ requiring

$$\lambda = \frac{d}{6\ell^2} \quad \beta \sqrt{3} \ell^2 = d \quad (200)$$

in agreement with equation (197). Therefore $S_{eff} = \frac{1}{2} N d \ln(\beta/2\sqrt{3})$. With these values for λ and ℓ , the model reduces to d -uncoupled Gaussian massless fields when investigating any observable depending on the coordinates $\mathbf{X}$. For instance, since $\langle (\mathbf{X}_j - \mathbf{X}_k)^2 \rangle = \ell^2$, the radius of gyration defined as

$$R_N^2 = \frac{1}{N} \left\langle \sum_j \mathbf{X}_j^2 \right\rangle \quad (201a)$$

assumes the value

$$R_N^2 = \frac{1}{\pi} \frac{\sqrt{3}}{4} \ell^2 \ln N = \frac{1}{\pi} \langle \text{elementary area} \rangle \ln N \quad (201b)$$

with a logarithmic behaviour arising from the corresponding infrared singularity of the two-dimensional propagator. This is to be contrasted with the linear behaviour in N for Brownian curves. In equation (201b) the prefactor $(1/\pi) \langle \text{elementary area} \rangle$ is specific to the triangulation, and of order d . It is corrected by terms of order $1/d$. In the present case, the leading correction in $(1 + (2/d) + \dots)$ has been observed in numerical simulations.

We conclude that a model based on a fixed topology and fixed triangulation, with a Boltzmann weight in $\exp(-\text{Area})$, yields as a consequence of equation (201) completely collapsed surfaces (of infinite Hausdorff dimension) and no simple modification of the action, as for instance by including a discretized Liouville term, seems to modify this conclusion. What is required to obtain a more sensible picture, closer to the continuous theory, is to re-examine the discrete analog of reparametrization invariance. This can only be to admit for a fixed genus all possible inequivalent triangulations $\mathcal{T}$ rather than pick a fixed one, which led us to the above catastrophic collapse. These triangulations appear as random lattices on an abstract base space of fixed topology. In order to simplify matters, and in analogy with the continuous case, a reference metric is assigned to such a triangulation with each link of equal length. The triangles are equilateral and therefore to each site is assigned an element of area $\sigma_i = \frac{1}{3}n_i$ in appropriate units (the equivalent of $\sqrt{g}$), where n_i is the number of its neighbours (equivalently the number of incident triangles or links). The total (intrinsic) area is equal to the total number of triangles N_2

$$A = \sum_i \sigma_i = N_2 \quad (202)$$

The curvature at a site in base space is given by the defect angle

$$\sigma_i R_i = 2\pi - \frac{1}{3}\pi n_i = \frac{1}{3}\pi n_i \left(\frac{6}{n_i} - 1 \right) \quad (203)$$

and is such that

$$\frac{1}{2\pi} \sum_i \sigma_i R_i = \chi \quad (204)$$

A term involving the square of the curvature would read

$$\sum_i \sigma_i R_i^2 = \sum_i \frac{1}{3}\pi^2 \frac{(6 - n_i)^2}{n_i} \quad (205)$$

To construct the action as was done in the continuum, one supplements a contribution proportional to the area (202) with a quadratic term in the embedding fields X_i using the base metric. For a fixed characteristic χ , one sums over all inequivalent triangulations $\mathcal{T}$ (annealed average) with arbitrary many triangles

and in addition one integrates over the X_i with a measure

$$\mathcal{D}X = \prod_i \left[\left(\frac{\sigma_i}{2\pi} \right)^{\frac{1}{2}d} d^d X_i \right] (2\pi)^{\frac{1}{2}d} \delta^d \left(\frac{1}{A} \sum_i \sigma_i X_i \right) \quad (206)$$

omitting for simplicity the integral over the center of mass. Thus

$$Z_X = \sum_{\mathcal{T}} \frac{1}{k(\mathcal{T})} e^{-\beta A} \int \mathcal{D}X \exp -\frac{1}{2} \sum_{(ij)} (X_i - X_j)^2 \quad (207)$$

The sum over triangulations replaces the sum over metrics, and $k(\mathcal{T})$ is a combinatorial factor equal to the order of the symmetry group of a given triangulation. An extra weight including R^2 -terms such as in equation (205) could be added to test for their irrelevance. The coefficient β is the analog of μ_0^2 in the preceding section or, equivalently, of a bare string tension.

The discrete operator occurring in the quadratic X -piece of the action has diagonal matrix elements $n_i = 3\sigma_i$ and off-diagonal ones zero or minus one if they correspond to a link. This is not quite the discrete scalar (minus) Laplacian obtained by multiplying both sides by the diagonal matrix $\sigma_i^{-\frac{1}{2}}$. Thus

$$\begin{aligned} (-\Delta)_{ii} &= 3 \\ (-\Delta)_{ij} &= \begin{cases} 0 \\ -(\sigma_i \sigma_j)^{-\frac{1}{2}} \end{cases} \end{aligned} \quad (208)$$

With the omission of the zero mode, one finds therefore that

$$Z_X = \sum_{\mathcal{T}} \frac{1}{k} e^{-\beta A} \left[\frac{\det'(-\Delta)}{A} \right]^{-\frac{1}{2}d} \quad (209)$$

where each piece depends on $\mathcal{T}$. That $\det'(-\Delta)$ is divided by the area A is a consequence of subtracting the center of mass integral. The dimension d appears as a parameter. Similarly one could consider local Green's functions tying the surfaces to a set of points $\{X_a\}$

$$\begin{aligned} G(\{X_a\}) &= \sum_{\mathcal{T}} \frac{1}{k(\mathcal{T})} e^{-\beta A} \int \mathcal{D}(X) \exp \left(-\frac{1}{2} \sum_{(ij)} (X_i - X_j)^2 \right) \\ &\quad \times \prod_a \left[\sum_j \sigma_j \delta(X_j - X_a) \right] \end{aligned} \quad (210)$$

appropriately normalized, assuming of course the lattice to contain more points than the set $\{X_a\}$.

For a fixed genus, the partition function Z_X is a power series in $e^{-\beta}$ since A is an integer and is expected to have a finite radius of convergence as some examples will soon show. Thus the model makes sense at least for $\beta > \beta_c$. At the critical point β_c , some singularities will occur, and one is interested as usual in the critical behaviour.

The one-point function G_1 is independent of X by translational invariance in target space and agrees with $-\partial_\beta Z_X$. The analogy with statistical physics is that the second derivative of Z_X is susceptibility (rather than a specific heat) and is proportional to the mean area. At criticality, one defines the exponent γ through

$$\frac{\partial^2}{\partial \beta^2} Z_X \sim \frac{1}{(\beta - \beta_c)^\gamma} \quad (211)$$

If γ is positive the mean area diverges, but remains bounded if $\gamma < 0$.

For β close to β_c , the two-point function is expected to scale as

$$G_2(X_{12}) \sim \frac{1}{|X_{12}|^{d-2+\eta}} e^{-m(\beta)|X_{12}|} \quad (212)$$

$$m(\beta) \sim (\beta - \beta_c)^\nu \quad (213)$$

with exponents η and ν . For a fixed triangulation one can define a radius of gyration

$$\langle X^2 \rangle_T = \frac{1}{A} \sum_{i,j} \sigma_i \sigma_j \langle (X_i - X_j)^2 \rangle$$

where the bracket means an average with respect to the X part of the measure. A generating function with respect to all triangulations reads

$$Z_{(X^2)} = \sum_T \frac{1}{k(T)} e^{-\beta A} \langle X^2 \rangle_T$$

If one performs the expansion

$$Z_X = \sum e^{-n\beta} a_n \quad Z_{(X^2)} = \sum e^{-n\beta} b_n \quad (214)$$

the ratio b_n/a_n is an average radius of gyration for fixed (intrinsic) area A (equal to n)

$$\langle X^2 \rangle_A = b_n/a_n \quad (215)$$

A possible definition of the Hausdorff dimension d_H is obtained in the limit $n \rightarrow \infty$ through

$$\langle X^2 \rangle_A \underset{A \rightarrow \infty}{\sim} A^{2/d_H} \quad (216)$$

Show that, if $m(\beta_c)$ vanishes, one expects the following scaling relations

$$\gamma = \nu(2 - \eta) \quad \nu = 1/d_H \quad (217)$$

For definiteness, we assume in the sequel planar topology (i.e. $\chi = 2$ with the plane compactified to a sphere to yield a compact surface). It was noticed by David that the model of randomly triangulated surfaces is equivalent, up to irrelevant terms, to a planar ϕ^3 theory with an exponential propagator, generalizing the discussion given in section 4 of chapter 10. Here ϕ is an $N \times N$ Hermitian matrix, and N tends to infinity. The ϕ -field theory is based on the (unbounded) action in Euclidean target space

$$S_\phi = \int d^d X \text{Tr} \phi e^A \phi + \frac{g}{\sqrt{N}} \text{Tr} \phi^3 \quad (218)$$

which is only used as a mean to generate a perturbative series. In the planar ($N \rightarrow \infty$) limit, the diagrams are dual to finite triangulations T of the plane (figure 6). One can modify the Lagrangian to eliminate tadpoles, and even self-energy insertions by including a second degree polynomial in ϕ . The ordinary combinatorial rules yield the factor $1/k(T)$ occurring in equation (207). The remaining contribution arises from the Feynman integral with a Gaussian configuration space propagator $\exp - (X_1 - X_2)^2$ between vertices of the planar diagrams.

One has to integrate over all X 's (but one). This is not recognized immediately as the X -part in the integral (207), the reason being that the X 's are indexed by sites of the planar graph rather than by those of the dual triangulation. Up to a factor, the Feynman integral here is of the form $(\det' K)^{-1/2}$, where $\det' K$ is any minor of the matrix K related to the incidence matrix on the planar graph, $K_{ab} = -1$ if a and b are neighbours, $K_{aa} =$ number of neighbours of vertex a , and otherwise zero. According to Kirchoff's theorem (chapter 7), $\det' K$ is the number of spanning trees on the planar graph, each one in one-to-one correspondence

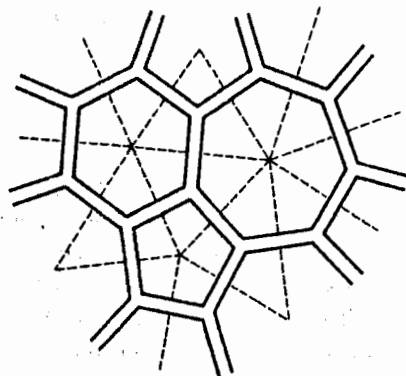


Fig. 6 Duality between planar diagrams (double lines) and triangulations (broken lines).

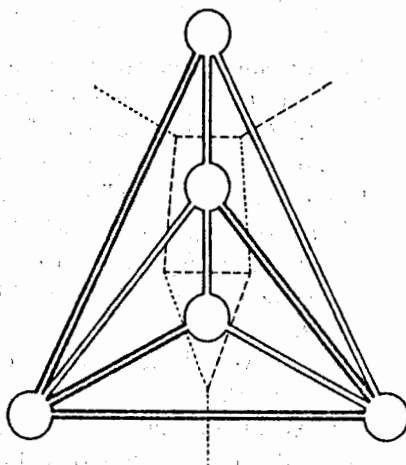


Fig. 7 An example of duality between spanning trees of planar graphs (indicated by heavy double or broken lines).

with a spanning tree on the dual lattice, i.e. the triangulation $\mathcal{T}$. This is a specific property of planar topology which would require modifications to be implemented on surfaces of higher genus. On figure 7, we show an example of such duality between spanning trees.

As a result, $(\det' K)^{-\frac{1}{2}d} = (\det' K_{\mathcal{T}})^{-\frac{1}{2}d}$ differs from the expected result $[(\det -\Delta)/A]^{-\frac{1}{2}d}$ only by factors arising in the

last case from the integration measure (206) i.e. $\prod_i \sigma_i^{\frac{1}{2}d} = \exp \frac{1}{2}d \sum_i \ln \sigma_i$. Since $\sigma_i = \frac{1}{3}n_i = 2[1 - \sigma_i R_i/\pi]$,

$$\prod_i \sigma_i^{\frac{1}{2}d} = \exp \left\{ \frac{1}{4}dA \ln 2 - (2 - \ln 2) \frac{1}{2}d\chi + \dots \right\} \quad (219)$$

using the properties of a planar triangulation ($2N_1 = 3N_2$,

$$N_0 - N_1 + N_2 \equiv \chi \equiv \frac{1}{2\pi} \sum_i \sigma_i R_i$$

incidentally, this shows that N_2 is even). The higher order terms involve higher powers of the curvature and are presumed irrelevant (or could be taken care of by suitably modifying the bare action in equation (207)). The first term on the r.h.s. in (219) is a shift in β and the second involves the expected topological invariant.

Thus from the point of view of renormalization, the planar ϕ^3 model defined through equation (218) is equivalent to the discretized Polyakov model. A renormalization of the propagator would be required to avoid ϕ^3 self-mass insertions (and triangulations with vertices having only two neighbours which could be omitted) and an added linear term in ϕ to eliminate tadpoles, where in the triangulations some vertices would be identified.

Consider the vacuum connected functional in the ϕ^3 version. The series in $e^{-\beta}$ is equivalent to an ordinary perturbative series in the coupling constant g for the planar theory where we have some experience in the convergence properties as well as singularities. A smooth extrapolation in the dimension d is afforded by equation (209). In particular, when $d = 0$, we can use the results of chapter 10 and in particular equation (10.290) which shows a finite radius of convergence in $g^2 = e^{-2\beta}$, with a behaviour

$$d=0 \quad Z_\chi = \frac{1}{2} \sum_{k=1}^{\infty} \frac{(72e^{-2\beta})^k}{(k+2)!} \frac{\Gamma(\frac{3}{2}k)}{\Gamma(\frac{1}{2}k+1)} \quad (220)$$

$$e^{-2\beta_c} = \frac{1}{108\sqrt{3}}$$

The general term in the series behaves as $e^{-2(\beta-\beta_c)k} k^{-\frac{1}{2}}$. As a result the singular part in Z_χ close to β_c is in $(\beta - \beta_c)^{\frac{1}{2}}$ and the exponent γ defined through the second derivative in β is negative

$$d=0 \quad \gamma = -\frac{1}{2} \quad (221)$$

analysis by O. Alvarez, *Nucl. Phys.* B216, 125 (1983) and the 1982 lectures of D. Friedan at Les Houches published in *Recent Advances in Field Theory and Statistical Physics*, J.-B. Zuber and R. Stora eds, North-Holland, Amsterdam (1984). The inclusion of extrinsic curvature terms in the action can be traced to the work of W. Helfrich, *Z. Naturforsch.* C28, 693 (1973) and *J. Physique* 46, 1263 (1985). See also A.M. Polyakov, *Nucl. Phys.* B268, 406 (1986).

For effective means to compute anomalies in the path integral approach see K. Fujikawa, *Phys. Rev.* D21, 2848 (1980), D23, 2262 (1981).

The study of interfaces between coexisting phases has a long history. Some recent representative works are D.J. Wallace and R.K.P. Zia, *Phys. Rev. Lett.* 43, 808 (1979), M.J. Lowe and D.J. Wallace, *Phys. Lett.* 93B, 433 (1980), *J. Phys.* A13, L381 (1980), F. David, *Phys. Lett.* 102B, 193 (1981).

Discrete models of surfaces were studied in the context of gauge theories by J.-M. Drouffe, G. Parisi and N. Surlas in a paper quoted in chapter 6, by D. Weingarten, *Phys. Lett.* 90B, 285 (1980), T. Eguchi and H. Kawai, *Phys. Lett.* 110B, 143 (1982), 114B, 247 (1982), by B. Durhuus, J. Fröhlich and T. Jonsson, *Nucl. Phys.* B225 [FS9], 185 (1983), *Phys. Lett.* 137B, 93 (1984). D.J. Gross, *Phys. Lett.* 138B, 185 (1984), A. Bilhoire, D.J. Gross and E. Marinari, *Phys. Lett.* 139B, 75 (1984) and B. Duplantier, *Phys. Lett.* 141B, 239 (1984) considered models involving a fixed triangulation, while a discretized version of Polyakov's model and its connection to the planar limit were investigated by F. David, *Nucl. Phys.* B257 (FS14), 543 (1985), V.A. Kazakov, I.K. Kostov and A.A. Migdal, *Phys. Lett.* 157B, 295 (1985) and J. Ambjørn, B. Durhuus and J. Fröhlich, *Nucl. Phys.* B257 (FS14), 433 (1985). Related computations were carried out in the continuous version by A.B. Zamolodchikov, *Phys. Lett.* 117B, 87 (1982) and J. Jurkiewicz and A. Krzywicki, *Phys. Lett.* 148B, 148 (1984). This list is by no means exhaustive.

As regards the literature on string (and superstring) field theory it is by now enormous. Many early papers are collected in reprint volumes such as *Dual Theory*, M. Jacob ed., North Holland, Amsterdam (1974 and 1984) and *Superstrings*, J.H. Schwarz ed., World Scientific, Singapore (1985). M.B. Green, J.H. Schwarz and E. Witten present a comprehensive view in the two volumes of *Superstring Theory*, Cambridge University Press (1987).

INDEX

A

A-D-E classification, 579, 754
 Abe, 157
 Abnormal dimension, 176
 Adjoint representation, 625
 Airy's equation, 653
 Alvarez, 780
 Amplitude ratios, 314
 Amputation, 239
 Anderson, 646
 Andrews, Baxter, Forrester, 588
 Annealed average, 802
 Anomalies, 268
 Anomalous dimension, 241, 274
 Anomaly, 516
 Anticommuting variables, 48
 Area law, 345
 Articulated diagrams, 250
 Ashkin-Teller model, 554, 594
 Asymptotic freedom, 329, 437, 727
 Atiyah-Singer index theorem, 793
 Avron, Seiler, 729
 Axial gauge, 333
 Background gauge, 340
 Baker, Nickel, Green, Meiron, 317
 Baker-Campbell-Hausdorff formula, 334

Bare parameters, 237

Baxter, 558

Baxter, Andrews, Forrester, 579

Becchi-Rouet-Stora (BRS), 647, 657, 717

Belavin, Polyakov, Zamolodchikov, 501, 578

Beltrami equation, 784

Bender, Wu, 447

Berezin, 48

Berezinskii, 200

Berlin-Kac, 140

Bernal, 739

Berry's phase, 734

Bethe approximation, 121

Bethe lattice, 419

Block spin method, 168

Bogoliubov, Parasiuk, Hepp, Zimmermann renormalization scheme, 264

Born amplitude, 282

Bosonization, 97, 722

Brownian motion, 1, 649, 778, 801

Brézin, Le Guillou, Zinn-Justin, 311

C

Caianiello, 95

Callan-Symanzik equation, 233, 273, 285, 309
 Canonical degree, 237
 Canonical dimension, 23, 241
 Cardy, 546, 601
 Cartan matrix, 622
 Cartan subalgebra, 619
 Cartan-Killing, 579, 619
 Casimir effect, 519, 564
 Casimir invariants, 583, 625, 635
 Casimir operator, 332
 Cauchy determinant, 84
 Cayley tree, 419
 Central charge, 515
 Character expansion, 374
 Cheeger, Müller, Schrader, 768
 Cheng-Wu, 88
 Chevalley, Serre, 622
 Christ, Friedberg, Lee, 738
 Classical Coulomb gas, 193
 Classical Heisenberg model, 26, 107
 Clausius-Mossotti formula, 109
 Clifford algebra, 52, 63
 Co-roots, 623
 Coleman, 197
 Configuration number, 439
 Confinement, 345
 Conformal anomaly, 773
 Conformal invariance, 502
 Conformal weights, 511
 Connected correlation functions, 239
 Contragredient form, 527
 Coordination number, 1
 Corrections to scaling laws, 301
 Correlation length, 40
 Coulomb gas, 203, 691
 Coulombic partition function, 593
 Coxeter numbers, 585, 624, 626
 Coxeter-Dynkin diagrams, 622
 Creutz, 473
 Critical amplitudes, 189

Critical exponent, 6
 Critical exponents, 125, 280
 Critical point, 40, 274
 Critical slowing down, 460, 471
 Cross ratio, 518
 Crossover exponent, 720
 Cumulant expansion, 77, 239
 Curie transition, 58

D

Darboux-Chritoffel kernel, 693
 David, 805
 Decimation, 176
 Deconfining transition, 354
 Dedekind's function, 528, 549
 611
 Descendant, 526
 Di Francesco, Saleur, Zuber, 577
 Diagram, 405
 Diffusion equation, 3
 Dimensional analysis, 236
 Dimensional regularization, 250
 Dimensional transmutation, 71
 Dirac equation, 93
 Dirac field, 721
 Dirac matrices, 394
 Dirichlet-Voronoi construction, 739
 Discrete Laplacian, 3
 Dominant weight, 635
 Dotsenko, 561, 563
 Dotsenko, Dotsenko, 716, 720
 728
 Dual complex, 348
 Dual Coxeter number, 624
 Dual group, 347
 Duality transformation, 180, 204
 Duality, 59, 345
 Dynamical dimension, 242
 Dyson, 310, 647, 665, 690

E

Effective potential, 156, 240, 249
 Eguchi, Ooguri, 578
 Elizur's theorem, 341
 Elliptic functions, 609
 Energy-momentum tensor, 507
 Enveloping algebra, 583
 ϵ -expansion, 235, 290, 311
 Equation of state, 166
 Equations of motion, 145
 Euclidean fields, 21
 Euler characteristic, 740, 765, 793
 Euler's formula, 705
 Euler's pentagonal identity, 549, 610
 Euler's relation, 409

F

Faddeev-Popov determinant, 785
 Faddeev-Popov ghost fields, 337, 436
 Feigin-Fuchs, 533
 Ferdinand, Fisher, 574
 Fermi levels, 714
 Feynman integrals, 246
 Feynman rules, 154, 244, 337
 Finite size effects, 478
 Fisher circles, 139
 Fisher, 233, 235, 290, 480
 Fisher-Gaunt, 157
 Fock-Bargmann space, 50, 676
 Fokker-Planck equation, 487, 649
 Free fields, 22
 Freudenthal formula, 626
 Friedan, 780, 798
 Friedan, Qiu, Shenker, 543, 592, 618, 638
 Frustrated partition function, 72, 86
 Frustration, 352, 588
 Fujikawa, 776

G

Gaudin, 699
 Gauge fields, 329
 Gauge fixing, 435
 Gauge invariance, 329
 Gauss series, 561
 Gauss sum, 612
 Gauss-Bonnet formula, 766
 Gaussian discrete model, 205
 Gaussian integrals, 22
 Gaussian model, 33, 550, 592
 Gell-Mann-Low formula, 522
 Generating functionals, 236
 Genus, 705
 Gepner, Qiu, 582
 Gepner, Witten, 581
 Glueball mass spectrum, 390
 Goddard, Kent, Olive, 636, 642
 Goldstone modes, 107, 118, 120, 127
 Goldstone's theorem, 298, 300, 435, 495
 Grading, 524
 Graph, 405
 Grassmannian integrals, 48
 Green function, 8
 Green functions, 238
 Gross-Neveu model, 490, 720
 Group characters, 331
 Group theoretical factors, 259

H

Haar measure, 427
 Hadron spectrum, 494
 Halfnian, 95
 Hall conductance, 729
 Halperin, 647
 Hard hexagon model, 558
 Hardy-Ramanujan formula, 548
 Harris criterion, 716, 720
 Hausdorff dimension, 6, 802, 808
 Heat bath algorithm, 460

Heat kernel, 332, 776, 790
 Hierarchical model, 192
 Higgs fields, 350
 High temperature expansion, 33, 421
 Highest weight vector, 526
 Hilbert-Einstein action, 768
 Hioe, Montroll, 715
 Hohenberg, 197, 219
 Homology and cohomology groups, 349
 Hopping parameter, 492
 Hopping probability, 759
 Huse, 588
 Hypergeometric equation, 561
 Hyperscaling hypothesis, 165

I

Importance sampling, 458
 Infrared fixed point, 171
 Interacting fields, 25
 Interface energy, 480
 Intersection probability, 16
 Irrelevant operators, 174, 307
 Ising model, 33, 58, 573, 605

J

Jordan-Wigner representation, 63
 Julia set, 193

K

Kac character formula, 640
 Kac determinant, 528, 532
 Kac table, 555, 559, 589
 Kac-Moody algebra, 526, 581, 619, 638
 Kadanoff, 168
 Kadanoff-Ceva, 87
 Kaufmann, 574
 Kazakov, Kostov, Migdal, 808

Killing field, 786
 Killing vector, 790, 794
 Kirchoff's theorem, 410, 805
 Kirkwood-Yvon theory, 110
 Kosterlitz-Thouless transition, 193, 200
 Kramers-Wannier duality, 61
 Kronecker's formula, 567
 Kubo's formula, 732

L

Lagrangian, 233
 Landau levels, 675
 Landau-Ginzburg criterion, 158, 166
 Langevin equation, 486
 Laplace transform, 153
 Laplace-Beltrami operator, 332
 Lattice fermions, 393
 Lattice gauge fields, 328
 Lattices, 43
 Laughlin, 688, 729
 Lee-Yang singularity, 125, 129, 131, 547, 566, 588
 Legendre transformation, 155, 159, 239, 415
 Level spacings, 606
 Level, 581
 Liapunov exponent, 666
 Lie algebra, 330, 579, 619
 Lifshitz tail, 654
 Liouville action, 773, 779, 795
 Liouville theory, 798
 Lipatov, 310, 447
 Localization length, 670
 Loop expansion, 246
 Loop group, 626
 Low temperature expansion, 430
 Lower critical dimension, 143

M

Ma, 456
 Majorana field, 93, 554, 721
 Majority rule, 183
 Marginal operators, 174, 218, 307
 Markov process, 5, 458
 McBryan-Spencer, 222
 McCoy-Wu, 78
 Mean field approximation, 108, 352
 Meijering, 743
 Mermin-Wagner theorem, 143, 193, 197, 219
 Metropolis algorithm, 462
 Microcanonical simulations, 465
 Migdal-Kadanoff approximation, 178
 Minimal coupling, 330
 Minimal models, 546
 Minimal subtraction scheme, 337
 Modular group, 567
 Modular invariance, 664
 moduli, 786
 Monodromy, 562
 Monte Carlo renormalization group, 456, 481
 Monte Carlo sweep, 460
 Multicritical points, 317
 Möbius transformation, 510

N

n -vector model, 25, 118, 259, 427, 436, 595
 Nambu-Goto action, 781
 Neveu-Schwarz and Ramond boundary conditions, 577, 616
 Nickel, 311
 Nielsen-Ninomiya theorem, 396
 Niemeijer-Van Leeuwen, 183
 Noether current, 508
 Noether's theorem, 97

nonlinear σ -model, 26, 107, 436, 769
 Number of loops, 247

O

One particle irreducible Green functions, 239
 One-loop correction, 155
 Onsager, 48, 58, 68, 75
 Orbifold, 554, 594
 Order-disorder variables, 87
 Oscillator wavefunctions, 698

P

Padé approximants, 448
 Painlevé equations, 98
 Parisi Sourlas, 686, 715
 Parisi, Wu Yong Shi, 486
 Pasquier, 579, 588, 595
 Path integral, 12
 Peierls' argument, 432
 Perimeter law, 345
 Perturbation theory, 242
 Perturbative renormalization, 264
 Peter-Weyl theorem, 427
 Pfaffian, 57
 Phenomenological renormalization, 224
 Pippard-Ginsberg, 165
 Planar approximation, 703
 Planar limit, 805
 Plaquette, 60, 331
 Poisson distribution, 665
 Poisson formula, 204, 580
 Polyakov loops, 469
 Polyakov, 738, 779, 780, 798
 Potts model, 176, 557, 595
 Power counting, 261
 Primary divergent diagrams, 262
 Primary field, 512
 Pseudofermion method, 494

Index

Q

Quadratic differential, 518
 Quantum Chromodynamics, 329, 490
 Quantum Hall effect, 675
 Quasiprimary field, 512
 Quenched approximation, 491

R

Radial quantization, 521
 Radius of gyration, 801, 804
 Random geometry, 738
 Random lattice, 664, 739
 Random matrices, 690
 Random number generator, 463
 Random potential, 647, 679
 Random surfaces, 764
 Random walk, 1
 Real space renormalization, 162, 168
 Regge, 768
 Relevant operators, 174
 Renormalization flow, 213, 270
 Renormalization group, 270
 Renormalized Green functions, 265
 Replica trick, 658
 Reproducing kernel, 684
 Resolvent, 680
 Riccati equation, 649
 Riemann's ζ function, 569
 Riemann-Christoffel symbol, 785
 Roots, 620
 Roughening transition, 91, 205, 385

S

Scaling regime, 89
 Schwarzian derivative, 517, 551
 Schwinger functions, 238
 Schwinger term, 631
 Second order phase transition, 70
 Self-avoiding walk, 31
 Semicircle law, 690, 708
 Shankar, 721, 729
 Simply laced Lie algebra, 579
 Sine-Gordon model, 206
 Six vertex model, 554
 Slater determinant, 698
 Solid on solid (SOS) model, 82, 205, 386
 Spherical model, 140
 Spin wave approximation, 197
 Spontaneous magnetization, 61, 72
 Spontaneous symmetry breaking, 112
 Staggered Kogut-Susskind fermions, 399
 Star-triangle transformation, 183
 Stochastic quantization, 486
 String tension, 314, 385
 Strong coupling expansion, 371
 Strong coupling series, 33
 Sugawara, 634
 Sum over paths, 9, 27
 Sum over surfaces, 780
 Super-renormalizability, 262
 Super-Schwarzian derivative, 615
 Superconformal algebra, 613
 Superficial convergence, 262
 Superfield, 683, 718
 Supersymmetry, 682
 Surface tension, 85

T

't Hooft loop, 352
 't Hooft, 703
 Tadpole graph, 244
 Teichmüller space, 786
 Thermalization time, 460
 Thermodynamic limit, 23
 t -series, 609
 Toeplitz matrix, 74
 Topological complex, 345
 Transfer matrix, 37, 62, 136
 Tree diagrams, 154, 248
 Triangulated manifold, 765
 Tricritical Ising model, 618

U

Universality, 33
 Upper critical dimension, 144

V

Vacuum diagrams, 244
 Vandermonde determinant, 77
 Verma module, 526
 Vertex functions, 239
 Vertex operator, 554
 Villain action, 204, 332
 Virasoro algebra, 521, 523
 Virasoro characters, 547
 Virial theorem, 663
 von Neumann-Wigner theorem, 734
 Vortices, 197, 201, 553

W

Ward identity, 145, 508, 513
 Ward-Slavnov identity, 337
 Watson's quintuple product, 611
 Wavefunction renormalization, 264
 Wegner, 679
 Weierstrass function, 612
 Weights, 624
 Weinberg's theorem, 274
 Wess-Zumino-Witten model, 626
 Weyl character formula, 625
 Weyl group, 622
 Weyl invariance, 782
 Wick ordering, 247, 536
 Wick's theorem, 24, 57, 94, 99, 243, 434
 Wigner, 647, 690
 Wilson fermions, 396
 Wilson loop, 344, 765
 Wilson, 233, 235, 290, 311, 328, 456, 482
 Wilson's action, 331
 Wulff construction, 89

X

XY-model, 193, 594

Y

Yang, 75
 Yang-Mills, 328, 325

Z

Z_n clock model, 474
 Zamolodchikov, 436, 587, 803
 ζ -regularization, 568